NATIONAL QUALITY FORUM

Moderator: Cost and Resource Use SDS Call
October 27, 2015
2:00 p.m. ET

OPERATOR: This is Conference #: 24987965

Ashlie Wilbon: Good afternoon, everyone. This is Ashlie Wilbon at NQF. I'm joined by Ann Phillips and Karen Johnson, my fellow colleagues here at NQF, and just probably a couple of other NQFers on the phone as well. So thank you all for joining us and welcome to the committee. This is our second call in the installment of our kind of post measure evaluation to further evaluate the sociodemographic and other socioeconomic factors, the SDS factors for risk-adjustment as a part of our trial. So, thank you all for joining us.

Just as a quick agenda review, we're going to do a quick roll call to see who's on. We'll do a very brief background review of how we got here, and then talk about the goals and the purpose of call and then kind of jump right into it.

So I did want to confirm that Nancy or Susannah Bernheim from Yale is on the phone to join us. Are you guys able to access the line? If you're talking, we can't hear you?

Nancy Kim: Yes, we can.

Ashlie Wilbon: Oh OK, Nancy?

Nancy Kim: Yes.

Ashlie Wilbon: OK, great. OK. I just wanted to make sure – yes, I can hear you now.

Nancy Kim: OK.

Susannah Bernheim: And Susannah is here also.

Ashlie Wilbon: Oh, OK, great. Thanks guys. I just want to make sure you were there. We're going to go right into our committee roll call real quick and then we'll jump in.

Brent Asplin?

Brent Asplin: Here.

Ashlie Wilbon: Lisa Latts?

Lisa Latts: I'm here.

Ashlie Wilbon: Ariel Bayewitz?

Ariel Bayewitz: Here.

Ashlie Wilbon: Larry Becker?

Larry Becker: Here.

Ashlie Wilbon: Mary Ann Clark? Cheryl Damberg?

Cheryl Damberg: Here.

Ashlie Wilbon: Jennifer Eames Huff? Nancy Garrett?

Nancy Garrett: Here.

Ashlie Wilbon: Andrea Gelzer? Stanley Hochberg? Martin Marciniak? Matthew McHugh?

Matthew McHugh: I'm here.

Ashlie Wilbon: James Naessens? Jack Needleman?

Jack Needleman: Here.

Ashlie Wilbon: Eugene Nelson? Janis Orlowski? Carolyn Pare? John Ratliff? Andrew Ryan? Joseph Stephansky? Lina Walker?

Lina Walker: Here.

Ashlie Wilbon: William Weintraub?

William Weintraub: Here.

Ashlie Wilbon: Herbert Wong?

Herbert Wong: Here.

Ashlie Wilbon: Dolores Yanagihara?

Dolores Yanagihara: Here.

Ashlie Wilbon: OK, great. Thanks, everyone, for joining us. Hopefully some others will pop in as we get going.

So just to give a brief overview of how we got here, so I think this is probably actually now our third call. We've got a couple of calls in the interim but before we got into the trial period. But this two-year trial period began in 2015 of this year, January of this year, and it was prompted by a report that was done by an expert panel that was convened to look at the impact of socioeconomic factors and other demographic variables, and the impact risk-adjustment. And the recommendations that came out of that report and that expert panel were that NQF should actually start considering measures that have included those factors in the risk-adjustment model if there is a conceptual reason for doing so.

So with that, we decided to try this variation on our policy for the next couple of years to see how the evaluation process would go and what that actually means for measure developers and for NQF. So within this trial period, we will be considering SDS factors as potential factors risk model. If there's a conceptual reason for doing so we would then ask the developers to do an empirical analysis if it's been decided that there is conceptual relationship to determine the contribution of those factors to the actual risk model.

So if in fact they do contribute and the developers decide to include those factors in the model, we then ask them to provide specifications for both risk, SDS adjusted measure and the non-SDS adjusted measure with stratification. And so because these measures were already endorsed as a part of a prior process, they got to board and the board realized that these measures were on the cost of this trial period beginning we're kind of retroactively or retrospectively going back to look at these measures and focusing primarily on the risk-adjustment as a part of the validity, the evaluation of the validity criterion.

So the three measures that are under review are the hospital-level, risk-standardized payment associated with the 30-day episode-of-care for AMI, heart failure and pneumonia which were developed by Yale and stewarded by CMS.

As a part of the conditions at the board put forth when they endorsed these measures, they wanted us to do a look back assessment of unintended consequences, consider these measures as a part of the SDS trial which we have done or in the process of doing and have it further examine the issue of attribution of which we are actually going to be actively doing, we launched the project yesterday which I encourage you all to look into and take a look at. It should be a lot of fun and excitement. We're expecting a lot of attention on that but it's really going to be allowing us to take a deeper dive on the issue of attribution and doing an environmental scan and some additional research on this issue and hopefully come up with some guidance for the field on how we can help with the measurement issues in that phase.

So again, these three measures, we have been talking with the developers when the recommendations came forward and we decided to review them in a Tuesday's approach across two webinars, the first of which we focused on the conceptual analysis that webinar was back in May and the committee had a very thoughtful discussion, made some recommendations which will review in just a second. And then today's webinar is the second of those two webinars which will be focused – look a little bit at the conceptual analysis as follow up to the previous webinar and the majority of this today's discussion will be

focused on the empirical analysis of which Yale has submitted a memo which was sent out to you guys last week. And so we will be discussing that in detail today and hopefully coming to some consensus around the validity criterion as well as the endorsement – recommendation for a continued endorsement.

So following today's call, we will be working to get all of the committee's deliberations written up. It will go out for 14 days comment periods, a little bit shorter because this is technically within our ad hoc review process which is a little bit slightly different process which goes a little bit faster than our normal consensus development process but it will be a 14-day comment period. We'll go to CSAC for review, the executive committee of the board and then a 30-day appeal process.

So any questions about that before I dive in? We're going to kind of jump in to reviewing what we've done already for the conceptual analysis and then we'll give Yale an opportunity to give us an overview of what they've done for the conceptual analysis as follow up to the previous call. So any questions before I hop in to that? OK, great.

So for our last call in May, we reviewed the conceptual analysis of the variables that Yale put forward that they identified as a part of their research and literature review and based on their access to variables for their own purposes of testing which were educational payment of – or income which they were going to use some census data and using the patient's ZIP code, Medicaid status as a proxy for low-income and insurance coverage and black or white race.

And so, we also asked the committee to determine whether or not there was a conceptual analysis – conceptual relationship and whether further empirical analysis was warranted and to give some more guidance on which variable should actually be pursued for further analysis and then any additional recommendations that the committee had on how the – what Yale should be considering in their conceptual – I'm sorry, their empirical analysis.

So as a quick summary, the three recommendations, the three key recommendations that came out to the first call were too broad in the conceptual model. The committee had quite a bit of discussion around the idea that the model, the original model that Yale presented was very medical-focused and they wanted to take a broader approach and consider other factors that may come up or be identified as a – of a more kind of public health or community-focused taking to account environmental and other patient factors. They also – the committee also made recommendations for the developer to do some additional literature review particularly on within and between hospital effects of race on hospital …

Male: We lost you.

Female: We can't hear you.

Female: Ashlie?

Female: We can't hear you, Ashlie.

Female: Oh Ashlie, I think a lot of us on the call can't hear you.

Female: Ann, just one moment.

Male: Is Yale developers on the call? This is Brent.

Susannah Bernheim: Susannah Bernheim is here. I think those of us who called in are on but the NQF folks dropped off.

(Mike): Hello, it's (Mike). Ashlie was just getting ready to hand it over for an overview. Lisa, would you agree with that may be we can have developers kind of walk through what they did while she's dialing back in?

Lisa Latts: Yes. It sounds like that would be a good use of time while they're dialing back in.

(Mike): Yes, if there's no objection.

Male: No objection.

Female: Keep going.

Male: All right.

Female: Yes, keep going. That's fine.

Male: Keep going, let's hand it over to Yale and keep going.

Female: Nancy, are you there? Nancy Kim?

OK. I – it was supposedly mostly any supporting mode for Nancy and so it's unfortunate if we walk through as well but I'm happy to talk to you guys to what we did. I don't have the materials in front of me but I will pull them up.

(Off-mike)

Nancy Kim: No, no, no. We got to call in.

Female: Oh Nancy, now, I can hear you again. You're on.

Nancy Kim: OK. Can you guys hear me? I'm sorry, we've been here the whole time, I don't know what's going on so let me just speak as loudly as I can.

Male: Yes. We got you.

Nancy Kim: OK. So essentially the committee wanted us to more explicitly include community, environmental, or patient factors, differentiate lack of patient resources, and lack of community, and reflect resources available for care within individual hospitals and there's up to a specific request that we change one title "patient behavior". It seems to blame the patients for their poor outcomes.

Does everybody have the materials in front of them? There should be an attachment of the new conceptual model and the memo.

Female: Yes.

Nancy Kim: So I'm speaking quickly just because we have so much to cover. But in response to those recommendations to broaden the model, to include more community factors, we made several modifications.

We changed the titles within pre-admission and post-discharge to capture the many patient and community factors that reflect differential SDS and can impact episode of care payments and we've categorized then as such. We made the specific change. We changed the title "patient behavior" to patient factors.

