NATIONAL QUALITY FORUM

+ + + + +

DISPARITIES STANDING COMMITTEE

+ + + + +

THURSDAY
JANUARY 21, 2016

+ + + + +

The Standing Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 9:00 a.m., Marshall Chin and Ninez Ponce, Co-Chairs, presiding.

PRESENT:
MARSHALL CHIN, MD, MPH, FACP, Co-Chair
NINEZ PONCE, MPP, PhD, Co-Chair
PHILIP ALBERTI, PhD, Association of American Medical Colleges
SUSANNAH BERNHEIM, MD, MHS, Yale-New Haven Health System Center for Outcomes Research and Evaluation
MICHELLE CABRERA, SEIU California
JUAN EMILIO CARRILLO, MD, MPH, New York-Presbyterian Hospital; Weill Cornell Medical College
LISA COOPER, MD, MPH, FACP, Johns Hopkins University School of Medicine*
RONALD COPELAND, MD, FACS, Kaiser Permanente
JOSE ESCARCE, MD, PhD, University of California at Los Angeles*
TRACI FERGUSON, MD, MBA, CPE, WellCare Health Plans, Inc.
KEVIN FISCELLA, MD, University of Rochester
NANCY GARRETT, PhD, Hennepin County Medical Center
ROMANA HASNAIN-WYNIA, PhD, Patient-Centered Outcomes Research Institute

LISA IEZZONI, MD, MSc, Harvard Medical School
DAVID NERENZ, PhD, Henry Ford Health System
YOLANDA OGBOLU, PhD, CRNP-Neonatal, University of Maryland Baltimore, School of Nursing
ROBERT RAUNER, MD, MPH, FAAFP, Partnership for a Healthy Lincoln
EDUARDO SANCHEZ, MD, MPH, FAAFP, American Heart Association
SARAH HUDSON SCHOLLE, MPH, DrPH, National Committee for Quality Assurance
THOMAS SEQUIST, MD, MPH, Partners Healthcare System
MARA YOUDELMAN, JD, LLM, National Health Law Program

NQF STAFF:
HELEN BURSTIN, MD, MPH, Chief Scientific Officer
ELISA MUNTHALI, Vice President of Quality Measurement
JANINE AMIRAULT, Project Analyst
KAREN JOHNSON, Senior Director
ERIN O'ROURKE, Senior Director
MICHAEL PHEULPIN, Project Manager

ALSO PRESENT:
DAVID HUNT, MD, FACS, Office of the National Coordinator for Health IT, HHS
CARA JAMES, Centers for Medicaid and Medicare Services (CMS)

* present by teleconference

TABLE OF CONTENTS

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

(9:01 a.m.)

CO-CHAIR CHIN: Is there anyone on the call right now? Anyone calling in? So we know Lisa Cooper will be calling in. Poor Lisa. She lives in the Baltimore area and so it took her three hours to get home last night after the meeting. And she just gave us a call that she's already been on the road two hours and she said I just don't think I'm going to get here because I think it's the ice that's causing the problems and so she's turning around and going back to Baltimore and calling in then.

And Jose, is Jose on the line? Okay.

MEMBER COOPER: I'm on the line. It's Lisa.

CO-CHAIR CHIN: Hi, Lisa.

MEMBER COOPER: Hi, everybody.

CO-CHAIR CHIN: We're all commiserating with you right now.

MEMBER COOPER: This is crazy.

CO-CHAIR CHIN: So we want to just do

a brief recap of yesterday and then we'll launch into today's agenda.

So first, Helen, Ninez, and I were just commenting that this is incredible to say that everyone on the committee spoke up frequently with very helpful comments which isn't always the case. So that was great. So I think the group dynamic was terrific and we really appreciate everyone building upon each other's comments and you know, it was good, constructive conversation.

Remember yesterday that the focus of yesterday was the work on the first charge of the committee to develop this road map for a variety of end users on what could be done to reduce disparities. And a lot was covered. I'm just going to go over a few of them, this measure, big picture themes. And what's going to happen then is that, over the next week or two, Ninez and I will huddle within the NQF Team and we'll try to come up with sort of a more flesh on the skeleton for you guys to give feedback on and to further build as we develop the process.

Probably the first major theme was that -- it was about the scope of whether it's health or healthcare. There was a long discussion about -- I think Helen said it was a violent agreement that we want to have a broad scope of what health is. It's not just healthcare within the healthcare sector, but the goal was overall health, overall population health. And so it has implications in terms of the metrics, for example.

Some will be probably like the traditional clinical performance measures that we're used to within the healthcare system per se, but others that then perhaps capture other elements of a broader health which actually fits well with NQF current work of population health and other groups that are working on population health, IOM, Healthy People 2020, et al.

Sort of related to this though, there's the fear about well, it's just too much. We're sort of biting off too much. But then a number of folks made the point that you could have an expansive goal so health is more expansive than

healthcare, but if you focus on the leverage that we -- well, our target audiences have influence and control over, well, that's the way to limit it. In other words, the example they gave as well, something like housing and education, they clearly are the key roles then for the housing community or the education community, the leverage that we have in terms of health are through the mechanisms of our target audiences which was one of the areas of discussion where the conclusion there was that it was broad so that still a broad variety of stakeholders and you remember like these show that have like four or five different particular stakeholders, but there is probably some type of prioritization. And so there was a lot of talk about the payers, for example, and CMS in particular being a key audience member.

And Cara is back. Welcome, Cara. It's great to have you back. So there would be a way to basically pass sort of a broad vision that I think everyone agreed with, but in a way that wasn't pie in the sky, that was a realistic given

who the end users would be et al.

Sort of related to that, a number of folks, I think the last was like Ron who talked about the idea about accountability. And so we talked -- Susannah, Sarah, I think Eduardo, sort of led off that conversation that in some ways -- like anything else, I mean you've got to have logic model to help guide things and the way Sarah described it was thinking about who are the key actors which will be everyone from NQF to a lot of the target audiences. Then thinking about them, what are their levers? What are the mechanisms they have for action leading to the outcomes. And so that will be fleshed out over time.

One of the handouts on the table, there was the conceptual model that -- I can't remember -- was that Eduardo's, yes, Eduardo, they talked about yesterday were these holistic models that involve the patient, the family, the healthcare system, the broader community, the traditional non-health sectors and all. And a key also then is how these all interrelate. So I forgot who it

was who had mentioned but the idea of connectivities, inter-relationships that -- because there were different levels of action. So if you had some of the payers, they're going to influence, of course, payment policy. Another one of the big stakeholders was like the health organizations and the providers, clinicians, I think is the way it was worded on the slide. And they have their own actions, but they're also influenced by these other key parties, not just the payers, so this model, when you think a little bit about how this all interrelates.

We also had a good discussion about what is such target population. The conclusion was it's broad. It goes beyond race and ethnicity and my take on like the people's conclusion was that we should have some type of example list of example groups so that there's some concreteness to it, but at the same time and maybe it was like -- I'm going say it was Kevin, I think it was Kevin who said that, but worded in a way that it's clear that this is not necessarily an exclusive list. These are

examples. So the idea is to be inclusive, but to have some type of specifics so that it's concrete.

There was also a discussion of what is the comparison group. So it was a nice discussion about the importance of wording on things. The discussion was like the meaning of like vulnerable populations, at-risk populations, social advantage, historically-advantaged groups, and so I think like when there's the first draft of that it's the nuance that's important and the importance of what words can connotate. So that will be taken into account. We have this first draft and then iterate on that.

Next steps that -- I guess it's not -- right now it's not another in-person meeting planned over the next year, so it all will be internally funded by NQF. And so that the plan was then to try to do most of this by webinar. So Erin and Michael and the team will go back and we're going to probably have to huddle in the next week or two, you know, like Helen and Ninez, I, and the team, just to sort of debrief a little bit and come

up with a first crack so that Erin and Michael and team start doing the drafting, they have that road map.

We're also going to distribute the meeting notes and the notes that Erin and all have been taking, so people have that background to --

DR. BURSTIN: The time line.

CO-CHAIR CHIN: The time line, too. I think there was agreement that less than a year, probably six months or more, Lisa said, she thought nine months we're going of experience probably the right range, but it was this idea about like some degree of urgency that I think everyone agreed that this is really an important time and it's actually very good timing and probably would like to mention returns over time too. We'll probably have a better idea as this goes along, but somewhere between 6 and 12 months is probably the range that we're talking about.

Keep in mind, too, the major goal of yesterday was just some general sense, either agreement or at least like issues that need to be

discussed further on the webinars on things like the overall goal, the mission statement, the principles, some of these target areas. But there's a lot of things that have to be fleshed in much more detail. So for example, we were talking about the five or so key actors and then the actions they can do. It's good what we did. We came up with some general things, but they were really general still. Payment being a great example where talk about using payment to help reduce disparities, but you know, there could be a good number of ways that could be done and a good number of issues need to be discussed in more detail. It's not going to help. CMS and other payers say well, you know, use payments to reduce disparities unless we have more meat on it.

So that's my impressions. Ninez, Helen, things to add?

DR. BURSTIN: I was going to say I know Cara's back with us this morning from CMS and she was willing to make a few remarks perhaps.

CO-CHAIR CHIN: That would be great.

So Cara James is great. She's the person at CMS that heads the Office of Minority Health at CMS and so Cara, we're delighted you're here and --

MS. JAMES: Thank you. I feel like I shouldn't say anything after you said I'm great. I think obviously, as I said yesterday in the introductions, we are very, very excited and interested in the work that you guys are doing. And I think that as Marshall said yesterday, the timing for this is really truly excellent. And several of you are working with the American Hospital Association on the work that they're doing and as well with the Joint Commission on America's Essential Hospitals, but I think with the health equity challenge that they have that we, CMS, has really taken a bold step in this with our release of our CMS equity plan for improving quality. Really, it's like the first time, I think, that the Agency has said we care about equity and not only that we care, but we're really trying to work on it. And that's the thought and the feedback that we received from several of you in the room on how

we should guide that and we thank you for that and the continued participation as we go forward with that.

The easy task was coming up with the six priorities and activities and now the work is actually implementing that towards change. And I think that hearing from you guys in vigorous agreement that this body not be another body that's convening, but really one that's driving towards change, I think really aligns with what we are interested in as well. And just one of the things that I think can really help us is the push/pull, sort of having us kind of work on the inside with CMS, but you guys being a voice on the outside that's really helping to push us to where we need to be, so very much looking forward to what comes out of this. And I'm sorry that I missed the most interesting pieces of yesterday of the who and the what, so I'm glad to hear that you are trying to get down to the specifics, because I think that a lot of people we find are interested in reducing disparities, but not sure what to do.

And it's one of the things that we're thinking about how do we give them the tools and we think about those tools kind of on a continuum for those people who are just starting this and sort of not sure where to begin. Those who may have started and are trying to look for that next step to really take that next set of progress.

So I'm hopeful that the work that you guys are doing, it will really help to move us with that and help guide us as well as what we can do to reduce disparities in the future. So thank you for allowing me to listen in.

CO-CHAIR CHIN: Not on the agenda, but if Cara is willing to, some people had a couple of questions.

MEMBER SANCHEZ: So about a year ago, no, September 8, 2015, I participated in a Medicare and Medicaid at 50 Pass, Present and Future Impact on Health Equity. I'm wondering is there something, a report or anything, about that that might inform us in some way as we do our work?

MS. JAMES: Yes, so I think -- so there

were a couple of you in the room who participated in that and one of the things was both reflecting on Medicare and Medicaid at 50 and what those programs have meant in the future. But one of the driving forces of that event also was the release of our CMS equity plan for improving quality in Medicare. And so that was the six priorities and for those of you who may not be familiar with it, number one is something that came out of the conversations yesterday, data. So the collection and analysis and reporting of standardized data across our CMS programs and that is not racial and ethnic minority data, but also sexual and gender minorities and people with disabilities, as well as how do we improve other demographic data.

Our second priority is really embedding disparities into our programs as we're moving forward, but also measuring the impact of those and some of that conversation happened as well yesterday, We've been very pleased with our partnership with CMMI with the Accountable Health Communities because this is the first model to

include a disparities impact statement, but not only does it include the impact statement, it's actually part of the scoring, so we're very excited about that even if it is a small part.

We also are working on our Transforming Clinical Practice Initiatives and our looking at the SAN that's focused on minority populations there, sorry, the Supporting and Aligning Network, which is helping with the implementation of that. And are working with our colleagues in CSQ on the Partnership for Patients and work that's going forward there. So that's a couple of the things that we're looking at.