And also, given the chance to revisit the model, we chose to reorient the model to emphasize the potential pathways by which low SDS may be exerting influence on care provided by hospitals that we also put a layer on top of the blue arrows moving horizontally that sort of reflect the patient's community, pre-admission, post-discharge and in the bold yellow, the hospital influence and episode and care payments. So we really did try to broaden this conceptual model in response to the committee's comments.

I'm sure if we're going to get NQF but I'm happy to move on to part two which was our updated literature review and their (asset) because we do have so much empirical analysis to discuss …

Ashlie Wilbon: Yes. Hi, Nancy, this is Ashlie. I'm not sure what happened. We somehow got cut off. Apologies to everyone. I didn't realize we were cut off until someone from staff right in and told us so apologies for that. Go ahead and go over your literature review as well.

Nancy Kim: Is that OK?

Ashlie Wilbon: Yes, please.

Nancy Kim: I'm sorry Ashlie moving quickly through the model and lit reviews so they spend the bulk of a time on the empiric analysis but I think …

Ashlie Wilbon: Yes.

Nancy Kim: … inline with what you're intended as well. So …

Ashlie Wilbon: Oh great. Thank you.

Nancy Kim: Thank you so much to the committee and specifically Andy Ryan for sharing with us a number of articles in this arena. I think if you remember the Standing Committee had wanted us to expand the literature search specifically to look for other SDS variable that should be considered in risk-adjustment beyond income, race, et cetera. And it focus specifically on the within and between effects of race and/or income on hospital performance. So Andy, if I may call you that since there are a number of articles on our way and we also had several articles of interest that we have in our core repository.

Essentially in total, we evaluated 14 relevant articles. You can find them at Appendix two and based on the committee's recommendations, we organized the articles conceptually by their focus on the following categories, quote, unquote, within hospital papers, those examining differences in quality or outcome between populations of different SDS cared for in the same institution and between hospital differences that focused on between – that focused on papers examining whether minorities or patients of low SDS are cared for at lower quality hospitals based on their outcomes.

So in sum, the body of literature reveals an inconsistent and complex association of low SDS and health outcomes. Most studies, in fact all but one, used race as their independent variable with less attention to income or other measures of poverty such as Medicaid status. And the literature, in sum, demonstrates both within and between hospital differences and outcomes among different racial/ethnic groups that can be partially not wholly explained by the use of lower quality hospitals by minorities. So that was our sort of broad brushstroke take on our updated literature review.

Cheryl Damberg: So Nancy, this is Cheryl Damberg. Can I ask you a quick question?

Nancy Kim: Of course.

Cheryl Damberg: So when I looked at the set of articles that you reviewed, I was kind of struck by how few actually focused on payments or costs and it seemed like there was more emphasis on the quality side, do I have that right?

Nancy Kim: You do. I don't know Cheryl if you – I'm sorry, was there more to your question before I respond?

Cheryl Damberg: No but I just was wondering, you know, I mean I know that you guys, you know, were looking to do an exhaustive search but may be there's just not work that's been done on that space but I was kind of struck by that.

Nancy Kim: OK.

Cheryl Damberg: And then I just wanted to clarify for people because in particular the NQF SDS document talks about, you know, focusing on the within differences as opposed to the between but it seems like your literature review tried to look at both aspects, right?

Nancy Kim: Yes. So to your first point of (ask) the literature review, not including many papers with payment or cost as an outcome, that's true. If you remember back to webinar one, in webinar one, we have the very narrow focused literature review for the purpose of this actual trial period and we really focused on payment or cost or resources as an outcome and found there was a (jars) of high quality studies out there and in response to that committee asked us to broaden our research to include other outcomes or quality outcomes.

Cheryl Damberg: OK. Thanks for that reminder.

Nancy Kim: No problem. And the second, it's true also in that webinar number one, the discussion that we had based on the conceptual model which really focused on this indigenous, exogenous within and between hospital factors. And I'm referencing the NQF summary of our meeting, there's a memo, a meeting summary. So based on that we did – you're absolutely right, we looked at those within and between hospital differences.

Jack Needleman: Yes. This is Jack Needleman. And the – part of that discussion it may not been a large part, I'm looking at the Appendix one update, a conceptual model which I think is improved but there was some discussion about the fact that the standardized payment structure for computing the estimates actually hides whether there are actual differences in the resources available, whether the resource is used across within – across different hospitals in treating patients.

So we've talked about low-quality hospitals, the research showing quality hospitals or hospitals where African-Americans are more likely to be treated. But part of the issue in terms of the long-term issue, I'm not thinking about value, is whether those hospitals are low-quality because they have fewer resources available to them and the standardized payment structure hides that.

So I'm just wondering whether in the course of literature review, you were able to find anything that looked at the actual resources available to hospitals with higher proportions of minority patients or lower income patients relative to hospitals with a smaller share of those patients.

Nancy Kim: Thanks for the question. This is Nancy Kim. So just to remind everybody, our standardized payments are based on a DRG payment structure so we don't get itemized, you know, you had a lab, you had a center line, you had this, this, so we can never see that. So in that sense …

Jack Needleman: But you also can't see is the average actual cost per patient at the hospitals which is perhaps a probably better proxy of – a second proxy, not better, but a second proxy for how many resources are available at LA county USC versus Cedar Sinai.

Nancy Kim: We cannot see any cost and in fact we're measuring payments, so you're right on that. We don't see any of that stuff and that is true across all the payment measures because we are based on the inpatient side on a DRG. And I have not come across to anything. There's a number of issues with measuring costs as you – I'm sure you well know because costs are not, "standardized" and I don't mean that in a statistical way across hospitals, meaning, you know, if syringe may costs differently or they may charge differently for particular nursing care. It's all embedded in that DRG payment. But cost for an MRI for example, if one hospital will not be the same in a different hospitals. So there are number of issues with cost but your point is correct. Your larger point is correct and that we do not see any itemized costs.

Nancy Garrett: This is Nancy Garrett. I just wanted to throw in a question as well. At the bottom of page three of the summary of the literature review, it says, taking together, the body of literature reveals an inconsistent and complex

association of low SDS and health outcomes, and then just a little bit concern with that summary because when I looked at the literature that you guys looked at, eight of the 10 or it looks like may be 12 of the 14 had race/ethnicity dependent variable and I would not say that race is a proxy for low SDS. I mean obviously race and income and educational across together in complex ways but I think we need to be really careful about what exactly we're talking about and I wouldn't say if we're talking about race/ethnicity we can classify that as low or high SDS.

Nancy Kim: OK.

Nancy Garrett: I just think we need to be really careful there.

Nancy Kim: Thank you, Nancy, that's a good point.

Ashlie Wilbon: And I just wanted to follow on because Jack, I think your point is a really good one. But I also kind of want to either remind ourselves kind of what our goal in here with kind of exploring this SDS piece, which is trying to potentially fix what I call a mismeasurement problem versus trying to equalize resources available to care for different populations. So that starts to get at those between hospital differences.

Jack Needleman: I'm not sure I'm following that.

Ashlie Wilbon: Well so, let's think about how these measures are used and practice. So they're use for public reporting, they're use for financial incentive but if fundamentally there's a difference between hospitals and kind of the base resources that are available to manage patients and for them to invest in quality improvement activities, kind of fixing it up a margin, you know, with the performance metric and the incentive payment, that's not going to fix that problem.

Jack Needleman: Yes, that's correct. That's not going to fix that problem. It's just that I view it's critically important to acknowledge that problem exists in our AMI measurement, particularly when you put the cost against quality measures and that two dimensional graphing and grid structure, the standardized costs do

not reflect the actual resources that are necessarily available within a specific hospital to care for those patients.

Nancy Kim: Yes. And I would concur with that point and I think, you know, the challenge in trying to capture that is there's so many different cross subsidies so you can imagine at theaters, there's lots of private pay, you know, patients who are, you know, providing additional resources to the organization which don't at the county.

Jack Needleman: Right and we know from MedPACs analysis that hospitals with more private patients report a bigger Medicare shortfall which simply reflects the fact that they're investing a lot more per patient in their care. So all that is, you know, the measures we have in the measures we have I just think in terms of providing guidance to the public on how to interpret this providing guidance to CMS on how to interpret it, the distinction between what you see is a resource measure and standardized payments. And what may be the resource is available in a given community or a given hospital or given set of hospital plus post-admission people for actually caring for that – those patients may not be as highly correlated as the high correlation between including (SES) measures and not in (price).

Nancy Kim: Yes.

Lina Walker: So this is Lina. So can I – I'm just trying to understand what this means for somebody who is not immersed in (the) technical work, so are you saying then that between we're using standardized cost measures, we might be attributing high values to hospitals who might have fewer resources by virtue of where they are, in the low (SES) neighborhood or because of who they are treating. And so there is mismeasurement on that front and that we would need or we would hope that an (SES) adjustment would somehow be able to correct for that mismeasurement by virtue of using those standardized costs. Is that what you're trying to …

Jack Needleman: Yes.

Lina Walker: OK.

Jack Needleman: Roughly, Lina. But I heard that the directionality I heard in your question I think is the exact opposite of …

Lina Walker: Exact opposite.

Jack Needleman: … which is that, you know, when you get the standardized measures for the inpatient admission for an AMI, UCLA Harbor which is a county hospital or LA County which is a county hospital will get the same or standardized DRG payments as Cedar Sinai and Cedar Sinai is the wealthiest hospital in Los Angeles County. So, the issue there is when you do the member – people are being penalized for having higher cost or lower quality as measured by the quality measure in the two dimensional payments (draft) measure.

Lina Walker: OK.

Jack Needleman: But if part of the reason – what people have been concerned about is that part of the reason why the safety net hospitals are being – are not able to deliver as high quality care as measured by the quality measures is that they have fewer resources those differences are hidden when you use the standardized payment and that's been the concern.

Female: And this is the …

Male: Yes, so (inaudible). So let me amplify on that because I think, you know, it's really the key point because not only will that happen but you end up rewarding the riches hospitals and punishing the poorest, so it's a really dangerous business.

Nancy Garrett: And this is Nancy Garrett.

Female: So can I just …

Nancy Garrett: This is just to understand a little bit, but so – but really the way it shows them is only if you use the quality and the cost measures together because what you're saying is two hospitals that apparently looked comparable in the amount of resource used. One hospital might have lower quality just by virtue

of being the less wealthy hospital or the hospital on the less wealthy neighborhood. So …

Male: They're not actually worked that way. What may happen is you'll end up having more readmissions and more problems at the less wealthy hospital so it looks like you're spending more money so then penalize them.

Nancy Garrett: All right.

Jack Needleman: Yes, so you've got to double hit, right?

Nancy Garrett: Oh like the …

Jack Needleman: Like a …

Male: That's a real danger of exactly what we're doing here and, you know, I think the key thing here is, you know, can we get access in some of kind of realistic way so we don't end up hurting the most vulnerable.

Nancy Kim: Ashlie, it's Nancy Kim, can I say something?

Ashlie Wilbon: Yes, please.

Nancy Kim: Just to remind everybody, the payments are based on an inpatient DRG but the DRG reflects those comorbidity and intensity of care so it's not that every AMI has the same DRG, in fact there are many DRGs for an AMI. It's based on whether or not you have a CABG or cath et cetera and those procedure intense, DRGs will bump up you so you're more expensive than an AMI that may be got TPA. So it's not that every AMI has the same DRG or every heart failure has the DRG or every pneumonia has the same DRG.

Jack Needleman: But Nancy, we all understand that …

(Crosstalk)

Male: … doesn't get at the issue that's been raised.

Cheryl Damberg: Can I add one more thing? This is Cheryl. So I think all of these are valid points but I just kind of want to, again, try to get some clarity because, I think we're being asked to consider whether this measure should be adjusted for some set of socioeconomic, you know, characteristics of patients. And I agree that the standardized payment doesn't adequately represent the full set of resources available to a hospital but it seems to me that if, you know, one word to adjust the performance measure that really isn't going to fix the issue of kind of equalizing payments across hospitals but that seems to me that that's more related to like a DRG effects so like if there an additional payment, you know, if you're on a hospital that have fewer resources overall as opposed to the performance measure fix, so I'm …

Nancy Garrett: And this is …

Cheryl Damberg: … trying to get some clarity. That would be helpful.

Nancy Garrett: This is Nancy Garrett. So I just – I think that's a great point Cheryl and as a member of the Risk Adjustment Committee, this is actually something that we put in our final recommendations. It's gotten a little bit loss because of the main recommendation of we should be doing this risk-adjustment when it conceptually makes sense on those empirical evidence but one of the recommendations or findings on that report was that risk-adjustment alone even we had perfect data and we did that really well which we're not – we're far from that obviously.

Risk adjustment for SDS factor does not going to address the fundamental issue that people are raising here of equalizing payment across hospitals to really get the resources so that we can reduce disparities in a real way so that's really more a payment policy issue and the risk-adjustment is not going to fix that. So it doesn't mean we don't do it – try to do it as well as we can but we still have to remember we're going to have to do some other things to make sure that we don't end up in a downward spiral in taking the resources away from the patients we need at the most.

Jack Needleman: Yes. And I got no problem, I recognize the limitations of the standardized payment structure in terms of dealing with the real resource issues but since

we're called the Resource Committee, Resource Use Committee, I think the limitations of those measures need to be at the measures we've got as a measure of resource use and resource is available just need to be – we need to be constantly reminding folks of that limitation of the way we're measuring even if we accept the measure that is based upon those limited elements.

Ashlie Wilbon: Thanks, Jack. This is – thank you everyone. That's actually really helpful. I think it helps to have – the committee to have time to kind of reengage about why this is so important and why this has been such as a tricky issue and ongoing issue but I think Nancy made some really good points and I think in the interest of time, I'd like to move on to the empirical analysis and I think that based on what Yale has done so far in terms of additional literature review and updating their conceptual model that, you know, we can know in the summary that there are clearly some limitations and gap to the literature on this issue but that given, you know, the specificity of the topic and even go and doing a broader search that there are, you know, there are some gaps in that. It has sufficiently I take to lead the foundation for what the committee needs to determine that.

You know that we made the right decision in terms of moving forward with the suggesting further empirical analysis. So everyone is comfortable with that, I'd like to move forward to the empirical analysis piece which is the really the meat of this webinar and hopefully, we can come to some consensus with that analysis and where the committee is comfortable in making some recommendations at the end.

So I'd like to just hop forward to the next slide to just kind of give a few points of guidance on what this discussion is centered on. It's really to review and discuss the empirical analysis of the risk-adjustment approach for decision to include or not include the SDS adjustment in the measure based on the analysis that they provided and ultimately at the end of this call, all of the committee members should have receive an e-mail from me with the links with SurveyMonkey and also in the memo that you received from me where you should be able to vote for each measure on the validity criterion. There'll be more guidance on that towards the end of the webinar as well as making an

endorsement recommendation for continued endorsement of the measure specified, continue endorsement or should we endorse the measures.

So with that, I just want to give a few highlights in terms of guidance that we have provided to developers in general about what we're looking for as a part of this empirical analysis. So, we're looking for analysis and interpretation that resulting on a decision to include or not include the factor in the measure.

So performance score, a risk model performance of the model with and without the SDS factors, interpretation of the results and if the developers decided to include the SDS factor in their measure, we're asking that they do update the reliability and validity testing. In this case, Yale, based on their analysis, has decided not to include those factors in their measure so that information was not submitted which is fine but the committee will be voting on the measures as they were specified basically at the initial evaluation that you guys did last year, that last year when you recommended the measures which is to have the measures continue endorsement as they are without the additional SDS factors that they have looked at in their empirical analysis to be included in the model.

So in terms of where we landed at the end of the call in May, where the committee landed in terms of their recommendations on the variables in the empirical analysis. In terms of race, the committee wanted the developers to review the data and to consider including other variables beyond just the black variable within the race codes.

Income – for income and educational attainment, the committee ultimately decided that the data that Yale had access to at the time which was five-digit ZIP code was not sufficient and they wanted a little more specificity and that it was – they would rather the developers wait until they had nine-digit ZIP code data to perform that additional analysis for that variable.

For the Medicaid, the eligibility status, the committee decided that they would like to see empirical analysis on the Medicaid status variable but only in combination with the low-income subsidy data as a proxy for the insurance status and income.

So with that, I would like to turn it back over to Yale to give everyone an overview of their empirical analysis and ultimately what supports their decision to not include those variables in the risk models for those measures.

Nancy Kim: Thanks, Ashlie. It's Nancy Kim again. First, I just want to acknowledge Cheryl Damberg, thank you so much for sharing your knowledge and experience to the low-income supplement variable. We did examine the use of both the Dual Status Indicator and the cost sharing group of which the low-income supplement is a subset, LIS, low-income supplement is a subset of the cost sharing group, a second variable in addition to the Dual Status Code that we ended up using.

We ended up using the Dual Status Indicator alone based on a technical guidance from the chronic condition warehouse data which was the data that we received because it was much more restrictive but included the poorest of poor when we took a very close look at what Cheryl sent to us it was a little more generous. It included also subset of the Dual Status Code, but it included folks that were unknown in relevant Medicaid AB but no data et cetera and the low-income supplement also did that.

So we chose to restrict it to you the smaller subset that we're comfortable with that we did however, we looked at the overlap between the RAND method and our method and there was really a 97, 98, 90 percent overlap across conditions so they aren't much different. But based on the restrictive the poorest of poor nature of our Dual Status Code and the technical guidance from CCW and the literature, we decided to go with what we assume. Thank you so much for the committee for sharing those data with us.

OK, now to the meat. Ashlie, (help) me to go through a condition by condition everything we did?

Ashlie Wilbon: Yes, let's go measure by measure.

Nancy Kim: OK. So, if people have the memo in front of them, I'm going to really start going into the meat at page seven. Before we do that, I just want to share with the committee, we performed a number, a lot of analysis and to facilitate discussion of the results. We're just providing a summary of the key findings.

All the analyses are there. They're in the appendix. We can provide more data if you need it. But we're beginning with the hospital-level of risk-standardized payment. You're going to hear me refer to that as RSP because that's actually the central question posed by NQF, whether the measure outcome or RSP is affected by the addition of the SDS variables.

So we provide everything but we're talking about the hospital-level results first and overall with them that there's a minimal association between race and Medicaid status in the episode of care payment across conditions. I'm going to walk you through condition by condition. So I'm on page seven and these are the – this is the AMI (episodic) care payment first.