The third priority is really developing and disseminating solutions to address disparities. So some of the things that we have are -- that's not out just yet. We're working on -- any day now, literally. Maybe today, tomorrow. We've been working with Massachusetts Disparities Solution Center at MGH on a guide to help reduce disparities in readmission rates as well as are working on a mapping tool to help people understand

disparities at a local level. That's a little bit like CDC's mapping of their cardiovascular disparities, but drills down to a number of other conditions as well as some of the payment pieces.

Our fourth priority is making sure that we have a workforce that's able to provide culturally competent care and taking care of vulnerable populations and so some of that is looking at both ensuring that providers are educated on the class standards and how to do that and what tools we can use to support them, but also looking at other types of providers who are not like social workers, community health workers, those sorts, and how do we integrate them into care teams.

Our fifth priority is looking at making sure we have access for language, for not just people with limited English proficiency, but also people with disabilities and that we're able to communicate effectively with them to help them navigate the system.

And our sixth priority is one that came out of our listening sessions that we had in

developing it. I'm looking at one of our people who participated and that is making sure that we are improving functional accessibility to -- physical accessibility to healthcare facilities for those with disabilities.

So that's what we're going to be driving towards for the next four years. We also have other work that's going on, some that we've just begun with our colleagues over in our Office of Rural Health Policy for implementation of MACRA around education for chronic care management services.

We also have other work going on, but that is kind of what we'll be sort of reporting out on as we're making progress on implementation of the equity plan. And one of the other things that we're working very hard on is stratified data. And as Sarah mentioned, really embedding these things into our kind of payment models and so thinking about how do we put class into value-based modifiers or a physician quality reporting system.

And also the last thing I would say in

terms of embedding that we've been working on and are excited is the physician fee schedule, NPRM, and the MIPS request for information solicited comments. And we thank those of you who provided those on how we embed equity into the new MIPS program. We are leaning towards looking at the clinical practice improvement activities as where we can do that, but also looking at reporting stratified data on physician compare for quality measures.

MEMBER SANCHEZ: Was that an okay question?

CO-CHAIR CHIN: Great question, yes.

MEMBER COOPER: I have a question for Cara. It's Lisa.

CO-CHAIR CHIN: Go ahead, Lisa.

MEMBER COOPER: Cara, you were talking about one of the priorities on the equity plan being the workforce separation and I know there's a work group that is focused on integrating community health workers into care teams. And I'm just wondering how this committee could help -- I don't

know, inform the work of that group or sort of push for what's needed in terms of payment reform in order to allow better integration of community health workers, payment, as well as like training standards, things like that. Is there anything that we can do specifically with regard to that work?

MS. JAMES: Yes, so let me think about that because I think the challenges of that are not necessarily new challenges in the sense that one of the issues is standardization of that work across the board. So as we're thinking through how we move forward with that, it's being clear on what we would be asking people to do and requiring them and holding them accountable for.

But I think -- I'm hoping a little bit as well that we'll see a little bit more evidence, if you will, coming out of the accountable health communities because of the importance of their role in helping to support people and connecting them to some of those other services. But I think -- I will take that back and think about that a little

bit more and get back to you.

MEMBER COOPER: Okay. Thank you.

CO-CHAIR CHIN: So we've got Bob, Kevin, and Romana. And aim to go to maybe to 9:30 and then go back to the regular agenda.

MEMBER RAUNER: Two questions. One is what disparity-related variables have that you can merge with your claims data sets so that there is better ability to do the stratification?

The second one and goes back to attribution. I used to work with an ACO and that was our biggest challenge, especially in rural areas of making sure that the right patient was linked to the right primary care doctor. And one of the problems we were running into which was making it worse was the fly by night screening companies who would pop up in a church parking lot. They're actually doing annual wellness visits and billing them which makes the attribution even worse. Some of it I think actually borders on fraud, so maybe you ought to look into that. One of them was even going to start trying to do chronic

care management codes as a fly-by-night screening by getting -- convincing these elderly people to sign up for it which made it even harder for the local primary care doctor to keep tabs on them because everything kept getting stuck away by these screening programs. And so if you wanted to go looking for fraud, I would suggest that you look at those companies.

MS. JAMES: So definitely would want to follow up with you on that one. We are very interested in that one. And that is, I will say for our office, the fraud area is an area we haven't been really able to get into because we are a small office, but we have been talking to our colleagues over in CPI about how we can step into that sphere.

For your data question, so we have -- one of the things that we did is, as people know, our Medicare data is slightly problematic on race and ethnicity. So slightly problematic. It's kind of interesting. I don't think people actually know how bad it is because -- and by that I mean Social Security in their infinite wisdom

decided to stop collecting race and ethnicity in 1989, so any Medicare beneficiary born after January 1, 1990, we do not have race and ethnicity data for them. And that's clearly just going to grow as we move forward.

The IMPACT Act of 2014 requires us to come up with a plan, so we will be submitting a report to Congress in two months, two and a half months that details how we can move forward on that and we're hoping to actually put that plan into place.

In the meantime, we have as many of you probably know, a couple of imputation models. So RAND has developed one, the BIZG as well as RTI has one. We have taken our Medicare enrollment data and those two imputation models to the Census Bureau and matched them up to compare, see how they perform and are actually working to strengthen the imputation model in the meantime as we move forward.

I am actually pleased to say they do perform pretty well, so that's good. But we also

as part of that data, it is educating about how do you get better data in. So why it's important that people when they sign up for their coverage through the marketplace that they check the box. How do we get people over the uncomfortableness of asking the question, that we do find when you explain to beneficiaries and consumers about why we're asking their likelihood that they don't report goes down significantly.

This is also a challenge for sexual and gender minorities and it's something that we've been using our administrative codes and actually testing how we can identify transgender individuals through that data and have had some success with that, but we've also been working to cognitively test questions about sexual orientation and gender identity on our Medicare current beneficiary survey. So those cognitive tests just ended last month. But I will definitely -- I do want to follow up with you on that because the issue of fraud is something we very much care about and obviously for the populations we focus

on there.

CO-CHAIR CHIN: We'll do Kevin and Romana and then we'll go back to the regular agenda.

MEMBER FISCELLA: So two comments and two questions. My first comment is wow. This is like a sea change. I mean from where CMS has been to producing this plan, I am really impressed. The second comment is I'm a little bit embarrassed. I actually have not read this plan before. I wasn't even aware it came out which I think raises a broader issue is when CMS does this, it's like where are the press reports? Where is the fanfare? And I think it merits that. I think this is a very thoughtful and well-done plan.

My first question is will there be annual reporting on progress on each of the goals, will that be publicly reported?

MS. JAMES: That is the plan. So we will be looking at both as we implement the activities related to the goals, as well as the overall progress on the plan itself that we will be reporting back on what we've done. We're still

fleshing out what that form will look like so it's not so onerous that it detracts us from actually doing the work, but that is the plan that we will be doing that.

To your earlier point, there was a press release, but a CMS press release. This actually did get a lot of pickup which was really interesting because it actually made USA Today and so within the halls of CMS there were people that were surprised and also asking why there was such interest.

MEMBER FISCELLA: I know.

MS. JAMES: I think it just shows the opportunities we have to educate some of our colleagues about why this work still matters and is important.

MEMBER FISCELLA: And my last question is not to put you on the spot too much, but what are your thoughts about how we can help move this plan forward?

MS. JAMES: The equity plan or your road map?

MEMBER FISCELLA: Yes. I mean that we're talking about the same things here.

MS. JAMES: So in terms of the equity plan specifically, I think that there is still opportunities for input. So we have the broad buckets of what we're doing. There are specific activities under there, but if there are other activities you think we should be engaging in, and thinking through as we develop this, we did listening sessions and we talked about the different levers that CMS has and we have the levers of payment and policy, but we also have the lever of the bully pulpit. And so how do we use sort of everything in between to move forward with that.

How do we partner with people, stakeholders who are interested and that's also been good thing of how do we engage with our stakeholders as we do this. But if there are things that you say within this sphere of language access, and I know Mara has a litany of things that we could be doing, but here's what you could be doing to help move the needle. Because there are,

as I call some of those one-
year things and there are some of those four-year things. And there are improving CMS data is not a one-year thing. It's going to take us many years, but we can start laying the foundation for what needs to be done and at least at the moment, I don't see myself leaving just yet. So I commit to stay there and work on these issues.

MEMBER HASNIAN-WYNIA: So I just want to echo Kevin's enthusiasm in terms of wow, this is fantastic work. And hearing you kind of just talk about it really brought it home for me. So I just really appreciate where CMS is going with this.

I wanted to piggyback on the workforce and Lisa's comments, mostly just offer some information to you so hopefully you'll tap into it and the work group that's working on this. So at PCORI within the Addressing Disparities Program, about 44 percent of our portfolio focuses on randomized control trials, comparative effectiveness trials, looking at the effectiveness

of community health workers. So part of what we're doing right now is these are a lot of studies. There are about 30 studies, so developing a taxonomy across things such as training, you know, frequency of interaction, etcetera. So to your comment about evidence, we're really trying to develop that evidence base and some of those projects are going to be closing out in the next year, next two years. So I really hope that you tap into us, both in terms of the taxonomy, as well as some of the kind of specific projects and the lessons that can be learned from there, because we keep hearing about well, you know, we know that community health workers, there are a number of names for community health workers, but they really do make a difference, but do they really help to improve outcomes. So that's what we're really trying to hone in. So I really wanted to highlight that for you and really encourage you to please use us as a resource for that evidence.

MS. JAMES: Absolutely. And one other thing, two other things I'll say. One is that in

terms of the evidence, evidence of reducing disparities is something we're also looking for. So as we're thinking through how do we embed this into programs, we're looking for what actually reduces disparities because that's kind of the argument we're going to have to make. The other thing I know, yesterday, from the part that I heard, CMMI kept coming up in the conversations and I think it's worth just remembering CMMI's statutory charge is reducing costs as well as sort of improving quality, but at the end of the day it has to kind of reduce costs.

And I think that one other thing I would just say is not to put all of our eggs in CMMI because there are -- because of that cost of factor and equation, there are participation barriers. So not all providers are making it into those demos. And when we think about disparities, one of the things that kind of concerns us is who's engaging in these models and who is not and what is the impact on disparities for those providers who don't have electronic health records and so can't meet the

threshold of participation or we've done a better job with some of our rural providers in making sure we have models that are looking creatively at how we incorporate rural providers. But some of those who may not be rural but are just small and can't meet that threshold, what's happening to them, their patients, and the quality of care that they're receiving.

CO-CHAIR CHIN: Nancy, just a real quick question and then we'll move on.

MEMBER GARRETT: So I just wanted to mention in terms of priorities around collection of data, the factors that you have listed, I don't see anything about socioeconomic position. It was just such an important variable to consider. And like in the accountable communities for health, it's very much about trying to understand vulnerable populations from that perspective, so screening for risk factors whether it's housing and food and security and that kind of thing. Do you feel like the plan, the equity plan will include those factors or is it explicitly not there for a

reason?

MS. JAMES: It's not missing, I would say. It's our three priority populations are those, race, ethnicity, sexual and gender minorities and people with disabilities. We are an office of 23, so we can't do 11 priority populations like AHRQ, and so that's what we focused on. But CMS takes care of the most vulnerable people and so when you think about those who are duly eligible, those in low-income subsidies or Medicaid, it's not that SES is missing. It's just not the priority population that we're kind of looking at because it's so woven throughout all of what we do at CMS. And I think that's also with us. We didn't call out specifically rural because geography across CMS, as well with just Medicaid, where you live matters. So it's an underlying current of everything that we look at and why we do mapping and drilling down and looking at the rural areas that we work on. It's not missing, it's just these are the three that we opted to focus on.

CO-CHAIR CHIN: So Cara, thank you very

much for your leadership and your presence and participation in this meeting.

So Helen is going to introduce the session.

DR. BURSTIN: Good morning again. So we're going to shift gears a bit this morning and talk a little bit about the work we've done over the last couple of years and I'm very pleased you've got Dave and Kevin here at the table who were the chairs of our prior efforts to talk about what we've done to date on the SES risk adjustment. A good number of you were part of that panel, several of you in the measurement development space are living through this trial period, so thank you.

Essentially, we'll walk through today is an overview of what we've done to date in terms of risk adjustment with Dave and Kevin. We're then going to have Karen Johnson, you can introduce yourself, who is one of our lead measure methodologists who is going to help us with the evaluation approach for the trial period. And this is where we really need your input. There's

only -- this is where particularly we would really welcome your thoughts about exactly what can be evaluated and you'll hear from, particularly within the two-year time period.

As you'll hear as we introduce this, part of the agreement when this went forward after the work of the prior SES Committee was done was that NQF would allow measures to come forward with adjustment for SDS, sociodemographic status, and that was intentional to get beyond SES only as being too narrow a lens. Under certain conditions as Kevin and Dave will go through, but essentially a conceptual, logical model for why you would include those factors and empiric data. And if a measure is deemed to be appropriate for adjustment, it needs to include as part of the specifications both an adjusted model, but also an ability to see the stratified data, so there's no masking of what is kind of going into and being adjusted in the model. You'll hear lots more about that from Dave and Kevin.

And so one of the things we're

responsible for at the end of two years is to determine whether this trial should stop being called a trial and simply be built into the actual criteria NQF uses for endorsement and whether we would continue to allow measures that have adjustment for SDS to continue to be submitted, whether we would, for example, continue to require that if they are submitted that they include the stratification, so these are the exact issues you would really welcome your input on today. We've got a good amount of time to go through this.

A little bit later after we talk about the trial period and the evaluation plan, we'll also hear from a couple of our committee members. Actually, I know Christie is on the phone today, but Christie Teigland will talk about some of the initial work they've been doing on doing SES adjustment for the adherence measures that are part of Medicare Part D, just their experiences. And I think probably it would be interesting for Cara as well, since really the challenges we've heard to date are about the ability to access the needed data

and they were able to get nine digit ZIP Code as you'll hear, which was quite successful.

And you'll also hear from our very own Jose Escarce who will soon join us when perhaps it's a little bit later on the West Coast. And Jose was on the --

MEMBER ESCARCE: I'm here. I've been here on all along. You guys can't tell who is on the conference?

DR. BURSTIN: Actually, you know, the person who can tell who is on the webinar is across the table from me, Michael can.

MEMBER ESCARCE: Okay, good enough.

DR. BURSTIN: So good morning, and thank you for joining us. But then Jose will give us some background -- he's on the IOM Committee that's been funded by ASPE, the Assistant Secretary for Planning and Evaluation. So we just want to make sure, as we heard from Cara, we just want to make sure there's a good amount of cross fertilization of what's happening across the federal and private space just so you're fully

informed.

So with that, I am going to turn it over to Kevin and Dave. I'm sorry, Tom, go ahead.

MEMBER SEQUIST: Quick question. Why are there multiple efforts going on at the same time? Like I don't understand while there's an IOM Panel and an NQF Panel debating the same thing. What if they come up with different answers?

DR. BURSTIN: Right. All good questions. And that's one of the reasons we explicitly included Jose. We actually reached out to Jose because he's on the panel to make sure we at least understand. And my understanding is their charge, and again, we'll hear about this from Jose, who is more narrowly focused on the data, the SES data, what's available now, what could be available in the future. So that seems like a key piece of it, but it is not the whole package certainly. And also, ASPE does have some additional work. They couldn't present at this meeting, but they'll present to us in a subsequent where they're doing analyses at the programmatic

level in terms of understanding different approaches you could take in terms adjustment or different payment approaches to handle some of these adjustment issues.

I think we're trying our best to stay aligned, but you know, this is the reality of -- right, so those on the phone, so part of the other issues, that's purely Medicare focused and we're much broader. And Cara wanted to weigh in, please.

MS. JAMES: So the ASPE work is coming out of the IMPACT Act. So the IMPACT Act of 2014, improving Medicare post-acute care and treatment, is requiring that ASPE and CMS look at the impact of socioeconomic status as well as other demographic factors on quality ratings and scores. And so that's what's driving the ASPE work. So ASPE and CMS, we've been working kind of hand in hand on that.

There's a report that's due that the IOM Committee is focusing on that's due in three years, and there's another one that's due in five years. So the three-year one is specifically SES. And the

five-year one is looking at other demographic characteristics and they called out things like health literacy, resubmissity, and other things.

MEMBER COPELAND: Just on that same point, during one of the breaks yesterday, I asked Erin is there a document or something that exists or can be created that just kind of identifies the top five, six, seven, whatever the right number is of committees, councils, agencies, whatever, that are federally organized and are working on the same problem, just so that as we think about our work for those who are -- have got that understanding that -- have a lay of the land of who's doing what, who's narrow, who's fine. Because I think as we wrestle with SER, how we carried out work and avoid duplication and cross fertilize where it's necessary, that's hard to do if you don't have a map that kind of tells you who the players are across the federal process that are working on this and what's different, what's unique, what's duplication. Folks who work in this space all the time it's probably second-hand knowledge because

they interact with people all the time. For myself, and I would imagine for others, that's uncharted territory, so it's hard to -- you hear about these things and you kind of say to the same question, why are they doing that? Is it different than what we're doing, to really understand where we can add value.

DR. BURSTIN: I think that is absolutely the right question. And in fact, I think even for those of us who live in this space, it is very hard to figure out who is doing what. I mean our SES report came out, then Congress put IMPACT in and decided ASPE would do this work.

So I think there's been a little trying to catch up to make sure that everybody who has now been directed to do one piece of this is staying alive which is why we intentionally wanted Cara to join us. We intentionally wanted ASPE to join us, ONC. But I think it would be great for us to try to do a little bit of a lay of the land and maybe Cara has some of this, but if not, it would be a good effort we would be happy to do

collaboratively, so we have one thing that shows how all of these pieces come together. I would welcome it. Probably hang it on my wall.

MEMBER FERGUSON: I think one other thing is just the time frame that we have a narrower time frame for our committee versus like you said, the three and the five year time frame.

DR. BURSTIN: Actually, that's a great thought as well. Maybe as we put this together to do it on a time line of when deliverables are due would be especially important, I think, yes.

MEMBER NERENZ: Well, let me just first of all, I'm going to preface the preface. If you look at the slide, you'll notice that I got an unjustified, undeserved promotion in terms of degree. And Kevin, on the other hand, got dinged a little bit. So I'm not as sure of this. The way we're going to do this is Kevin is going to do the major walk through and I'll just make some emphasis points as we go along. So we'll try to do this quickly and then leave as much time as we can for discussion.

MEMBER FISCELLA: You're going to have to make that font bigger for my eyes. Can we make the font bigger so I can see the screen?

So this slide summarizes some of the basic issues that the committee really dealt with and I can assure you for those who are not on the committee this was really a complex task with very really, I think, often very nuanced and complex arguments on both sides.

I think I can read this. Okay. I'm putting my 63-year-old eyes to the test here.

And this was really set up from the beginning that when the committee was formulated that there were really, I think, very compelling and thoughtful arguments on both sides of whether to adjust or not to adjust. And on the support side, we see accurate and informative quality measurement, the idea being improved user, buy in, adjustment is necessary to avoid penalizing providers, serving vulnerable populations and communities. And this one came up over and over again, a risk adjustment allows for comparative

requirements.

A performance score alone whether or not adjusted for SDS factors can't identify -- I'm sorry, this is -- okay. This is support for adjustment, just to be clear.

Hospitals caring for are already being penalized and no evidence of disparities would be reduced to further negative financial incentives. Adjustment generally does not mask performance by providers caring for higher proportions of low SES patients and there was lots of discussion. In the end, the statisticians felt that, in general, that did not happen.

On the other side, there were arguments that some providers may deliver worse quality care to disadvantaged patients. Adjustment could make meaningful differences in quality no longer apparent, in effect, excusing worse care. Worse outcomes could be expected. No expectation to improve. In other words that the adjustment would essentially lower the bar so that there would potentially be a dual standard of care again

implies a different standard and lack of adequate data for SDS adjustment which you've certainly heard a lot about and prefer payment approach to help the safety net.

In other words, to just keep the measures the way they are and to address the problem in under resourced practices and organizations that serve essentially disadvantaged groups, fix that problems perhaps in other ways. And don't mess with the adjustment.

MEMBER NERENZ: Just a couple quick points of emphasis before you move to the next slide. On both sides of the line there are a couple of subdomains. There was a -- I'll call it a technical set of concerns about just what's good measurement. What gives you accurate, unbiased reflection of quality care which then begs the definition of what's quality of care? What's a confounder? What's within the sphere of quality? There are those considerations living on both sides of the lines here.

And then there are practical

considerations about what are the effects on the ground of either not adjusting or adjusting and it was interesting here because that group, like this group, shares the idea that disparities are bad. We want to get rid of them.

The question is what is the pathway? And does lack of adjustment add to disparities? And we have talked about some ways in which that could occur. We don't want that. But adjusting could and we talked about that. So in both of these different points of view that we started with, there were the technical and the practical and we had to wrestle with all of that.

MEMBER NERENZ: Next slide, please. This is just a background on the panel. I think most of you are familiar with the process. There were multiple stakeholders represented and there was a period of public comment. And I think we got thousands of comments. Is that right, Helen? And then those comments were revered and taken into consideration. The most ever. Wow.

Next slide, please. And so, you know,

given the task we thought it was important to start with some core principles and the first was that outcome performance measurement is critical to the aims of the NQF strategy. Disparities in healthcare, health and healthcare should be identified and reduced. Performance measurement should not lead to increased disparities in health and healthcare. And outcomes may be influenced by patients' health status, clinical and sociodemographic factors in addition to the quality and effectiveness of healthcare services, treatments, and interventions.

Next slide, please. When used in accountability applications, performance measures that are influenced by factors other than the care received, particularly outcomes, need to be adjusted for relevant differences to avoid incorrect inferences about performance. Risk adjustment may be constrained by data limitations and data collection burden. And seven, the method factors and rationale for risk adjustment should be transparent.

Is there another set of -- next slide, please. And the recommendations may apply to outcome performance measures include resource use and patient-reported outcomes and some process measures used for comparative performance assessment. So in other words, they have broad application. Each measure may be assessed individually to determine the appropriateness of SDS adjustment. That's an important one. I'm just going to read it again. Each measure must be assessed individually to determine the appropriateness of SDS adjustment. Recommendations may apply to any level of analysis.

MEMBER FISCELLA: Maybe just before we move on there, I want to emphasis the focus in our discussion on outcomes. We may choose to carry that into our own deliberations in this group. But that was very important because we recognized in the traditional spectrum of quality concepts, you've got structure process and outcome, that when you look at many outcome measures, whether these are hospital related, physician, plan, what not,

one view and it's the one I adhere to very strongly is that outcomes are a combination of quality of care and something.

Now the "and something" may be really small and tiny and so trivial you don't have to worry about it, but as the outcomes get further away in time from the clinical events and as they move say to different settings and hospital readmission would be a classic example of this. A month goes by and the patient is now out in the community somewhere.

There are these "and" factors. And it was a significant part of our discussion about how do we take that into account and what factors are appropriate to consider as outside the scope of quality of care in this "and" domain? And if there are such factors, they are probably best considered to be confounders or things that need to be adjusted, if the goal is to have the clearest possible measure of quality of care. And again, we emphasize the "if."

Now process measures and some outcome

measures don't have that same problem because there just aren't a lot of other "and" factors to worry about. If a guideline is very clear to say that in the presence of a certain set of conditions X thing should be done, the measure is pretty straight forward of patients who met those characteristics is X done or what percent of time.

And if there are concerns about the applicability of the measure to a particular person, that's usually taken care of in the denominator specifications. So once you've done that the measure is pretty interpretable. So most of our discussion is really not about process of fair measures. It was really much more focused on outcome measures.

And then within that domain, as I said, there's some outcomes that are so tight in time and so completely under control of whatever entity is being measured that there just aren't a lot of "and" factors to worry about. Now in some of my discussions and I told one of our clinicians, I think it was maybe Kevin corrected me. I used

central line infection as the example of that. The events leading to that are presumably directly under control of say the hospital being measured. I used that and then Kevin said no, wait a minute, sometimes conceptual factors can influence that. So maybe there's a better example than that.

I think it's important for this group to understand that the whole meat of the discussion was about the situation in which social or demographic factors could be in this "and" territory. And then the debate is is there an "and" territory. Then should we consider the social and demographic factors in essentially the same way that we currently commonly treat clinical risk factors.

I think people are well aware in most of these measures, we do clinical risk adjustment all the time. It's routine. You've got HCC type adjustment. You've got comorbidity, another form of adjustment. So we say well, you get adjustment if the patients are older, if the patients have more comorbidity, if the patients smoke. We're used to

that. And a lot of our discussion then is should we think about some social and demographic factors essentially in the same way and ultimately treat them the same way. And I think with caveats that we'll hear about in just a second, the group concluded that yes, that makes sense.