Just generally, we had about 380,000 index admissions. Based on that, about eight percent were black, 19 percent were categorized as Medicaid admissions and 3.3 were – percent were categorized as both. So the take home point there is that there was not a lot of overlap between black race and Medicaid status. They're not proxies for one other.

When we look at the distribution of percent black and Medicaid index admissions in the measure cohort across hospitals, we focus on the hospitals with at least 25 cases in (sense) that we report on in hospital compare and the median percent of black admissions was 3.4 percent. The median percent of Medicaid index is 17 percent.

I'm going to move on now to hospital-level results and feel free to interrupt me. I'm just moving quickly for the sake of time.

Focusing on table four, when we look at the risk-standardized payments calculated with the current model versus the current model with either the addition of black race alone or the addition of Medicaid race alone or black and Medicaid together, we really found no difference in the distribution of RSPs across hospitals.

You can see in table four, we had 2300 hospitals and compared to the current model which is on the top, the most full model current plus black plus Medicaid which is in the bottom row really looks no difference. Minimum

payments are still about $13,700 or so. The median payment is $21,640 or so et cetera.

We've almost – we then looked at the percent change in risk-standardized payments calculated with the addition of the SDS variables compared to the current model and that's table five. And this is a little more complicated so I'm just going to walk you to the columns.

The first column shows the current model plus black alone, second row shows current plus Medicaid alone, the third row shows the current plus both black and Medicaid. That second column represents the percent change in risk-standardized payment when these variables are added. Third column is the number of hospitals affected by that percent change and the fourth column is the percent of hospitals affected by that change.

And what we saw was that focusing on that bottom row, the current plus black plus Medicaid, the most full SDS variables included, we really found that 97 percent of hospitals is changed by less than one percent and either direction, either got more expensive or less expensive. So again, focusing on that last row, the first, I guess header is a change of negative one percent bureau hospitals fall in to that category.

The second row there and that subset rows is a change of zero to one percent cheaper and about 71 percent of hospitals got cheaper when black and Medicaid were added to the AMI payment model. That third row is zero to one percent more expensive and about 26 percent of hospitals got more expensive by only by less than one percent and that fourth column that one plus is indicating hospitals that change by one percent or greater in a more expensive direction and only 3.4 percent of hospitals increase their risk-standardized payments by one percent or greater. No hospitals got cheaper by greater than one percent but most hospitals, 97 percent of hospital got less than of one percent change in either direction, cheaper or most expensive.

Are there any questions about that because we're going to see that table again for heart failure and pneumonia?

OK. Table six is …

Female: I just …

Nancy Kim: Yes.

Female: I just find that shocking with the numbers are so low. I mean is that different that what you would have expected?

Nancy Kim: You know based on the bivariate which we're going to get to, obviously, we worked in step-wise fashioned, we did bivariate first representing the hospital-level results first, if we're focused on a hospital-level question. Based on what we saw on the bivariate then throwing SDS variables into the multivariable model, we were not really that shocked and I can go on and show you those big data as well. We're going to get there next.

Joe Stephansky: Nancy, it's Joe.

Nancy Kim: Yes.

Joe Stephansky: So since these are 30-day models, what proportion of the standardized payment is the hospitalization itself?

Nancy Kim: For AMI, it's about 70 percent. 70 percent is at the index payment and 30 percent is opposed to discharge payment. It changes for heart failure and pneumonia.

Joe Stephansky: OK. So roughly 14,000 and the 21,000 is …

Nancy Kim: Yes.

Joe Stephansky: OK.

Nancy Kim: Yes.

Joe Stephansky: So did you look at the non-hospital component and whether it was changed by the inclusion of the SDS variables?

Nancy Kim: No, we didn't.

Joe Stephansky: OK.

John Ratliff: Can I ask a quick question please? This is John Ratliff.

Nancy Kim: Sure.

John Ratliff: Did you happen to specifically look at readmit rates because I agree with one of other earlier comments that it's literally stunning the rates of readmission in AMI and CHF exacerbations are not affected by socioeconomic status …

Nancy Kim: No, we didn't look at any of the post-discharge settings by sociodemographic status. We only looked at that.

John Ratliff: But readmissions would be captured by an episode of care payment correct?

Nancy Kim: Yes, they would be captured in the post-discharge portion of the episode of care payment so we have the episode of care payment equals index plus what we called post-acute and the readmissions, we considered part of the post-acute care after discharge.

John Ratliff: And the readmission regardless of facility would be captured within the episode of care model?

Nancy Kim: Yes, that's true. And I think the question before you were saying do we look at SDS for the post-acute care settings including readmission and rehab and all that, we did not look at the association of SDS and post-acute payments. We only look at the association of SDS and total episode of care payments. Does that answer your questions?

Susannah Bernheim: Nancy, can I have one thing? This is Susannah Bernheim from the Yale team also. I mean I just – I understand the surprise on the committee and I think one of the things that we have looked very early on development of this measure was sort of how hospitals did prior to the risk standardization on the different chunks of the payment and I think part of what we talked about in our first webinar is that SDS can play out in large different ways

So in AMI, it may be that there are fewer procedures being done. Certainly, we know that in race, on the black patient, and that brings cost – indigent cost

down and may be that there's higher needs in the post-discharge period which brings cost up, some of that may be representing better care, some worst care. I think what happened is that the faster to the extends that they have any influence payment going on a lot of different directions and in some we aren't seeing much difference between the hospital based on their composition and (inaudible) when you add the risk-adjustments.

Nancy Kim: OK, should I move on Ashlie with the bivariate?

Ashlie Wilbon: Yes, please.

Nancy Kim: OK. So I'm on page nine. I hope people are following on their packet. I'm on table seven. So those – that was the bottom line of the hospital-level result which I view as essential question of the committee, how do – how does the addition of SDS variables affect hospital-level risk-standardized payments.

Of course we did bivariate because we build these models in step-wise fashion as per NQF guidance, we treat these SDS variables like other clinical entities although there are some debates about that. But when we look at the bivariate looking at black alone compared to our total payment, when we just look at the bivariate, when we add the black alone, there was a one percent higher total payment compared to non-blacks. When we did the same with Medicaid so how is the Medicaid status affects total payments. No other variables in the model. There was no change in the total payment compared to the non-Medicaid patients.

When we added black to the currents model, we found that payments were six percent less expensive than non-blacks. When we added Medicaid alone so without black to the current model, payments were two percent less expensive than non-Medicaid patients and when we added both Medicaid and black to the current model individual results tell that is to say that black admissions were six percent less expensive than non-black. Medicaid admissions were two percent less expensive than non-Medicaid patients.

So for AMI, when we add black and Medicaid to the current model, black and Medicaid admissions are less expensive than non-black and non-Medicaid admissions.

We also looked at model performance. This is table nine on page 10.

Jack Needleman: Just to clarify, that could – just building on the prior – that can reflect either post discharge differences in resources used or differences in the DRG mixed between blacks and the rest of the other population?

Nancy Kim: That's correct. They could occur anytime in the episode of care. For AMI specifically as (Anna) said, we have empiric analysis within our data to suggest the blacks in both Medicaid – both blacks and Medicaid patients receive fewer procedures then non-blacks and non-Medicaid patients but you're absolutely right. Those payments could be decreased at anywhere in the episode of care payment.

So again, for the sake of time, I'm going to just move to table nine which is evaluating the model performance with the addition of the SDS variables. So if you're looking at table nine, you'll see our first column is the diagnostics that we used predictive ratios and the Quasi-R squared. The second column is our current model and then you have – moving horizontal, you can this a lot, there's the current plus black alone then current plus Medicaid alone and then the far most right last column, you'll see the most full model, our current model indicates.

And the take home point is really that would be addition of black in a Medicaid variable. There was no change in model performance.

Male: The other take home of course is that the overall performance of the model is, isn't that great which, you know, we brought up no this – on the – with this group before, I think it's the biggest worry.

Nancy Kim: Should I move on Ashlie to the residual analysis?

Ashlie Wilbon: Yes, please.

Nancy Kim: OK. So, we also want to assess whether the addition of the SDS variables produced a better model fit that is improved the predicted total payment for sub-groups of patients. So we compared the residuals from the current model

to the residuals after the SDS variables were added. And what we found was that the addition of sociodemographic SDS variables residuals improved for all subgroups so they improved for blacks, they improved for non-blacks, they improved for Medicaid, they improved for non-Medicaid when we parsed them out into four different groups, but they did improve more for black and Medicaid subgroups.

We also tested the collinearity of our SDS variables with each other and they were not collinear with each other or the other variables in the current model.

Should I pause, Ashlie, for any other questions?

Ashlie Wilbon: Yes. So that was the end of the AMI measure, right?

Nancy Kim: Yes.

Ashlie Wilbon: So is there anything else about the AMI measure you wanted to summarize? I think I'd like to take a pause here just for the committee to discuss this measure. Since there are some similarities for the other measures, it might be useful to just kind of talk through any questions or issues and then see how does or does not apply to the other two measures or so. If we can pause here, that would be great.

Are there any questions from the committee member or clarifications around the analysis or any other kind of questions and clarifications needed to understand why they decided not to include, ultimately not to include those factors in the model?