CO-CHAIR CHIN: Also, one point, David, the ultimate recommendation there was also some process measures because there was a lot of discussion about there's quite a few process measures where you should probably adjust, like mammography screening, for example. There's going to be a whole variety of patient factors that --

MEMBER NERENZ: Maybe I just drew this too tightly now, so I guess from the outcome you sort of work back. If there's reason to think and then we're going to get into the text of the recommendation in a second, if there's reason to think that a social or demographic factor might influence the outcome in a way that it behaves like a confounder, then it may make sense as well. But

I still think I was on solid ground saying most of our emphasis was on outcomes. That's where these issues are most salient most of the time. Is that -- okay.

CO-CHAIR PONCE: David, one of the core -- and I was on this and I don't remember this, but the core principles was that risk adjustment may be constrained by data limitations and data collection burden. So that's sort of a constraint that was kind of the reality, but I'm not sure -- I thought our core principle was to try to promote more data collection.

MEMBER NERENZ: I think both of those are correct. I think that principle is just simply a statement of the realities of the world that you can say we'd like to adjust for factors X, Y, Z, and Q, but if in a particular data set used for the measurement in a particular program X, Y, Z, and Q don't exist, you can't do it.

Now five years from now we could say that X, Y, Z, and Q ought to be in there and I think that was -- so we were just saying that any one point

in time the ability to do this is going to be limited to data available. But now more in the spirit of this group, I think, we would be pushing for the appropriate data to become available if we could.

CO-CHAIR PONCE: Thank you.

MEMBER FISCELLA: Okay, so these are the recommendations of the committee.

Recommendation 1. When there is a conceptual relationship, i.e., a logical rationale or theory, between sociodemographic factors and outcomes or processes of care, and empirical evidence, e.g., statistical analysis that sociodemographic factors affect an outcome or process of care reflected in a performance measure, those sociodemographic factors should be included in risk adjustment of the performance score using accepted guidelines for selecting risk factors unless there are conceptual reasons or empirical evidence indicating that adjustment is unnecessary or inappropriate; and the performance measure specifications must also include specifications for stratification of a clinically-adjusted

version of the measure based on the sociodemographic factors used in risk adjustment so that one can see the disparity.

Next. Recommendation 2. NQF should define a transition period for implementation of the recommendations related to sociodemographic adjustment. During the transition period, if a performance measure is adjusted for sociodemographic status, then it will also include specifications for a clinically-adjusted version of the measure only for purposes of comparison to the SDS adjusted measure. So one can see both the unadjusted and adjusted measures.

Recommendation 3. And David, if you want to interject on any of these, feel free or we can wait until the end.

MEMBER NERENZ: I want to say a little bit about one, but why don't we go all the way to the bottom. I'll come back to that. One is sort of the main --

MEMBER FISCELLA: Okay. Yes, yes. A new NQF standing committee focused on disparities

should be established. Welcome. All the way through, okay.

Recommendation 4. The NQF criteria for endorsing performance measures used in accountability applications, e.g., public reporting, pay for performance, should be revised as follows to indicate that patient factors for risk adjustment include both clinical and sociodemographic factors.

2b4. For outcome measures and other measures when indicated, e.g., research use and some process measures. An evidence based risk adjustment strategy is specified is based on patient factors including clinical and sociodemographic factors that influence the measured outcome and are present at the start of care and has demonstrated adequate discrimination and calibration or rationale data support no risk adjustment. Risk factors that influence outcomes generally should not be specified as exclusions.

MEMBER NERENZ: I just emphasize that italicized text here is actual literal NQF policy

text and these are changes in -- word-by-word changes we recommended.

CO-CHAIR CHIN: Wait, say that again?

MEMBER NERENZ: The italicized text here that you see, it starts with the 2b4, this is text that exists in NQF a policy document. What's its name exactly?

DR. BURSTIN: It's in the measure valuation criteria document.

MEMBER NERENZ: Okay, and the wording before we did our work was slightly different than this. We don't have the strikeouts included here. This is the change. So the non-italics is sort of explaining, but the italicized text is new NQF document wording that we recommend.

MEMBER YOUDELMAN: Was it adopted or not yet?

DR. BURSTIN: Yes, oh, yes. This is now actively --

MEMBER YOUDELMAN: It was recommended, I just wanted to be sure of that.

DR. BURSTIN: Yes, yes. Then we'll go

through on what we've done so far on that.

MEMBER FISCELLA: Just to remind everybody and Helen can correct me. I think it was 2006 where NQF adopted a policy prohibiting adjustment and so this text comes from that policy.

Next slide, please. Recommendation 5. The same guidelines for selecting clinical and health status risk factors for adjustment of performance measures may be applied to sociodemographic factors, and include the following: clinical/conceptual relationship with the outcome of interest; empirical association with the outcome of interest; variation in prevalence of the factor across the measured healthcare units; present at the start of care; is not an indicator of characteristics of the care provided (e.g., treatments, expertise of staff), sort of in that causal pathway; resistant to manipulation or gaming; accurate data that can be reliably and feasibly captured; contribution of unique variation in the outcome, not redundant; potentially, improvement of the risk model (e.g.,

risk model metrics of discrimination, calibration); potentially, face validity and acceptability.

Next slide.

MEMBER YOUDELMAN: Is it all of the qualities or some of the qualities?

MEMBER FISCELLA: All of them. Thank you.

MEMBER BERNHEIM: Just clarifying, the committee didn't come up with this list. The committee's recommendation was only that when you think about SDS factors, you should think about them using the same guidance that is used for clinical factors which was already established by NQF.

MEMBER FISCELLA: Next slide, please. Recommendation 6. When there is a conceptual relationship and evidence that sociodemographic factors affect an outcome or process of care reflected in a performance measure submitted to NQF for endorsement, the following information should be included in the submission: A detailed

discussion of the rationale and decisions for selecting or not selecting sociodemographic risk factors and methods of adjustment (including a conceptual description of relationship to the outcome or process; empirical analyses; and limitations of available sociodemographic data and/or potential proxy data) should be submitted to demonstrate that adjustment incorporates relevant sociodemographic factors unless there are conceptual reasons or empirical evidence indicating that adjustment is unnecessary or inappropriate. In addition to identifying current and planned use of the performance measure, a discussion of the limitations and risks for misuse of the specified performance measure.

Next slide. 7. NQF should consider expanding its role to include guidance on implementation of performance measures. Possibilities to explore include: guidance for each measure as part of the endorsement process; guidance for different accountability applications, e.g., use in pay-for-performance

versus pay for-improvement; innovative approaches to quality measurement explicitly designed to reduce disparities.

Recommendation 8. NQF should make explicit the existing policy that endorsement of a performance measure is for a specific context as specified and tested for a specific patient population, e.g., diagnosis, age;, data source, e.g., claims, chart abstraction; care setting, e.g., hospital, ambulatory care; and level of analysis, e.g., health plan, facility, individual clinician. Endorsement should not be extended to expanded specifications without review and usually additional testing, so that the measure is used as it was intended for.

Next slide. 9. When performance measures are used for accountability applications such as public reporting and pay-for-performance, then purchasers, policymakers, and other users of performance measures should assess the potential impact on disadvantaged patient populations and the provider/health plans serving them to identify

unintended consequences and to ensure alignment with program and policy goals. Additional actions such as creating peer groups for comparison purposes could be applied.

Next slide. 10. NQF and others such as CMS, Office of the National Coordinator for Health Information Technology, and the Agency for Healthcare Research and Quality should develop strategies to identify a standard set of sociodemographic variables, patient and community-level, to be collected and made available for performance measurement and identifying disparities.

Next slide. Okay.

MEMBER NERENZ: If you could just run back to Recommendation 1, please. It's five, six slides back. All the recommendations are there for a reason. We talked through them. They speak to -- through different domains, but the first recommendation I think is the core one. It's the one that represented or at least recommended a very significant change in NQF policy and I guess by

extension the CMS and others as well.

When our group formed, it was actually an interesting dynamic. It was very much like this one. We had a conference call and then we met in this very room and we were sitting around this very table and a lot of the same people. In fact, I think Susannah was more or less in the same seat. There's a reason.

And I think it's important for this group today to note that the group like this one came from a variety of perspectives by plan and intention. We had some measure developers. We had some providers. We had some purchasers. It was an intentionally diverse group. And I think there was a mix of opinions starting about this issue of should there or should there not be adjustment.

And that first slide we showed about the different views, I think those were represented around the table. In fact, some of us came with both that whole set of thoughts jumbled around in our own heads because -- because they're valid

concerns.

So the general consensus in the direction of this positive recommendation about adjustment emerged probably mid of our first in-person day. There were still a couple of folks with very strong reservations of the consensus in this. And I want to single out Susannah in the most positive way I possibly can. I've done this before, so she knows. Susannah did an amazing, wonderful job of sort of a bridge across these two sets of considerations because what we came up with we think is sort of careful and nuanced and conditional. And it says under a certain set of conditions, adjustment would make sense, but that doesn't mean that adjustment always makes sense.

And I think we're comfortable moving forward with that and Susannah and I thank you for your unique contribution sort of defining that. And although the words don't appear here because we wrestled so much, there's some underlying concepts here of things like variance explanation and R squared and signal and noise that we struggle

with how to word it correctly and we ended up backing off a little bit. But we considered wording like, for example, if the outcome is predominantly influenced by social factors and then we couldn't quite figure out what does predominantly mean, but I think the concept was still floating around in there.

In my own mind, I tend to think of the kind of box scenario diagrams that people who do path analysis use where you've got coefficients leading from one thing to another thing. And the coefficients are either really little or they're really big reflecting the strength of the causal path. And to this day I think that that's part of what we were trying to get at here that if what you have in front of you is a strong -- well, not strong, who knows how strong -- a signal about quality and then you have some other factors that represented in some ways noise, adjustment is about getting rid of the noise so that you can see the signal clearly.

The empirical question is how -- what's signal and what's noise? That's the conceptual

part. And then the empirical part is how big are the relative contributions? And sometimes there's so little noise you don't have to worry about it. Sometimes there's so much noise, there's hardly any signal.

But I did want to credit Susannah for helping us sort of bridge across this and say that if we make this recommendation conditional, we may be at a point that at least most of the group could feel comfortable because there was a yes component to it that really was a major significant change for NQF. But it's not blanket, unconditional adjust all the time.

DR. BURSTIN: Ron.

MEMBER COPELAND: I think your third recommendation was that this group be established. So it says the standing committee focused on disparities should be established, but it didn't say for what purpose, so I'm curious when that recommendation was made from that body, why did you make that recommendation? What was the intention?

MEMBER NERENZ: Well, and I think the

main thing we had in mind was that this group could oversee the effects of this recommended policy change if adopted. Now at the time that we wrote the report and made the recommendation, we didn't know what the NQF Board would do with it.

MEMBER COPELAND: Right.

MEMBER NERENZ: The NQF Board could have rejected it flat out. We didn't know.

MEMBER COPELAND: Right, right.

MEMBER NERENZ: Now that's not what happened and here we are. But I think the idea was that if a change like this occurs, there should be a group who watches what happens, particularly through the lens of disparities. Now there may be other lenses with which you can evaluate the effects of it. There could be highly technical lenses, but I think we thought it was important since so much of the concern was what is the effect of adjustment or lack of adjustment on disparities? We thought a group with disparities in its title should be looking over what happens. And then judging, eventually, are good things or bad things

happening?

MEMBER COPELAND: So predominantly a monitoring role is what you had in mind?

MEMBER NERENZ: Yes.

MEMBER COPELAND: For unintended consequences.

MEMBER NERENZ: Yes, or intended consequences.

MEMBER COPELAND: Intended consequences.

DR. BURSTIN: It was actually also Marshall, I think, and I just want to credit Marshall with a lot of this. Marshall had done so many of our committees over four years and I think his feeling was and I think we agreed was why we embraced this recommendation; in particular, that we couldn't keep doing these one-off disparities efforts, that to really make real progress, we had to have a group that had more of a longitudinal, cross sectional view of NQF, all of our work on the measurement enterprise. So that was really a big part of this. So every time we talk about trying

to find this, we're just going to do it. I can hear Marshall going you have to have this cross-sectional longitudinal group. So thanks to Marshall.

CO-CHAIR PONCE: Emilio, Romana, and Mara.

MEMBER ESCARCE: Can you put me in the queue as well, please? This is Jose.

CO-CHAIR PONCE: Jose, yes.

MEMBER SANCHEZ: Well, I really have to just commend all of you that have worked on this. This is really ground breaking. It's really a very difficult topic and it was handled very intelligently. Building the bridges and being persistent, I mean all the elements that went into it, it's a terrific piece of work.

My question is at first blush, what's the initial feedback from CMS and other bodies that would be affected? Just wondering.

DR. BURSTIN: I think you'll hear some of that when we start going through the trial period and the measures submitted to date. So let's hold

that if we could.

Romana?

MEMBER HASNIAN-WYNIA: So you know, when the recommendations came out and I read them, they just intuitively made so much sense just on kind of first reading. So I really appreciated that.

My question is in terms of kind of the no absolutes, not everything has to be adjusted, not every measure has to be adjusted and what I remember from at least the executive summary where there were some examples of which measure should be stratified versus which measures should be adjusted.

I'm curious about the 800 plus or so public comments or maybe even discussion within the committee around was there tension around the types of measures where -- I'm curious about the guidance. So if I were in the delivery system and the guidance was not everything has to be adjusted, but here are some examples of what should and what shouldn't, was there more guidance? And was there

tension in that and I'm curious about the public comments related to that.

MEMBER NERENZ: Well, I'll let others answer. I'll just speak to the public comments. I don't recall the public comments being very much about specific measures or nuances in measures. The public comments, I think, were either yes, great, you did the right thing, and just numerically those were vastly -- I think we counted the organizations. There was something like 150 organizations and it was 143 to 7 if you did like the score board. But the seven were not trivial. The seven included CMS, NCQA, Consumer Reports. So you have to kind of decide how you want to weigh it. But that was the initial comment during a three-day window on the draft report.

Now, about a year and half or so has gone by since then and some additional thinking has gone on, but that was at least. But I don't recall those comments being much about nuances or we think you're on the right track here, but not here. It was just pretty much are you on the right track

period or are you on the wrong track was my recollection of that.

MEMBER FISCELLA: We did grapple with this issue and I think it was -- it's a really difficult issue of could we provide more explicit guidance and we did make attempts at that. I think in the end we felt comfortable with where we went. I'd be interested in what Susannah's -- I think Susannah was very much a part of that, of those complex discussions.

MEMBER BERNHEIM: Do you want me to speak generally to the question of which kinds of measures? MEMBER HASNIAN-WYNIA: I was just wondering whether there was -- I'm thinking about the impact of this report, so if I were reading it and I was sitting at a large delivery system thinking and I was focusing on disparities and I was really thinking about this dialogue around risk adjustment versus stratification, so there's some general guidance. But a lot of what happens in delivery systems is just lack of technical skills to even know what to do.

So my question was more just directed to was there tension around providing specific guidance and were there certain measures where it was very clear that stratification would make sense and risk adjustment wouldn't? Did the committee grapple with or recommend very specific guidance even by way of example?

MEMBER BERNHEIM: So I think we didn't and I think part of that has to do with how much you can accomplish in two days with a really diverse, engaged group and then follow up afterwards. Also, there's some of the details that take a lot to take consensus around. And I think, in general, there was a strong feeling that measures differ. The way that you think about socioeconomic and race factors playing into measures really do depend a lot on the kind of measure. And so it was going to be challenging to sort of numerate the best things.

MEMBER HASNIAN-WYNIA: I'll make just a -- there was a lot of discussion and I'm pretty passionate about it and spend a fair amount of time

thinking about it which doesn't mean I'm always articulate about it, but I don't want to take a lot of time, but I'll just name a couple of quick things to help put it in context.

So that committee, as opposed to this one, was constrained by the request, right? So we sat around this room dealing with the same issues this group has, but our box to work in was risk adjustment of quality measures, right? Guidance sends you up on risk adjustment of quality measures. It was not payment policy. So that was a tension in the room and I think that's where a lot of my concerns come out because we were sort of stuck making a decision about what you should do about risk adjustment and quality measures which is not where I think the right lever is.

That said, the committee worked really hard and was really thoughtful, but I really give NQF a lot of credit for putting themselves behind this committee and expanding the scope because I think it gives an opportunity to change the conversation and that wouldn't be happening and

this is really their work. So I think that that's a really important thing.

I have a lot of reservations about putting patient level factors into the measures that my team works on which are outcome measures and it comes to the thing that David said which is what you're trying to do when you look at an outcome is parse out the part that is sort of inherent to the patient when they walk in the door which SES and race feel like they are that might affect the outcome and quality.

And as everybody sitting around this room knows, race and socioeconomic status are deeply enmeshed with access to high quality care in this country and so every analysis we do that tries to take apart sort of what how much of this is the patient and how much of this is quality says both are happening. And so if you take that coefficient and put it into your model, you're risking carrying the quality piece along and lessening the quality signal. And if you don't, the concern that everybody has is that you're

risking not accounting for the fact that some of that risk is outside of the providers' control and they may look worse on this measure.

So there's just not a simple solution to this, particularly if your only choice is do I put this in the model or do I not? So I think we can be more creative than we're trying to be. But that's where my reservation comes from and I always bring race to the table because when you sit in front of the CSAC, they say, we're not talking about race and I say why are you not talking about race? Because we're worried that actually there's nothing inherent about race that should make my risk of being readmitted higher because we know that that has to do with communities of color having access to poor quality hospitals. But when you talk about SES people say oh, well throw that in the model. And I'm like do we not think that communities of poverty have access to lower quality hospitals? I mean they're not the same.

There are very important differences between SES and race, but if we pretend that one

is simple and the other is simple and they're not deeply enmeshed with quality in this country, this is a hard conversation and I will stop my soapbox now. I will put it back in the box and I'll walk away, but this is a real struggle and I'm really interested in other ways of handling this that, I think, are more effective at protecting the safety of hospitals and less likely to ignore and sort of just build into our models disparities in outcomes. Not disparities in quality, disparities in outcomes, sort of changing how we think about what should be the expected outcome if you're a poor patient walking into a hospital. Okay, done. There you go. That's where the tension was.

MEMBER NERENZ: If I could just highlight and I think what the very last thing you said is just crucial to our continuing discussion because it's easy to just say we're interested in disparities, but I think that sentence always has to continue further and say disparities in what? And a disparity in outcome is not always the same as a disparity in quality and whatever discussion

we're having, we need to be as precise as we can about what disparity exactly we want to see, which one do we want in public discussion and how do the two relate to each other. So just to emphasize that.

MEMBER BERNHEIM: Right, so this relates to these conversations about visibility. We got a little bit bogged down in this visibility because some of NQF earlier language, which I'm afraid we influenced and I think actually confused the picture, said you wouldn't want to risk adjust because you would hide disparities and the concept there was if you put SES risk adjustment into a model and in general hospitals that care for poor patients had worse outcomes, but you adjusted them, you would no longer see that those hospitals had worse outcomes. They would suddenly look on your sort of adjusted measure closer to other hospitals.

So the concern that had been expressed that had led to some of the early NQF guidance was that risk adjustment would make disparities between the hospitals different, but it had been

interpreted as making disparities in quality invisible. And the truth is the current measures, which is part of what this committee is going to help us do don't make disparities and quality visible, except to the extent you look at hospitals.

CO-CHAIR PONCE: Can we hold off on that because we have a queue. I think that's relevant for our Disparities Standing Committee. So Mara, Jose, Traci, Nancy. Now Mara.

MEMBER YOUDELMAN: Well, I'm confused and also scared a little bit now that this committee has a pretty big charge and to try to do it all remotely with a couple of webinars a year is going to be really tough if given what you guys have just said about the idea of the standing committee. So I don't know how that gets fixed, but I think given just the depth and breadth of these issues it's going to be hard for us to sort of achieve some of that and so I won't put people on the spot.

I wish there was a way and maybe we could help with that, but look at ways to have longer

meetings, more in-person meetings something like that. Because otherwise, I think we're named as the Disparities Standing Committee and looked to to be the -- not the be-all, end-all, but have a really strong role.

And we're not going to be able to sort of live up to it and it's going to be to the detriment of, I think, NQF and CMS and a lot of the other quality measures whereas if we had more quality measures where as if we had more ability to grapple with some of this it would help, but that wasn't really my question. I'm sort of trying to figure out when this is implemented, so maybe it's a question for the pilot, but maybe it's not.

Did the committee discuss, figure out at what point you have enough of a population within an entity to risk adjust these measures? So if, for example, a measure is in for risk adjustment, when it gets applied, does it then depend on how many of X group are in this population? So if it's a hospital measure, that the hospital has to have a certain percentage of people of a certain race

in order to then risk adjust it or is it like if it's person, that's enough?

I'm trying to figure out sort of like the population factors and then how it doesn't get impacted because you might have a couple of outliers one way or the other and if that's something that was discussed and how it sort of came out.

DR. BURSTIN: It was probably discussed in the context of stratification, more so I think than risk adjustment, that at times the cell sizes would be too small, but others should weigh in.

MEMBER FISCELLA: On adjustment, we didn't discuss it in the context of adjustment, more on stratification due to cell sizes.

CO-CHAIR PONCE: Jose, you're on.

MEMBER ESCARCE: Yes, I think it was probably Helen. She's probably answered a lot of my question, but I'll go a little bit further. What I was going to ask is any adjustment of either quality measures or resource use measures have to

use data that reflect what's going on now. And so for example for resource use you can imagine a situation where disadvantaged patients don't increase resources, but that's, of course, because of the incentives inherent in the current system and so providers don't spend more resources on them and the consequence of that may be that they do worse than I'll just say "they needed to" and so adjusting the quality measures would be over adjustment and that's exactly what I think Helen if she was the one who was speaking was saying.

So my question isn't about that because I think that's pretty clear and Helen mentioned that people had thought about and had hard conversations about it.

My question is a little more specific. Were there discussions by the committee on exactly what to do about that? So adjusting the measures in a different way or adjusting them and then sort of for reporting purposes only, I mean I'm making this stuff up now, only going half way between the unadjusted and the adjusted. I mean there is a

million -- not a million, but a few things you can think of, so after them recognizing this is both a conceptual and an empirical problem, were discussions about actually what to do about it?

MEMBER FISCELLA: I was going to say I think the short answer is we didn't really get into that level of specificity.

MEMBER ESCARCE: Okay, that's good enough.

MEMBER NERENZ: Just to emphasize that and Susannah had made the same point that we took up as much of our charge as we could in the time and mental effort of everybody available, but we recognize there's a wholly immense territory where you start going measure by measure or domain by domain and you start thinking about is this model structure better than that model structure? Is this set of variables better than that set of variables?

MEMBER ESCARCE: I wasn't thinking about that. It was really just a question to think about a methodological approach, just even talk

about it, how you might deal with the fact that you might be "over-adjusting" if you just took the data that the world is giving you now because the world functions in a particular way. And it's a very difficult question. I can't imagine you had a right answer anyway. I don't know if there is a right answer. I just wondered if you had thought about it.

CO-CHAIR PONCE: Noted, and I think that this -- again, the impetus of forming this committee was to look at that robust trial and have more thoughtfulness about what to do with these different adjustments.

On this point, David? Okay. Traci, then Nancy, then Tom, and then Eduardo.

MEMBER FERGUSON: From our payers' perspective in response to this and also Christie and I, through email distribution, did send out the Inovalon study which again looks at Medicare Advantage beneficiaries from over 81 MA contracts. And it did focus some of the confounding differences.

And Christie, I'm going to take my notes because I did review that study. But looking at the plan all cause readmission, there was an over 81 percent disparity and over 54 percent was attributed to 17 chronic conditions and over 27 percent of this disparity was attributed to three sociodemographic factors. And it did show that there is again by pulling all of this data from multiple health plans that there is confounders. And I think that is one of the things that we're hopeful with this committee is that looking at some of the existing measures that are out there, what are the ones that we could, when they come up for review for endorsement, we would sort of in the sense maybe recommend to developers, they should either risk adjust or stratify. But I think that's where the importance of where we come in is to identify those areas where there's gaps and how we could do that in a way, since we're not the measure developers, how do we get that information to those who are developing these measures and how we review them.

CO-CHAIR PONCE: So talking about a dissemination plan, possibly this committee could spearhead that.

Nancy, who is also on the SDS Risk Adjustment Panel?

MEMBER GARRETT: Yes, I just wanted to add that I feel like when this report and the recommendations are discussed that recommendation 9 gets missed a lot. I don't know if you'd mind going to that slide. And it may not be worded in a way that's very clear to everyone. But to me, what that recommendation is saying is that risk adjustment alone is not going to solve the problem of disparities. It's not going to close the gap.

And sometimes I feel like people use this report as well, we'll just throw in a few variables. We'll do risk adjustment and then everything will be fine. There's a lot more that we have to do. And I think that's what we're trying to capture here. We're just so early in having the data we need to really risk adjust for these a lot of unmeasured factors around sociodemographic

variables.