Cheryl Damberg: This is Cheryl. I have a question on the model itself. So I realized there is a fair amount of debate over there about how to do this but I was curious why you included a hospital random affect versus a fixed affect in the model because in doing so it's going to have more correction effect if you will on the small hospital as opposed to, you know, in closer to the mean and I don't know whether your team discussed, you know, pros and cons of a fixed versus a random affect in the model.

Nancy Kim: Yes. Thanks for that question. It's Nancy. So the random affect is the way (CORE) approaches all of its measure whether it's episode of care payment or mortality or hip/knee complication so that is our standard approach and yes, it was hotly debated a few years ago when we were developing the approach, but that is the current approach of our episode of care payment model without the (inaudible), our standard approach to our nonpayment measures as well including the publicly reported mortality readmission and complications measure. If there is – I'm sitting next to my senior analyst, (Leslie Ott), who conducted all these analysis probably has a lovelier technical explanation but that is the reason we used the hospitals (or) random effects.

Cheryl Damberg: But it ultimately is shrinking small hospitals in closer to the mean because there are estimates that are likely less reliable. So that's just – I want people to be clear on what the effect is of doing that. Because I think …

(Crosstalk)

Male: I have experience with this kind of things in other settings. I mean that is exactly what happens, you get that shrinkage whether it is technically the correct way to do it. And the problem with that is that it works in general for small hospitals but where you can fall down is in applying it to anyone specific hospitals but this is what you can do.

Nancy Kim: And we only public the report on hospitals just greater than 25 events with AMI.

Male: But even so a 25 event is a very small number of events.

Nancy Kim: But it is …

Male: To make any kind of generalization and that's why you can say something about small hospitals but for anyone small hospital to say that with this model, you know what's going on it's a questionable validity.

Male: Yes, but just in terms of the imposition of penalties or bonus payments for lower cost, higher cost, it also means by pulling the smaller hospitals in they are less likely to be in the tails that (inaudible) penalize or bonuses.

Male: Yes. That's true and that's the good part. In that part is you don't know what's really going on with them but at least you're not doing things to either extreme.

Lina Walker: And this is Lina. And just so my clarification, what I'm hearing is that this would be a better approach than a fixed affect model because that wouldn't adequately capture variations in the hospital affects.

Nancy Kim: This is course decision with our team of statistical analyst based for every measure we do so that is the reason we …

Jennifer Eames Huff: Yes, this is Jennifer. I think it's just what people are talking about being aware of the implications of that decision. I think from a public reporting in terms of late consumers seeing the information it shows providers that are average which they could be misclassified, you know, it could be the provider could be actually more at the tails than they are then what this model will show but on the other hand then it's sort of the balance of the reliability and validity of the results.

Brent Asplin: Nancy, if there is not any objection, it will be great if they did go to the pneumonia measure next because even if they ended up in the same place results were a little different, if I'm recalling correctly from the memo, I don't have it open in front of me.

Nancy Kim: Sure. This is Nancy Kim.

Ashlie Wilbon: Sure.

Nancy Kim: Ashlie, should I go through heart failure or what would you?

Ashlie Wilbon: If folks – Brent, is that – you want to go to the pneumonia so that we could see the differences …

Brent Asplin: It's purely selfish. I got to drop off and I'm just curious did they end up at the same …

Ashlie Wilbon: OK.

Brent Asplin: … final recommendation.

Nancy Kim: We needed up with the same final recommendation.

Brent Asplin: Yes, but the results look a little different. Heart failure looks almost exactly the same from the current model.

Nancy Kim: So heart failure basically, at the hospital-level, all three measures there was no impact on the overall distribution or same hospital risk-standardized payment when SDS was added and it looked very similar. The only thing with heart failures and the multi-variable analysis black admissions were less expensive than non-black but Medicaid admissions were lower expensive than non-Medicaid patients so that's slightly different from AMI where they were both less expensive.

And in pneumonia, it's the opposite. Black and Medicaid, well first of all again, at the hospital-level, there was no impact on the overall distribution or same hospital risk-standardized when SDS variables were added to the hospital-level model. When looked at the patient-level model, after a black end, SDS were added to the current multivariable model, black and Medicaid patients were more expensive than non-black and non-Medicaid patients so that is different.

So in AMI they were – once we added them into the multivariable model at the patient-level both black and Medicaid look cheaper in heart failure. In pneumonia they were both more expensive and in heart failure, black was cheaper and Medicaid more expensive.

With the hospital level, we found the same thing. There was no impact on the overall distribution of hospital risk-standardized payments or on the same hospital risk-standardized payment when SDS variables were added.

Ashlie, should I walk through item by item?

Ashlie Wilbon: I don't necessarily – unless there is a committee – a strong committee, you know, suggestion or desire for that, I don't think that is necessary. I think you

are – your overview in terms of distinguishing the differences from the AMI measure were very helpful. So maybe I'll open it up to the committee to ask any questions about either measures individually or the differences between the measures and, you know, whether that we've any additional questions or clarifications on the decisions to ultimately include or not include it in the model but I'll leave it to the committee at this point to decide how they'd like to discuss it going forward.

Male: Can somebody remind me while I'm looking at table 21, the percentiles go from less expensive to most expensive so the 10 percentile is less and the 99 is the most expensive?

Nancy Kim: Yes. So the table 21 is on page 17 and this is regarding the pneumonia payments and this is regarding the distribution of the percent black and Medicaid index admissions and the pneumonia measure cohort across hospitals. So you're right, the minimum – so the way to read this is let's take the percent black index admissions hospitals with at least 25 cases at third column. The minimum number of black admissions one hospital is zero percent so some hospitals are zero percent black admissions. The maximum number of black admissions at some hospitals is 100 percent. The median number of black admissions is about two percent at any given whatever hospital. Does that answer your question?

Male: Yes. And first, so what we're seeing is a very clear gradient that as the percentage of blacks or percentage of Medicaid increase the cost in the hospital's increase. Is that the way that we …

Nancy Kim: No. Thank you for your question. This is Nancy Kim again. Table 21 is a descriptive table to merely represent the distribution of the person who is a black patient going to any – going to the array of hospitals within our cohort. So basically, what this is showing you is that, well, there's a large distribution of – and now I'm looking at the Medicaid Index admissions in that last column. There are Medicaid Index admissions in almost all hospitals, right? Ninety percent of hospitals are having some proportions of their patients having Medicaid insurance. In the 10th percentile, the hospitals in the 10th percentile, least amount of Medicaid admissions of their patients who are

coming within AMI have Medicaid. So, it's just telling us that, you know, there – that there are number of Medicaid admissions across most hospitals. It's not like only 10 percent of hospitals have any Medicaid admissions, and there is variation.

Male: OK. This is cumulative table.

Nancy Kim: Cumulative?

Male: Cumulative, so it's …

Nancy Kim: No. It doesn't add up to 100. It is ordered from the minimum number of eligible Medicaid admissions. So, I choose Medicaid as an example at any given hospital. But it's really a distribution. So if you could see, this will be a histogram. If we presented this in bar graph, it'd be a histogram, just a distribution of admissions at any given hospital that are percent Medicaid.

Male: So again, staring at table 21.

Nancy Kim: Yes.

Male: Ninety nine percentile are – is this table – last column, is this table saying that for the hospitals that are in the 99th percentile in terms of cost, the most expensive, the percent Medicaid admissions in those hospitals is 83 percent?

Nancy Kim: Thank you for your question. No, it's not saying that. It's basically saying that among all the 100 percent of hospitals represented by whatever, 4000 hospitals. In the 99the percentile which is the 99th percentile of Medicaid admissions at any given hospital, it's 83.5 percent of Medicaid admissions are happening at the hospital. That's representing the 99th percentile of hospitals with the most Medicaid admissions. Does that make sense? There's no payment anywhere in table 21. It is merely describing the distribution of the percent Medicaid patients at any given hospitals. It's just – there's nothing – no information about payments in table 21.

Male: OK. So, it's ordered by – it's ordered by the – by whatever …

Nancy Kim: The percentage of black admitted.

Male: Whatever the column, their head column variable list.

Nancy Kim: Correct.

Male: Oh, OK. Never mind.

Nancy Kim: The percent, yes.

Male: OK. I'm misreading the table. Thank you.

Nancy Kim: No problem, of course.

Larry Becker: So, this is Larry. I have a question to a lot of these and there are lots of questions in – it feels to me a little bit like this takes a lot to understand this data. And so, and we take this beyond this group. It needs to be some really serious work on how to communicate this well.

Nancy Kim: Is there something I could clarify?

Female: Larry, I just think the point is that if we want – people aren't just going to take it at face value and so the detail to understand it is significant.

Nancy Kim: It's true.

Larry Becker: That's right. Exactly right. And so, you know, if we want to make sure that people get it and we don't end up getting a whole lot of misconceptions and wrong ideas out of the data, then, do a really careful and fastidious job of explaining what this data says.

Ashlie Wilbon: Hi, Larry. This is Ashlie. So, thinking more on the side of like NQF in terms of summarizing the discussion and tying in the developers' analysis to the committee's ultimate recommendations, or are you thinking in terms of just the way that Yale presented the information? I'm just trying to figure out how we can – how and where we provide that clarification. Because I agree, I think it's not easily digestible by most audiences. So, I think your point is well taken.

Larry Becker: Yes. I don't know that I have any more to say. It just feels to me like we've got to be really careful with this one.