We're early in the methodology. How do we actually do it in a accurate way? And so we really need to still be paying attention to these major gaps and the resources we're investing in vulnerable populations. So again, Minnesota, I think I shared the legislation with all of you that we worked on and we are stratifying measures by sociodemographic factors. We're looking at risk adjustment, but we're also looking at a payment enhancement for the safety net to try and address these disparities and get resources into the right place. And so those three things work together, but they're not necessarily the same. The payment enhancement might have nothing to do with a traditional risk adjustment method. It might be some new creative way. So I just wanted to highlight that. I think that's important to take in mind as well.

CO-CHAIR PONCE: Thanks, Nancy. And I think recommendation 9 gets a bit of what Jose was asking as well.

So Tom, and then Eduardo is going to be the last so we can move on with the agenda.

MEMBER SEQUIST: So I guess I have two questions or comments. One is that I still feel it would be more helpful just for lingo to understand what we mean by quality versus outcomes. I have sense of what you mean is that quality is a process measure and outcomes is a --

CO-CHAIR PONCE: I'm not sure how that got implied. Quality is the broad construct. It includes structure, process, and outcome. So that's how we view it. It's all of those.

MEMBER SEQUIST: Okay.

CO-CHAIR PONCE: I think the issue is that we had to bait in this report about whether these recommendations really applied only to outcome measures or would you also potentially adjust the process measure which didn't seem quite as obvious.

MEMBER SEQUIST: I guess what I was hearing a lot was folks saying and I don't know if it was Susannah or somebody else was saying you

know, if we adjust for these things on outcomes measures, and we don't adjust it for quality and it definitely sounded like we were putting them into buckets.

MEMBER NERENZ: I'm the guilty party here and I apologize and it was just talking fast. When I used the word quality and distinguished it from outcome I should have said process quality and I would have been much more accurate.

MEMBER SEQUIST: Okay.

MEMBER NERENZ: And I think that helps. And I apologize for the confusion. There are some other nuances we can still get into. I'm trying to keep us moving along. I think but --

MEMBER SEQUIST: No, that's helpful.

MEMBER NERENZ: I don't think I spoke completely wrong. I just didn't speak quite precisely enough.

MEMBER SEQUIST: Okay. My main question was I'm interested from the folks who are on this committee which is really great. I think when it came out we were partners, really excited

and supportive of this.

I really still have a lot of concerns about how we will go about ensuring that the data, the actual patient variables that we're collecting are done so in a consistent way and how we make sure that it's -- because we're going to be using this to compare presumably across hospitals and doctors and such. And these data are so much less standardized than an ICD-9 diagnosis for acute MI in the emergency room. And I'm wondering how much that came up.

And you know, if the role of this group here is going to be to provide guidance around that, not only sort of what should be collected, but the how to collect it. Marshall joked that he always sort of thinks about the collection of race and ethnicity data and the work that you had done with how do you collect. And that was so much work that went into that and it's still so hard to get accurate and consistent race data and to picture sort of asking our patients what was your income last year? What's your housing stability?

And then I wonder about the frequency with which that stuff needs to be assessed. So once you've assessed accurately that your patient self-reports their race as black, I'm not sure that's as big a deal in terms of frequency updating as you collected your patients' income or their housing stability or things like that that change with lots of frequency. So I'm just wondering if those conversations came up.

And if there's a role for this group in guiding that conversation, it's not really about the science of whether we -- at what level we put these into our models or whether we just process measures or outcome measures, but it's sort of the actual variable itself and how it's collected.

MEMBER FISCELLA: Let me respond and then maybe other members can weigh in. I think that there was a thought that we didn't want the perfect here to be the enemy of the good and that we needed to at least get the process going, even with very imperfect data that at the end of the day may not always have more than a trivial effect on

the amount of variance explained.

And at the same time, doing so may actually encourage the reliable and systematic collection of all of the variables that we need to account for this. And that as long as there was a prohibition against any sort of SES adjustment, sort of the idea of even collecting SES is a little bit marginalized. So if we could at least put that in the forefront that that would push it ahead and in fairness, this would be a role for this committee to continue to push the issue of collection of reliable and systematic data.

CO-CHAIR PONCE: I think that is important to push, but even among population survey developers there's inconsistencies in measures, but there is at least more of a dialogue in getting harmonized measures, but then you don't want to get it stuck in a harmonized measure that's wrong, too. But I think there is a role for this committee.

And Eduardo, are you going to --

MEMBER SANCHEZ: I'll be quick. I think this goes back to the comments that Nancy was

making on recommendation 9. As I think about the sorts of reporting that I've seen about the safety net hospital in Dallas, Texas, what sometimes is not part of the conversation, and maybe this is covered and people just don't know it, but what's not part of the conversation is the revenue or the dollars expended or the dollars received per patient.

In other words, the SES demographics, if you will, are characteristics of safety net hospitals vis-a-vis the funding they have to do the things that they're doing versus another hospital, never mind the SES demographic challenges of the patient population, but the safety net hospital in Dallas. And I suspect this is the case in many communities is underfunded. It's under resourced on a per patient basis. And I wonder if that is -- is that part of what we're trying to capture in recommendation 9? Is that something that we should at least put on the table that it's not enough to risk adjust for SDS of the patients because the resources with which different

hospitals are working may also be very different.

And so when you combine SDS challenge with an under resourced system, you end up with outcomes that don't necessarily fare well. You may attribute them all to SDS when, in fact, it may be because the system is underfunded. And it's not just that it should be what was the word, that there should be an enhancement. It really -- an enhancement almost sounds like we're giving you extra money. The language we might think about using is leveling the playing field.

So I just bring that up as maybe something that we want to be sure. The idea is to improve health outcomes. Just risk adjusting for SDS gets us part of the way there. We also need to think about the other side of that equation which is the resources being used to address the health issues in various systems.

MEMBER FISCELLA: To answer your question briefly, it definitely came up repeatedly, the issues that you raised. I think yes, we tried to capture here under recommendation

9 and I agree with -- I personally agree with the comments about leveling, leveling the playing field in terms of appropriate funding relative to need.

CO-CHAIR PONCE: I said Eduardo would be last, but Michelle.

MEMBER CABRERA: Thank you. You know, I think just sort of building off of Eduardo's comments, there's a flip side to this coin and I'm going to be really sort of crude in how I describe this, but I think in the safety net in having to deal with patient populations who come in the door with significantly more challenges, right, they've also developed a capacity and an expertise and an ability that I don't think is getting fully captured. So when we say safety net is just over quality period, end of story, I take a little bit of issue with that because I think it also depends on what you're measuring and what criteria you use to assess quality.

So if you plunked the same patient population at the doors of a provider who primarily

doesn't deal with safety net population, how prepared are they? What abilities do they have to provide high-quality services to this patient population? Do they have the interpreters? Do they know how to deal compassionately with some of the challenges?

And I would say safety net often falls short on some of these things because of those resource constraints and sort of just they overwhelm, right? But they also rise to the challenge day in and day out in other ways.

So I think somewhere in this conversation, I'd like to have us consider how -- what kind of the established core set of quality measures are -- may sort of also create this perception problem about what is quality vis-a-vis a safety net population. I don't know, maybe somebody is already doing that.

CO-CHAIR PONCE: Noted. And thank you. I also wanted to thank Dave and Kevin and Emilio is looking at me very -- on this point. No? Okay.

So thanks to Dave and Kevin for presenting and also for leading that committee expertly, masterfully as Marshall said. The other complimented them yesterday. It was speed and rigor, I think, exemplary at that, and we're very fortunate to have NQF invest in the Disparities Standing Committee and that there's a few of us who were in that -- that SDS risk adjustment that could help with the continuity and help fill in some of the holes in addressing the tension so that we -- we lift some of the tension.

Okay, so Susannah would like to have a clarification on her comment.

MEMBER BERNHEIM: I just want to make sure that the comment I made earlier about communities of color and communities of poor patients often having access to worse quality care didn't get misinterpreted because it sounds like a little bit of sort of like safety net hospitals aren't doing a good job.

So one of the remarkable things we find in our measures is that, in fact, despite all of this

concern, many safety net hospitals do remarkably well in this measure, on our outcome measures, despite having no doubt more challenging populations to get good outcomes in, worst funding, right? So no way disparaging many high-quality safety net hospitals, just reflecting that when you see there's an increased risk of a poor outcome for patients from socially-disadvantaged circumstances, embedded in that increase risk is that there's pretty strong evidence that there are places in this country where they are more likely to have lower quality care and that you have to be careful how you disentangle that increased risk and how much you assume it's inherent to a patient versus modifiable.

In fact, the success of some of these patient safety net hospitals shows how well you can do even with those populations. I just don't want to have gone on record saying safety net hospitals are providing bad care.

CO-CHAIR PONCE: All right, so let's look at some data, and the recommendation was a

robust trial, and so we'll be looking at data in the next two presentations.

Karen Johnson, do you want to introduce yourself, please?

MS. JOHNSON: Sure. Hello, I'm Karen Johnson. I am one of the senior directors here at NQF and I have the privilege and sometimes the scariness to help oversee our trial period for this SDS adjustment along with many of my colleagues.

So we wanted to go ahead and let you know what we did. So the SDS panel did all this great work, a year and some change ago, and NQF took it to heart, so let's go the next slide.

Our Board approved a two year trial period regarding SDS adjustment. So as has already been mentioned, we used to have a prohibition saying you're not allowed to do it for two years. That prohibition has now been lifted, so that's what we mean by our trial period. And during this trial period, as the panel recommended, if adjustment is determined to be appropriate for a given measure, NQF will endorse one measure with the

specifications to compute that SDS adjusted measure but also have this clinically-adjusted version only so that stratification could take place.

Next slide, please. We took the recommendations of the panel very seriously when we designed this trial. So none of this stuff I think will be a surprise given Kevin and David's discussion just now. Each measure must be assessed individually to determine if SDS adjustment is appropriate. Not all measures should be adjusted for SDS factors. And the recommendations from the panel apply to any level of analysis.

So going back a little bit to the discussion about process measures, outcome measures, et cetera, all of those are on the table per the panel's recommendations, and each measure will be considered.

Next slide, please. So when did we start the trial period? Well, our official date was April 15th, tax day and SDS trial day. We had to pick a date because, and we'll get into a little bit more, but just because NQF said our trial starts

today, doesn't mean that folks were really ready to do that, right? Rethinking risk adjustment approaches is not something you can do kind of turning on a dime, so we had to take that into consideration, and we also had some internal processes that we had to take care of before we could even do that. So April 15th. All measures submitted are part of the trial. So everything that comes in for evaluation, newly submitted measures that we've never seen before, previously submitted measures that are for maintenance, those with conditional endorsement, and I'll talk about that in a few minutes, as well as measures undergoing ad hoc reviews for this particular question would all be considered. They're all part of the trial. So it's not that only a few measures are in the trial. All measures coming forward for these two years are part of our trial. But clearly not all will be -- have risk adjustment, and not all will be adjusted for SDS factors.

Next slide, please. So I did want to give you a little flavor of some of the

infrastructure that we had to do to support the trial, and we spent a few months doing this. We had a lot of communications with our external stakeholders including various -- a number of briefings. We had a very well attended breakfast meeting at our annual conference, standing room only, to talk about rolling this out. 7 a.m., yes, people were interested.

This is a big change as I'm sure you all are very much aware. We put out a frequently asked questions document to try to explain what we're doing. We had to make some additions and modifications to our measure submission forms to make sure that we're asking the questions in the way that we needed to to get the information that we need to go forward.

We spent quite a bit of time coming up with some guidance and also training for our developers. Again, this is new, so we had a lot of one-on-one talks and discussions with various groups to try to prepare them for what was coming down the pipe.

We also had quite a bit of consultation with some external statistical experts to help us as staff understand what we should be seeing when things come in the door and how to look at those. And then finally, we have done some thinking, and hopefully you'll help us do more thinking, about our evaluation plan, what we're going to do in two years.

I would just point out that while we had to basically develop some guidance, et cetera, for this, we did not have to change any of our evaluation criteria. So just so you know, part of our scientific acceptability criterion that looks at reliability and validity of measures that -- we've always had this. The validity piece of that already looks at risk adjustment approaches. So again, risk adjustment comes under what we would consider questions of validity, potential threats to validity. So we didn't have to change our criteria to make this change.

Let's go to the next slide. So these are just a few questions that committees can

consider when they're reviewing the SDS adjusted measures. I won't go through all of these, but basically again, no surprise here. Is there a conceptual relationship or not? And sometimes the answer is actually no, there is not. And that's okay if that's the answer. Is the SDS factor or factors present at the start of care? Is there variation across measured entities? What does the empirical analysis show in terms of whether the SDS factor has a significant and unique effect? And also this whole question of data availability, which we'll revisit over and over again, are the data available and generally accessible?

Next slide, please. So, so far, we've had four projects to date that have contributed to the trial, and you probably know that we do our work in kind of project phases, so we have launched a lot of projects, but so far many of these projects that we've launched late last year have not yet got to the evaluation stage. Actually, most of them have not even gotten to the submission deadline stage. So these are what we're looking at right now in terms

of what we're seeing.

Cost and resource use and admissions/readmissions. Those two projects actually were ongoing at the time the SDS panel had their deliberations, so at that time, our previous policy was still in place, so the committees made their recommendations under our previous policy.

So when the Board endorsed the measures under these two projects in late 2014, they did so with conditions, and the conditions were that for cost and resource use, the three measures in that project would actually enter the trial. And for the admissions/readmissions project, the conditions put on for endorsement by the Board were that the standing committee would look at the measures and decide which of those measures would enter the trial. So cost resource use, admissions/readmissions are a little bit different.

Then we had cardiovascular phase 3. That one was the first one out of the gate, so it was really the guinea pig for us seeing how things

would work, and then pediatrics is one that is still fairly new. We've had the in-person meeting, but the measures are not all the way through the process yet. So let me give you just the highlights of these.

Next slide, please. So cost and resource use. There were three measures, and they were all hospital-level risk standardized payment associated with the 30-day episode of care for these three different conditions, AMI, heart failure, and pneumonia. And I won't say too much about these, but Susannah knows these intimately, so you can certainly interject if we need to.

So let's go to the next slide. So basically what we did for cost and resource use and for admissions/readmissions, we changed our process just a little bit, so rather than having the entire discussion in one big meeting, we worked with the developers to come up with a time line, and basically we split the evaluation into two pieces: one, discussions about the conceptual rationale, and then a second discussion later on after giving

the developers time to work and think to talk about their empirical analyses and results.

So initially, in the discussion of the conceptual rationale for the cost and resource use measures, the developers initially considered educational attainment or income and possibly using census data based on matching the patient ZIP Code; Medicaid status as a proxy for low income and insurance coverage; and, black or white race.

The committee actually asked the developer at that point early on to broaden their conceptual model and add to the literature review which they did. And they did come back, and when they presented their empirical analyses, they actually looked at race, operationalized as black/non-black and Medicaid enrollments, for dual status as proxy for low income.

Next slide, please. So based on the empirical analyses that Susannah and company provided, and I will say you guys did a really nice job. I think you hit everything that we asked you to do. They did choose not to include those SDS

variables in their models, and their reasoning was nominal impact on the risk model performance itself as well as very little impact on the outcomes that they were looking at, the cost outcomes. And ultimately, the committee voted to continue to endorse them.

Let's move to the next slide. Admissions/readmissions, that project is not quite as far along. Very early on, their committee did look and listen to developers about the 17 measures that were in that project. And they determined that of those 17, 16 should enter the trial period. So they had in September an additional meeting with the committee. And in that meeting, they talked about some of their approaches for their critical analyses. And the committee gave some input on those analyses, and I'll describe that in a little bit. But to be clear, they haven't actually presented those analyses yet to the committee. So those meetings are coming up in March and in May, so we're still yet to know what's going on there.

Next slide, please. So in terms of the

input from the committee on the admissions/readmissions measure, there was some tension in thinking about quote unquote the robustness of the proposed factors versus the data availability and accessibility. So again, I think this is the same conversation that you've already had, but the committee did speak about that.

They pointed out that there is not just one way to risk adjust and that they felt that NQF should not be prescriptive regarding either methods or factors. But they did provide guidance to the developers that if characteristics are present prior to treatment, and they are known or suspected confounders, then they should at least be considered for inclusion in risk adjustment approaches.

Next slide, please. They specifically encouraged consideration of age, gender, and some sort of measure of poverty, for example dual eligibility status. They suggested using community-level variables when patient-level data were not available, and they also asked the

developers to justify any decision not to include such factors. And there was discussion about proxy data, particularly data from nine digit ZIP Code matches versus five digit ZIP Code or county matches with the feeling of that committee at least at that time that the nine digit match may be best. And finally, they urged caution on use of race as a proxy for patient -- I should say SES in this case. And as we've already discussed and heard in this meeting, the construct of race is more than just about SES.

Next slide, please. Actually, can you go back to that slide, just to put in a little thing for you to be thinking about, especially as Jose goes through his slides a little bit later on, the IOM report did differentiate between the committee-level factors thinking about using them as proxy for individual data versus kind of a group level, contextual kind of factor, so that might be something that you as a committee can really help us understand. I know that the panel did talk about that in their report, but that in the field I think

might be still confusing to people out there, so that might be something you can consider helping us with.

Next slide, please. For cardiovascular phase 3, that one, 27 measures entered the trial and of these, 10 included risk adjustment. So I will say that it was our decision to -- even though the panel said we're mostly focusing on outcome measures, but process measures are on the table for adjustment, we are mostly focusing, at least in this trial period, on outcome measures. So basically our rule of thumb is if the measure is coming in with risk adjustment, then the SDS discussion is definitely on the table and something that we will point out. If a measure is not risk adjusted, we won't necessarily make it an issue, if that makes any sense.

Does that make any sense? Okay, let me try again.

If a measure comes in, and it's an outcome measure, it's usually risk adjusted. So we will do -- we will go through the whole SDS

discussion with the committee. If it's a process measure that is not typically risk adjusted, we're not going to make the committee have the discussion about should it be risk adjusted and should SDS factors be included in that risk adjustment. They can have that discussion if they want to, but we're not going to put it explicitly on the table.

DR. BURSTIN: Okay, to slightly finer print it, it's really complex, so think about how you add SES Adjustment when you haven't -- when you have no adjustment of any kind. So it feels like a bigger leap.

And so we're -- the committee is going to talk about it, but it's harder to -- kind of, hard wire into their discussions of, do you add these factors when there isn't a model to add them to?

MS. JOHNSON: And often what we'll see is, in the questions about risk adjustments, developers of process measures will say, not applicable, it's a process measure.

So, you know, in some cases, this has never even been a question. So it may need to be

a question further on, but for this trial, at least, for now, we're not, necessarily, going down that road.

So four of the measures did include information on the conceptual rationale, and they looked at variables, including race, dual eligibility status, and an art composite index of SDS, and that art composite was applied using the five-digit zip.

But, ultimately, none of these were included in the risk adjustment protest, for those four variables. Six of the ten, actually, did not include information on the conceptual rationale in the written submission, but the topic was addressed, at least briefly, during the discussion of the measures.

And where we are with this project is those six are actually coming back for additional consideration by the Panel. You know, this is a huge change in our philosophy and policy, and it can be a lot of work. But, at the end of the day, we have 30 minutes, 45 minutes, depending on the

measure, time to discuss things.

So some measures have other things that take up the Committee's time, and the risk adjustment approach is not always the biggest topic for discussion in evaluations. Next slide, please.

So Pediatrics. 24 measures included in the trial, and 11 included a risk adjustment approach. These 11 measures, actually, were based on two different instruments, so we consider them what we call patient-reported outcome performance measures.

But, in this project, actually, there was relatively little discussion of the risk adjustment approach. There were, again, other issues that occupied the Committee and risk adjustment was kind of low on the list of things that they wanted to discuss.

But one of the measures did include a conceptual rationale. And, you see the variables there. They actually did look at a few things, but only included age and self-reported health status

in their final model.

The remainder that came, actually, from one of the instruments, considered several variables, including child gender, age, and race, ethnicity, and then caregiver age, race, ethnicity, English proficiency, and educational attainment. But after analyses, they included only the respondent, or caregiver, education in their final marks.

Okay. And, I think that's the end of that slide. So that's where we are right now. I don't know if we want to have any discussion on this?

CO-CHAIR PONCE: Michelle.

MEMBER CABRERA: Thank you, for the overview and for all the work. It sounds like it's been a challenge to, sort of, chart these -- or try to navigate these unchartered waters.

I just wanted to make sure that I'm understanding correctly, and I have one suggestion for this Committee. Is it, is it true then that of the initial, all except for the Pediatrics, there was consideration or conversation, but SDS Risk

Adjustment either was decided, was decided against, or it was limited to like one factor?

MS. JOHNSON: You are correct. I'll tell you another thing that we kind of operationalized, and if you guys think that this is an incorrect way to think about it, it would be good to know.

Race and sex, or gender, are SDS factors, but they can also be considered biological factors or whatever, so some models come through, and they do have those two in there, and we're not really focusing on those, necessarily, for this trial.

So what I mean by that is, it is correct to say that SDS factors were not included in the previous models, except for the Pediatrics, except for possibly age and gender.

MEMBER CABRERA: Okay. I think one thing that would be helpful, as we're, sort of, doing our, I don't know, arm's length review of what's going on, is to have a chart that lays out each area that's been considered, whether it

included a conceptual -- you know, just something so that we could in kind of a snapshot get a sense of the -- what's gone through the pipeline.

The other question I had was, with -- because I don't recall hearing this: with one -- areas where we -- there are known disparities based on factors such as race or ethnicity, is there -- is there a crosswalk happening there, or a conversation, or is it being suggested or raised that those things be looked at in relation to these measures? Or would that be inappropriate? I don't know.

MS. JOHNSON: Often that is included in some of the information that's given by the developers, but it may not be, maybe, as robustly discussed, as we might like, in terms of whether SDS adjustment is appropriate or not. I'm not sure if that's quite answering your question, but that seems to be what we're getting.

CO-CHAIR PONCE: Okay, so we have Phillip, Kevin, and I don't -- Mara, Bob, and Traci. Phillip.

MEMBER ALBERTI: Thank you, so much, for that overview. I'm really excited to see the progress that's been made. I have a -- I have a concern kind of wrapped in a question.

And so to get into it, you know, there's a paper that came out about a year plus ago that showed that if you discharge a patient to a disadvantaged community, that risk confers the same risk of readmission as if that patient had COPD. They were looking at community-level variables.

And so I was very heartened to see in Recommendation 10 that some of the SDS factors be considered either at the patient or community-level.

And then, over the course of presentation, I got a little disheartened by community-level factors either being looked at very broadly at a ZIP Code level, which, for a place like New York or other places just doesn't make any sense, or only considered as a proxy when individual patient level data aren't available.

So from an epidemiologic perspective,

they are independent of facts of my income and my community's income at a census tract-level, and so I wonder, has that been a part of the conversation?

Why has, or -- I guess, you can only assess what kinds of variables are submitted and the logic behind the measurement, but I -- has there been thought to really look at some community-level variables that capture issues of neighborhood poverty, neighborhood resource that are often sometimes even more important than my own personal income or my own race, or as important, that aren't really collinear?

MS. JOHNSON: I would think that, right now, that has not been the case. I think it might be fair to ask, Susannah, your thinking, because you have helped in some of the, several of the measures that have come through, so are you thinking about it from the contextual point of view?

MEMBER BERNHEIM: Yes, I think this is a place that it would be great to have, sort of, a convergence of conversations. We couldn't really get into it in the last committee.

The primary expectation about this has -- so far, has been, you know, what can you do for patient level risk adjustment, so just being really concrete about terms.

You put it into the model, as if it went with the patient; even if the data's coming from their neighborhood district, just you're pretending that they're neighborhood tells you something about them, right? That's the way the models are set up right now; we have fixed effects of the patient's level.

This community issue, I think, is important, and there's -- you know, it's hard for me to -- I'll try not to get too lost in the technical, but when you bring it to the Committee, you have to get into the technical.

And so the question is, would we, literally -- we have two level models already, where we sort of add a third level. I mean, it's really a hard thing, so I don't want to get into it now, but I think it's important, and it's not well-fleshed out. We're thinking about it a little

bit.

You know, I don't know if you have a sub. I don't know how NQF can help, but right now it's not what we're bringing forward, and it's not what we think we're being asked to bring forward, but it comes up in all of the conversations --

MEMBER ALBERTI: Okay.

MEMBER BERNHEIM: -- very similar things to what you're saying. It's not about the patient; it's about the community. How do we count that? And there are only two ways this can move. I mean, we could bring something forward and have committees do that. You guys could help us frame guidance about it. But I think it's an, it's a important area that's on --

MEMBER ALBERTI: You know, on our SharePoint site, there was a paper from some work out of Missouri that was looking at kind of a hierarchical model that had fixed and random effects and that showed some real, you know, rubber hitting the road when adjusting for census tract-level data.