Lina Walker: This is Lina, just putting on my CSAC hat. You know when this eventually goes to the CSAC committee, I think it's that there will be a lot of questions about, you know, helping to understand the results and why the results didn't move as much as some people might have expected it to.

So, you know, we touched a little bit on that discussion earlier on, when Nancy, you went through the first AMI results. But I – you know, I think at least from the CSAC committee perspective, I think it would be really helpful to have a little bit more discussion about why there was so little variation with the inclusion of the Medicaid and the black variables when you would have expected a lot more.

And then I would say that the direction, you know, the sign change was surprising too. Just based on the earlier discussion that we had, I was surprised to see that it was actually a little bit low. So, I think – so, I mean that's not to say it was done incorrectly but I think that you will definitely have those questions, and so anticipating those questions and being able to have the material before them or at least have somebody explain or talk through those issues would be very helpful.

Lisa Latts: And this is Lisa. I think you're probably going to need, you know, sort of back to Larry's comment that, you know, assuming it passes us, passes CSAC and the board approves it, I think there's going to need to be in addition to, you know, access in the analysis which, you know, most people are probably not going to do. Some plain English summary and mechanism especially as NQF embarks on a premise for the first time so that folks can, again, in plain English, which is going to be very challenging. What does this – how it was done that showed?

Ashlie Wilbon: Thank you all. I think that's a really great point and we're sitting here nodding our heads as well in complete agreement, and I think that's something that as NQF staff, we're going to have to work on throughout this process. This is, as you guys know and I think I've stated on previous calls, this is the first time

committee, a committee has reviewed empirical analysis underneath these trials. So, I think any additional support that we can provide in terms of communication and helping others understand is going to go a long way with other committees and just sharing the results and our progress through this trial. I think very well taken.

On that note, Yale, I wonder if – I'm sorry, Yale. Nancy and Susannah, I wonder if to that point, is there – is any additional, maybe we could do a little bit of kind of between measures of comparison on – I know you – I know Nancy we're talking quickly to try to get through the information to make sure there's enough time for discussion, but I wonder if it might be useful to kind of go back and talk a little bit more about each of the measures and the differences in the results. And maybe to Lina's point, why there was or was not differences between the measures in terms of, you know, black patients. In some cases, having lower payments and Medicaid higher to some others, so could you maybe talk a little bit more about that?

Nancy Kim: Sure. Ashlie, did you want me …

(Crosstalk)

Ashlie Wilbon: … Reiterate, maybe just reiterate what you said.

Nancy Kim: I know. And I know, and I thank the committee for their patience. This is a lot of material to digest. And if you're not oriented or you're not deep within it like we are, it can be very disorienting. Ashlie, is it helpful for me talk a little bit about heart failure and differences? I won't go through step by step, everything, but I will try to walk through the broad brush strokes in a step wise fashion. Is that what you were thinking? Hello? Hello? Oh, my God.

Male: We're here.

Female: All right, it sounds like we may have lost her again. Go ahead and do the …

(Off-mike)

Nancy Kim: OK. So, I'm looking at heart failure and just to orient everybody, I'm on page 13 if you have your packet in front of you. I'm sorry. Not – wait. I now page 13 – yes, I'm on page 12. So these are the heart failure hospital-level of results. Again, so no impact on the overall distribution of hospital risk generated payments or the same hospital risk generated payments, SDS variables were added.

We found among the 740,000, there's admissions, that 11.5 percent of the admissions were categorized as black. And about 24 percent were categorized as Medicaid and 5 percent were both. So again, the take home is that there's not a lot of overlap between our black variable and our Medicaid variable.

In table 12, you can see the distribution again but there are Medicaid and black admissions at almost all hospitals. And median hospital for black admissions, there were only 3.3 black admissions to the median hospital. And for Medicaid, there were about 23 percent Medicaid admissions to the median hospital in the Medicaid admission distribution.

I'm going to move to page 13. And again, this is the hospital-level of results. If you focus on table 13, there's really no impact on the distribution of the overall hospital risk-standardized payments.

Again, in table 13, you'll first see our current model where the minimum payment is about $11,700, our maximum is about 22,100 and our median is 15,200. When we add black alone, it doesn't change much. When we add Medicaid alone, it doesn't change much. When we add both black and Medicaid to the current multi-variable model, we don't see a lot of change in the hospital distribution of payments.

Table 14 again is our percent change in heart failure risk-standardized payments, when we add the SDS variables to the current model. So this is trying to look at the same hospital, right? So when we're looking at hospital A with the current model, what is their risk-standardized payment? When we look at hospital A with the current model plus black and Medicaid added in, how much does their risk-standardized payment changed? That's what table 14 is telling us.

Obviously, in the first row, it's just current plus black. And the second row, it's current plus Medicaid. And the third row, it's "the full (inaudible) model", the current plus black and/or Medicaid. And what we see here again is that 97 percent of hospitals change by less than 1 percent in either direction. Sixty three got little less expensive. Thirty four percent got a little more expensive. But really 97 percent of hospitals change by less than 1 percent in either direction when you've added the SDS variables in.

And only two …

Female: So Nancy?

Nancy Kim: Yes?

Female: I'm sorry. Go ahead and finish.

Nancy Kim: No, no, no. It's fine. Only less than 3 percent of hospitals increase their risk-standardized payment by greater than or equal to 1 percent. And only 0.2 percent of hospitals decrease their risk-standardized payments by 1 percent or greater.

I'm done. What was your question?

Lisa Latts: This is Lisa. I don't know if you can answer this. But the question is going to come to everybody's mind, I think, is that is this (truth). Does SDS really matter so little in these measures or are we using the wrong variables?

Nancy Kim: It's a great question. I think that …

(Off-mike)

Nancy Kim: Go ahead. Did you want to say something? I agree that the variables that we have are blunt tools at best. We have this discussion webinar one that we don't have access to. A lot of them were granular tools. I know people were hungry for the nine-digit zip. We are too. And it's – we are not ready to use that right now. But we are looking forward to that variable. But we only have five-digit zip right now and I think the consensus' decision from the

committee webinar one was not to perceive because we know that is also a very blunt tool, and who knows what's that telling us.

Was there another comment?

OK. So, I'm just going to move on. So that's really the hospital-level effective SDS on the heart failure risk-standardized payment.

We looked at bivariates as well when we added black alone. So this is just saying how much is black (race) associated with total episode of payments? Black admissions were 1 percent higher in the total payment compared to non-black at the bivariate patient-level analysis. This is the same with Medicaid alone, and when we looked at the bivariate patient-level analysis for Medicaid alone, Medicaid admissions were 6 percent higher than the total payment compared to non-Medicaid admissions.

When we added black to current multivariable patient-level model, payments were 3 percent less expensive than non-blacks for heart failure. When we added Medicaid alone to the current multivariable model, payments are 1 percent more expensive than non-Medicaid admissions. When both were added to the multivariable model individual result held, that is black patients were less expensive than non-black patients and Medicaid patients were expensive than non-Medicaid but not by much, right. This is still a very small increase in payments in both directions.

We also looked at the model performance when we added the SDS variables into our heart failure model and there was really no change with the addition of black in their Medicaid variables. And these results are found in table 18.

We did the residual analysis again and that answers the question of whether or not the addition of the SDS variables improve model fit for subgroups. So we looked at black alone, non-black alone, Medicaid alone, non-Medicaid alone, and we found that adding the SDS variables improved the residuals for all of those four parsed out subgroups but it did improve it more for black and Medicaid subgroups.

We did our co-linearity analysis in heart failure as well and we found that the SDS variables were not co-linear with each other or other risk-adjustment variables.

Ashlie Wilbon: Hey Nancy?

Nancy Kim: Yes?

Ashlie Wilbon: This is Ashlie. I wonder if you could talk a little bit about – I know within the AMI measure, you talked a little bit about how the literature in terms of, you know, the number of procedures are being done on blacks for AMI may differ a little bit, which may have some impact on how, you know, how much impact the variables are having on the risk-adjustment model.

I wonder if there's anything in your literature review that you guys did for heart failure and pneumonia, if there's anything in there based on your analysis that you think might help folks wrap a little bit more with why there are no differences, or maybe what some potential factors maybe for they're not being as much differences people might have expected.

Nancy Kim: That's a great comment. So for AMI which is an acute condition that is treated both with medicines and with procedures, and procedures like surgery or catheterizations are expensive. It does sort of help to explain the differences. Although, I think (Jack) brought up that we're also having post discharge payments in this episode of care, so it doesn't explain all the differences. But we do feel that the knowledge that we have that black and Medicaid patients received your procedures does have to explain the fact that black and Medicaid patients were less expensive for AMI. Although again, at the hospital-level, there was no change. That's just at the patient we saw that.

For heart failure and pneumonia which are acute medical conditions that are treated really with medicines, there are a lot of procedures. And yes, in heart failure, you can have some procedures. But by enlarge, most heart failure patients and most pneumonia patients are being treated medically. So there's no invasive procedure like surgery or catheterization that's analogous to AMI. So we can't really have the same logic about explaining the findings that we

did discover when we did the SDS analyses. I don't if the committee has other thoughts as well. I know there are numbers of clinicians on the committee.

Susannah Bernheim: Yes. So the only thing to add – sorry – not for the committee. This is Susannah from CORE. The only thing I would add because obviously (inaudible) to say it is remember that we're adding these variables into a model that's already pre-thoroughly risk adjusted as much as you can for clinical issues. And so part of what happens with something like SDS is that if we built a model with SDS – you know, it's Medicaid alone, it might look to be a more powerful variable. When we add it to a model that's already adjusted for the extent to which patients are coming in sicker, it's going to be a less powerful table.

So I think that's the one other piece. I think the two important ways to claim this are one, that we've already adjusted for clinical factors and we're only seeing how much more we get. And two, that there is – I mean quite honestly, there's a lot of complexity to the way that raise in socioeconomic status contribute to cost, and they don't all wind up in the same direction, and so this ends up a little bit of a wash.

Nancy Kim: Yes. I'm glad you raised that because I think that the clinical factors so I don't know if those – I can't remember if there's a drug from age 50's. I mean there's already a lot of adjustment going on.

Female: Ashlie, did you want me to pause or move on or what?

Jack Needleman: This is Jack. I have a question about – or a comment. I'm not quite sure which. I'm trying to reflect on the – what we're actually including as SDS and what we're measuring, and they're all individual level variables. So we're looking at variations in treatment cost given the random affects model largely within hospitals. They might vary by rates or by the presence of a Medicaid eligibility status.

When we had the – reaching about the community context, and the SDS of the community that the hospital is serving and the impact of that on availability resources, on the cost of the resources, the risk of readmissions which got

raised earlier, and it's in the conceptual model, and I think in the conceptual model rather nicely. But there's no – there are no community level SDS measures in this model. And I'm wondering will they consider they're not included, were they consider – not considered appropriate that the community context for the treatment be directly measured in this. Because right now, to the extent that we have community factors, the individual variation – the individual-level measures of standing is they're very poor proxy for those.

Nancy Kim: Yes. This is Nancy. It's a great comment. We are always considering the impact to the community on these health outcomes. We just don't know how to capture that in the measure. And as you said, whether or not we should as there's pros and cons. But right now, we're not even there because there's no feasible way. I think that's probably where the nine digit zip may help but we're not ready to use that. The five digit zip, we don't think it's (civil).

So it is a great point, something that we continue to discuss internally. But we don't really have access to a variable of the national level that we feel comfortable using to represent the community factor. And then I think we're not quite sure how comfortable we feel, I think, including all of those community factors, yes or no. You know, there's – we haven't given that full thought because we don't have access to the variables there. Our first order of business is really to get that nine-digit zip ready, up and ready to try in our model.

Susannah Bernheim: And there's just two things there, just so – this is again Susannah. Just to follow on what Nancy said, right. So there are variables that are measured at the community level that we could introduce as a proxy for patient-level and put it in at the patient-level, right. So the first pass with the nine-digit ZIP code is to say this nine-digit ZIP code tells us something probably about this patient and certainly an aggregate for a hospital public client fair amount about the patient the hospital to take care. But we would put it into the model the same way we did with Medicaid, representing a patient-level variable knowing that it's a – I mean, perfect proxy for that, right. So that's one question.

I think your question, and Nancy addressed this as well, was more with the risk-adjust for, you know, this county has, you know, fewer primary care doctors than another county, right? And so, would we actually put in a variable introducing quite honestly another level to this model, right. We already have a patient-level model and a hospital-level model. Would we introduce another level for this model? And then what you're doing is setting a different comparison group for your hospitals. You're basically saying how are you doing compared to other hospitals and counties with few PCPs. And it's as Nancy said, but now a little bit more.

It's a big conceptual leap that we're not comfortable with. But if this committee – and this a question Nancy and I were both on that committee in. So with the question of the committee didn't really have a chance to get into that group. Under what circumstances would you add adjustment at a community level where the factor goes into the model as a community factor, and when would that make sense and when do that not?

And I'm – Ashlie, you can stop me because this is me just trying to take advantage of the smart people in the room. But I would love to know what people on this call thing of that because I think it's potentially very problematic what people are interested in, and it's a note a place we've gone yet. So if this committee has strong feelings, I'd be interested to hear that.

Nancy Garrett: Sure. That's fine. Thanks.

Lina Walker: This is Lina. So conceptually, I guess I'm having some difficulty understanding the rationale for adjusting for (inaudible). So if we're using this measure for accountability purposes, you know, the risk adjustment at the front end where we're accounting for, you know, their health conditions and the intensity of their conditions, this isn't in theory supposed to level up at that point. And the – so the discussion we're having about adjusting for SDS is because of missed measurement at that level.

And then we've been talking about some potential miss measurement with the use of these standardized DRGs. But just listening to the discussion about the use of hospital random effects – I mean, wouldn't that capture, like if you're

concerned about differences in the community, the hospitals are in the community, when that pick up those community different? I don't know why we would want to include an additional adjustment for community attributes. So I mean I, you know, would be glad to hear people talk about this more just trying to understand the arguments here.

Jack Needleman: No. I think that's a legitimate and appropriate and important question. I only go back to the whole arguments about value and resource use and quality that have been, you know, under guard this. And the (inaudible) has been – are in communities that have fewer resources to deal with post hospital homecare. Some patients maybe – so you've got the patient-level variables, right.

Some patients are less able to cope with self management or homecare and more likely to bounce back. We need more spends on them or more likely again institutionalized because of their individual condition, and that's at the patient-level.

But the only data we're seeing in this analysis is, of course, patient-level data for Medicare patients because that's what the measure is build upon – around, and it doesn't matter whether the hospital is 10 percent Medicare or 25 percent Medicare or 50 percent Medicare. We're analyzing just the Medicare patients.

But the argument is made that for some of the post hospitalization services, the community context matters is homecare available. It's high quality, nursing homecare available, or the follow up physician services available in the community, and that's going to be – and those may or may not be fully captured by the patient SDS given that it's a patient SDS for a Medicare patient.

And then the second argument that has been made on this is that, you know, the hospitals that are serving less advantage patients that have 10 percent Medicare and 50 Medicaid in non-insured or 70 percent Medicaid non-insured. Ten percent Medicare, 10 to 20 percent private pay and everybody else is basically Medicaid non-insured have or likely to have fewer resources available within them to treat within the hospital. And also the – and again, beyond community that have fewer resources available for the post hospital or

outside of the hospital care that might affect the cost or might affect the patient outcomes.

And that's I think a lot of the argument that's been raised about appropriate SDS risk adjustment. And the question is whether those risk adjusters are most appropriately captured at the patient-level or the community level in which the hospital is situated, and the patients – the communities that the hospital actually serves within that neighborhood.

Female: Thank you, Jack. So you're saying that the hospital random effects wouldn't pick up those differences (from) the community?

Jack Needleman: No. But what the hospital random effects does is it basically produces the hospital a hospital – estimated hospital effects in posting some distributional requirements on that. And then for smaller hospitals where there's more uncertainty in the measurement of that, it will shrink that measure towards the mean by averaging with the mean.

What random effects models allow you to do, includes more hospital-level measures in your model which the effects model would not allow you to do that. But a random effects model does, but we don't have hospital-level variables in this models, and I think that goes back to the CMS philosophy about what they think is appropriate to take into account in comparing higher and lower cost hospitals or higher and lower outcomes hospitals. And a lot of the argument about appropriate adjustment for SDS has been about whether that the model that CMS has and what's appropriate is in fact the appropriate one.

Female: So, but then from what you're saying, so then are you suggesting that where we would see the difference would be in the post discharge spending, not so much in the inpatient spending?

Jack Needleman: Well, I think we will see it in the post discharge spending more than the inpatient. I think that's in fact the case. So some analysis of cost variance around the non-hospital part of the about those statement I think might be a good thing to do. But also, the issue of the bounce back into the hospital for readmissions, I'd like to know what percentage of the variation of cost is

associated with the readmission differences and readmission risk towards the hospitals. And that might be a function of what resources are available in the community for the patients who were discharged back into the community.

Female: So Yale, do you have that data to share?

We did not look at the association of SDS with any post acute settings. So we only look at the association of SDS with the total episode payments.

Female: OK, I think I got that.

Female: Thank you for the discussion, Jack.

Ashlie Wilbon: Yes. So, thank you guys. And we're at about – I think we've got about 20 to 25 minutes left to the webinar. So I think the discussion up to this point has been really, really interesting and I think we're all learning a lot here. As a part of this discussion, I wonder if there's any final questions particularly for Yale on their analysis or decision not to include variables. Because if not, I'd like to move on to the next – kind of wrapping up for next the steps, what we want to committee to do, some guidance on voting, and then I'd like to open it up for public comment. I know we've got quite a few folks on the phone. So I'd like to give some people an opportunity to make some comments to the committee and potentially have some opportunity for discussion of those questions if needed.

So are there any final questions for the developer from committee members? OK, hearing none. It doesn't mean that it's over. You can certainly e-mail myself if you have any questions and I can, you know, pass this on Yale. We can see the discussion as the process is not over.

So we'd still have commenting and CSAC and all that stuff. So I do want to thank Yale tremendously for the amount of work that you guys put into this. I think being – kind of being the (guinea) pig in this process and being the first developer to kind of go through this process with a real focus on this analysis has been really helpful and we appreciate your work for this, so thank you.

Jack Needleman: Ashlie, this is Jack. I have probably been the most skeptical in the call, not about the work that's been done but about some of the limitation – inherent limitations of the measure given data and the CMS framework for it. But I also just want to say I deeply, deeply appreciate the work that the Yale team has done on this, both in terms of the overall quality of the work and originally and also the responsiveness to some of our concerns.

Ashlie Wilbon: Thanks, Jack.

Nancy Kim: Thank you.

Ashlie Wilbon: So with that, I do want to talk a little bit about the next steps for the committee. So again, you guys should have all received link, a link from me in an e-mail earlier this morning or late morning with the link to the SurveyMonkey tool as well as it's also in the memo that I sent out last week at the very bottom in that last committee actions box at the bottom of the memo. So you can access it two ways. It's open now.

And for those of you that are on the call and you're sitting there, a computer, I would encourage you guys to vote sooner that later so that – and kind of, you know, make any comments and vote while the discussion is fresh in you mind.

I would then like to encourage you to take a look, make sure you look at appendix A and B in the NQF memo that was sent out. I pulled out the validity criterion for the resource use measure evaluation criteria so that you can re-familiarize yourself while we're focused today on risk adjustment. Risk adjustment is a sub-criteria within the validity criterion. And we are going to have this overall in the validity criterion. So I've also, in appendix B, included excepts from the measure summary of your discussions of these measures, specifically on the scientific acceptability portion of the criterion for each of the measures so that you can re-familiarize yourself. And as you're voting on validity, just encourage you to consider all the elements and – of validity in the context of your previous discussions on the measures, your discussions today and from the risk adjustment. And also take note of, you know, the center of the committee's votes in terms of consistency. We want to try to – you know, the process is somewhat disjointed by nature of us kind of

going back in voting on this one sub-element of the measure. But I would like to just call your attention to that piece as well.

For example, the first measure in appendix B, they're actually kind of listed backwards from the way they are in the Yale memo but 2579 is pneumonia. For validity, the committee voted three high, 18 moderate, one low. The issues were around – in terms of validity we're around testifying a 30-day episode of care for the pneumonia measure. The attribution of cost in terms of transfer, patient being transferred from one hospital to another and the R squared which seems to resolved as a result of the developer's feedback on their analysis of why the risk – the R squared might have been as low as it was. And then the final concern was around the, you know, consideration of – for SDS factor which is what we're addressing now.

And I think those comments are very common throughout those – these three measures as they were very similar although different conditions. I think the committee's concerns were pretty consistent throughout. I know for heart failure, there was a few different concerns because of the nature of the condition being a more chronic condition and concerns about for 30-day episode. So please do take a look at that before you submit your validity votes and keep that in mind.

For each of the votes, we are asking that you type in a rationale so that we some sense of what is, you know, driving your vote, so we can justify and aggregate the committees' recommendation, either recommendation for the – for continued endorsement or not to continue endorsement but also on the validity criterion.

So with that, I'll just pause to see if anyone has questions about kind of next steps and submitting their – in submitting your votes. For those of you that are not in front of a computer and are not – do not have the time to submit your votes right now, we're just asking that you do that by the end of the week, so by COB Friday. That gives you three, four more day – three more days sort of to submit your votes to try and encourage you to do that while it's as fresh as possible. So, any questions about that?

OK. With that, I'd like to open it up for public comments. (Bridgette) are you there?

Operator: At this time, if you'd like to make a public comment, please press star then the number one on your telephone keypad. And there are no public comments at this time.

Ashlie Wilbon: OK. Can I – I'll just give it another minute because maybe people are thinking about what that question should be. I know we've got a lot of folks on the line.

So hearing none, so no additional prompts for comment (Bridgette)?

Operator: There is no one in for comments.

Ashlie Wilbon: OK. With that, I'm going to ahead and keep moving. If anyone does have questions, feel free to – that you'd like to submit, feel free to submit them in the chat box. We will do our best to respond. Oh, there was some comment in the chat box. Sorry, we're trying to go back.

(Off-mike)

Ashlie Wilbon: OK. But I want to try to ...

(Off-mike)

Ashlie Wilbon: OK. So there was a committee about – I'm sorry, there was a question about the committee's discussion about the ZIP code and talking about business track as a better unit.

The ZIP code is only for postal service regardless of how many digits. And then there was some data or some information put into the chat box which I won't read all of them, but some reasons for choosing census track over ZIP code data. So I can certainly share that with the committee, but are there any comments from the committee or others about the use of census track versus ZIP code data?

I know that you guys talked about that on the first call quite a bit but I just …

Nancy Garrett: So this is Nancy Garrett. I mean that's a really good comment. And my recollection is that we talked about that and we recommended using census black groups. And in order to do that, you'd had to have address data from Medicare and CMS wasn't able to provide that in time from we needed it. That's my recollection.

But I think as we, you know, go forward, you know, I don't know if we can ask that that would eventually be pulled in to be considered in the future for this measure because – just because we don't have the data now, it doesn't mean that we shouldn't be pushing to get better data. That would be one way to do it.

Ashlie Wilbon: Sure. Thanks Nancy for that reminder. So yes, that's certainly something we can include in the recommendations section that I think will be helpful for future evaluation of this measure but other measures as well, so thank you.

All right. With that, I'm going to go ahead and move on to the next item which is just to give a quick overview of the remaining milestones for this piece of the process in reviewing the SDS factor. So we will be putting out a memo summarizing this – the both of the webinars for the evaluation of the conceptual and empirical analysis for these measures on November 16th. That will be out for comment for 14 calendar days due to 30th.

We will be going to CSAC on two locations primarily because we want to give the CSAC an opportunity to discuss that in detail, their in-person meeting. However, the comments will not be back yet because it will still be out for comment during the CSAC meeting. So we will have an agenda item at the in-person meeting for CSAC in November but ultimately, their final decision will not be made until a discussion after – until after their discussion on the December 8th call during which day we'll actually have an opportunity to review the comments that were received during the comment period, and they will make their final recommendation after that call.

We will go to the executive committee in January and then out for an appeals process shortly following that. So depending on when the executive committee meets will be able to update those dates when we have them. So those are the next immediate steps for this process.

Just a reminder again that we would like you guys to submit – for the committee to submit their votes by COB on Friday, so we'll be pinging you for some reason we don't get those votes from you. But particularly for those of you that were on the call today, please try to get those votes in. It will help us meet our deadlines to get out for comment on time as well.

I did want to mention also that this committee has been convened I think for a little over two years. And if you recall, the first time that you guys met as a standing committee, I know some of you guys have been with us for sometime before we actually had standing committees and then were seated on the committee as a standing committee member where you were assigned terms of either two or three years. So those of you that were assigned to two year term or you picture two year term out of the hat, we will be reaching out to you over the next few months to see whether or not you'd like to continue on. There will be – we'll send you out some e-mail detailing what that actually entails and what the term roll over process is. Ad we will – for those of you who don't remember what your term is, we will let you know so don't worry about that.

Also, I wanted to just give you a little heads up on – for now, we don't have any funded work that we – that is scheduled to come in for the evaluation of cost to resource these measures. It is something that we are continuing to try to put on our list for upcoming projects, then we'll see how those next funding process goes, the next funding cycle goes.

In the meantime, for each of the standing committee topic areas, we're going to be instituting kind of maintenance activities for the – each of the committees to make sure that we're staying engaged and keeping on top of what's going on within the topic area, so within the cost and resources arena.

Because this committee has been busy with this ad hoc review, this has been kind of one of your "maintenance activities" to make sure that you're staying engaged and we're also kind of getting work done in the process as a standing committee.. So we'll be in touch with the committee members on – once we have establishment terms. We're going to roll over for individual and so forth and give you a sense of what those maintenance activities will be.

Generally, it will be at least one webinar within a quarter to talk about any pressing issues about what's going on, on the topic area and staff will work with co-chairs and the committees to figure out what those issues are and when those webinars will be.

So more to come on that, and I just want to thank everyone for joining us. There's not anything else I think at this point that we have on our agenda for you guys. I did want to just let everyone know that this is my last week at NQF, so.

Female: Oh no.

Ashlie Wilbon: Yes. So I've been working with some of you guys for a very long time and I just wanted to thank you for such an amazing experience. I've learned a lot through all of you, listening to your discussions over the years. And I've also gotten to know many of you as individuals and it's been a pleasure to work with you.

I'm moving on to – going back to clinical practice actually as a nurse practitioner. So I'll be in the area and around and I'll still be at NQF but I won't be working out on this project. I'll be in a much more limited capacity. And (Taroon) will be, who you all know very well, who will be carrying the process, this project through the remaining steps in the process.

So you'll, you know, continue to hear from him. And (Anne) will also continue to work on the project. But I just wanted to thank you all and it's been great working with you, so thank you.

Male: Best of luck, Ashlie, we will all really miss you.

(Crosstalk)

Ashlie Wilbon: Thank you.

Female: Ashlie, good luck.

(Crosstalk)

Female: Thank you so much, Ashlie.

Female: Yes. You've been great to work with.

Female: Really fantastic. We really appreciate it.

Male: We'll miss you. Best of luck to your new position.

Ashlie Wilbon: Thanks. Thanks guys. OK. So feel free to forward any questions if you have them. Thanks everyone for joining us today for a great discussion.

Female: Take care. Thank you.

Male: Bye-bye everyone.

Female: Bye-bye.

Female: Bye.

END