And, you know, the, the ONC and their panel that they've recommended for inclusion in their certification actually includes GO coded address that -- so if it's an issue of data availability, that might be changing. So I wonder if there are other models that we could look at as well?

CO-CHAIR PONCE: Yes. Okay, Romana, on this point?

MEMBER HASNIAN-WYNIA: Was the work in Missouri focusing on health literacy, kind of adjustment? You know, focusing on mapping for --

MEMBER ALBERTI: I've seen, so I'm not sure what the paper in the SharePoint was, but what a presentation that I saw actually looked at is literacy, but also, there was some readmission stuff as well.

CO-CHAIR PONCE: Okay, so we'll note, looking at other models, multi-level models and get at explicitly modeling the contextual effect versus approximate or the individual. Kevin, Mara -- just run through the roll. Kevin, Mara, Bob, Traci, and

now, Yolanda.

MEMBER FISCELLA: Just a quick clarifying question. When you looked at --

DR. CLARK: I'm sorry. It's Lisa. I just want to interject that I've had my hand raised for a while, and nobody's noticed.

CO-CHAIR PONCE: I'm sorry, Lisa. I didn't see it. Sorry. I'm going to put you after Mara, so -- oh, Kevin said he'd yield to you, Lisa, so go ahead. Lisa?

MEMBER COOPER: Okay. Sorry, I was on -- I muted myself, temporarily. So I -- I have a -- this has been a really interesting conversation, so one of my questions goes like a little bit further back, and then one of them is, is more up to where we are now.

So I guess, and we may not want -- I don't know if we want to backtrack to this now or later, but I didn't have a clear understanding about the process.

Like, I know that all the measures that are coming through are going to be part of this trial

period, but I guess, I don't -- I don't understand like whose decision it is, like which measure gets proposed for risk adjustment. Is that all up to the measure developers?

And, if they propose something for risk adjustment, do -- is it all on them to like provide the rationale, and if so, is that sort of a disincentive for people proposing measures for risk adjustment, because they know it's going to be like extra work for them?

Like I couldn't quite -- I didn't quite understand exactly who decided which measures get proposed for this process in the -- in the first place, so that was my one question which was about process that went back a bit further.

And then I guess my other comment is not really so much a question, but I guess one of the ideas that I think we might want to talk about later is to what extent these measures are being used only to -- for performance or for like determination of payment, or whether they can actually be used to guide resource allocation, or guide, like,

programming planning, and then, you know, that might determine when we think risk adjustment is appropriate, as well.

MS. JOHNSON: I'll answer the first part. This is Karen. Which measures are risk adjusted, well, that is actually up to the developer.

We have, you know, quite frankly, an expectation that outcome measures, or resource or cost measures, we generally expect to see those to be risk adjusted, so we actually have a question in our submission form that says, if you're not doing risk adjustment, you need to justify why you're not doing that.

So it is up to the developer to make their best decision. You know, they know their measure, they know their data, but then they have to make their case to the Committee, and the Committee may or may not agree with that decision. So that's how it works.

DR. BURSTIN: I think, I think Lisa's asking more so the question of is it up to them to

decide whether or not they submit variables with SDS adjustment, and so in that instance, what we've done is, as part of the submission form now, there is a section where they have to include a conceptual basis.

So they have to at least put forward why they did or did not choose to include SDS, and the Committee's purview is to actually challenge that. So we --

MEMBER COOPER: Okay. Yes, I didn't understand --

DR. BURSTIN: Yes.

MEMBER COOPER: -- that they had to --

DR. BURSTIN: We didn't explain it --

MEMBER COOPER: -- submit a justification, either way, one or -- one way or the other, for each measure.

MEMBER BERNHEIM: For people who haven't gone through the process, looking at the NQF forms, they are impressive. And I think it really helps people. I mean, in a good way.

I mean, they are, there's a lot of detail

about what's expected. There's a lot of information in them, and I recommend, for anyone who hasn't done it, to just one.

And you can take one that's filled out; they're all public, you know. You can look at ours, if you want to and give us some advice, or you can just look at the empty forms.

But it's worth seeing what we're asked to bring. It's everything from, sort of, why is this important to measure, to details of how you assess the reliability and the validity of the measure and how you selected the risk variables. It's very comprehensive. Who's going to answer how is it -- how is it used?

CO-CHAIR PONCE: And I was going to put that one to Helen.

DR. BURSTIN: The second part is really interesting, because it is this question of the intended use of measures, and some of you may know, we just completed a very large body of work trying to decide if endorsement should actually be tied to whether it's a measure of payment, whether it's a

measure of public reporting, and ultimately decided there's just, frankly, not enough science to really differentiate what that criteria be for one or the other.

But we are recommending a new designation within the NQF portfolio of measures that exceed our current criteria and may be measures that, for example, the Measures Application Partnership might consider differently because they exceed criteria in some key areas, like testing at the measure score level, always a big issue for measures that are used for comparative performance and also input from end users effected by the measure, as a part of the solution. So more on that to follow.

But, you know, much of that work around how the measure's actually used is the work of the Measures Application Partnership, which meets in this room next week, so it will be an interesting time.

CO-CHAIR PONCE: Thank you. Kevin and Mara, Bob, Traci, Yolanda, and now, Lisa Iezzoni.

Kevin.

MEMBER FISCELLA: A quick question. When you look at at whether SDS makes a difference or not, are you just looking at the amount of initial variance explained, or are you also looking at changes in potential ranges?

MS. JOHNSON: We've asked developers to provide both of those to us.

MEMBER YOUDELMAN: I'm struggling a little bit in trying to figure out how to formulate my question or my concern. So it seems like, to date, 50 or so measures have been considered, but only two actually have, sort of, formally been approved for the pilot period.

And so I guess, I'm a little bit -- and those don't include some of the factors that I think the Committee was looking at to a greater degree.

So I guess, I'm trying to understand, do you expect that to change, as things progress, and measures coming up soon are going to be better suited in some way, I don't know what that is, because I don't know how this really works because

I'm not enmeshed in this world.

Or, to me, I guess, I would think, part of the trial period is to, sort of, make some measures have to do this. Because, if the literature doesn't yet support it, we might not know, until we look at the quality measurement, that there is a disparity.

So to some degree, some of the earlier slides, which sort of said, well, there isn't literature on this, and therefore, we're not going to do it, seems to be a little chicken and egg of well, isn't part of a trial period to see? If we do it for some of these measures and -- I just -- I'm sort of struggling with this because I don't quite see, therefore, all the work that was done, and sort of wanting to do it, but then not having a really robust number of measures yet, and worried that we're going to get to the end of a trial period and not really have it.

And I don't know how to fix that, but it's just -- it's a little bit worrisome to me, because I don't understand how that plays out and

what the next steps are, and therefore, when you get to the end of a trial period, if you only have a handful of measures that actually have gone through that, we don't really have any results. So I don't quite know what the real questions are, but --

DR. BURSTIN: And we'll talk about this more when we hit the evaluation plan for the trial period, but I think your point's well-taken.

And the conceptual basis isn't just, necessarily, grounded in the literature. You could make, I think, a strong rationale or a case for why a particular measure, given its conceptual area, would logically be adjusted. So again, it isn't fully dependent on the literature. That's, obviously, the strongest way to put it.

And, secondly, I think, as you'll see, one of the biggest issues we're encountering, not surprisingly, is I think what we're going to find as we go through the evaluation plan is there will be many more measures with a conceptual basis for which the empirical data isn't holding up.

And one of the questions is going to be,

why -- how often is it because frankly there isn't a real relationship, and how much of it is really that there may be a relationship, but the variables we've got, currently, to assess SDS/SES, are not adequate for the task. So more on that to follow. Did you want to add something, Marshall?

CO-CHAIR CHIN: Yes, I think I'm -- not to read too much into it, but your tone of your voice when you're presenting, you almost implied that, like some of the committees didn't take it seriously.

So you said that like, from example, the pediatric group, out of 24, it's hard to believe that there weren't some of those 24 for the kids that weren't appropriate, but you said that what -- they had other issues that they wanted to talk about instead, and so it was like low in priority.

Or, you said, like for example, the cardiac one that the answer was, well, these are process measures, end of explanation. Therefore, by the process means, well, you know, you can't possibly hit relationship.

Can you expand on it, like, were people taking it seriously? Was there sort of a good reason why you decided these were not appropriate for the trial period, or they really didn't take it seriously?

MS. JOHNSON: So apologies, if I made it sound like the Committee didn't take it seriously. I think they do take it seriously. However, the way we structure our evaluation process is we have criteria, and we go in a particular order.

So for example, and I don't recall if this happened with these measures or not, but the first thing we talk about is evidence. And that's not a huge thing for outcome measures, but if a measure doesn't make it through evidence, we don't even go forward to talk about the scientific acceptability.

So that could be it, that we just don't even get to it, or there may be other things. For example, some of the measures from the cardiovascular, there were real concerns with basically what was included and being included in

the measure.

And that, I think, was a -- it was also thought about under the validity criterion, but it over -- that discussion was a little bit more basic than the risk adjustment.

So if you don't even agree about what was being put in the measure and what's being measured, then the risk adjustment becomes secondary. So that's a couple of examples of what's going on.

DR. BURSTIN: And just to add to that, I think Karen's right. Most of the time -- most of those measures that didn't get this discussion, as for Leslie, partly, because they had already -- kind of didn't make it through based on the earlier criteria, so we'll see that.

And in the pediatric example, a lot of those were survey-based measures where they had included education; as you saw, there was consideration about the factors, and, obviously, education is an SDS factor that did make it through. But the other factors, they found were unnecessary.

But, we do also have that table. I

think it was Michelle who asked for that table, so we'll share that with you just to make it a little easier to understand.

CO-CHAIR PONCE: Bob.

MEMBER RAUNER: I was going to ask if you could give us a specific example of what made it through the first cut, how far it got through?

Like, say for example, pediatric asthma admissions where you would see potential racial ethnic access area issues that went through. These variables were added, but then they all failed out, after adjustment showed no significant difference. Can you give us an example like that, or are some of those protected?

MS. JOHNSON: I don't think I can give you an example off the top of my hand because I just don't remember the details that well. The pediatric -- the two pediatric ones were not asthma EDUs.

As a matter of fact, I think those are coming up in an upcoming project, a pulmonary project, so stay tuned for that to see what happens

there. Do you have any examples, Helen? Nothing's --

DR. BURSTIN: There's one of them from CV3. It was -- they -- sorry, the third phase of the cardiovascular project, sorry. One of the measures was an inpatient mortality measure, so they actually did go through the process of looking at SDS variables, but not surprisingly, I think, they found that it wasn't really affecting anything, so --

CO-CHAIR PONCE: Traci.

MEMBER FERGUSON: Yes, going through the discussion, it said that NQF would not be prescriptive in what they would do. Is that sort of the role that you would see this standing committee being, somewhat prescriptive in what we would expect, like certain areas in terms of measures? So that even though you couldn't, could the Committee be prescriptive?

Because I feel like there are some measures where they didn't have a conceptual rationale. Is it because it was just so daunting

to them to put it together, even thinking about it, or they -- because like, it just seems that -- maybe it's more training; I don't know what it is. It just seems like it's -- for cardiovascular to not be able to come forward with something, it just seems almost like a -- almost like a waste. I mean --

DR. BURSTIN: Yes, you know, one thing that might be useful is, we asked all of the developers, particularly those that were required to be in the trial period, to submit the conceptual basis for each of those measures. And maybe we should share one or two of those examples so you can get a flavor of it.

I think, in general, they did a really good job explaining a conceptual basis. And, interestingly, we went through these 17 readmission measures, 16 out of the 17 develop -- 16 out of the 17 measures, the developers agreed there was a conceptual basis.

So I think that conceptual issues are actually pretty minor compared to the empiric

issues. So either the measures were just not appropriate; they were process measures, so they didn't make it through the criteria in the first place, or they may have had a conceptual basis, but I think what we really need your help on is how do we handle the lack of what we're finding, in terms of the empirical analysis.

And this may come back to the earlier comments that Tom and others and Phillip had around, you know, maybe you could really help with greater clarity, for example, around these neighborhood characteristics.

So it isn't, I think, so much that we don't feel like we have the power to do that. I think there isn't a whole lot of clarity that we feel we can put forward.

And so if you can offer additional clarity and what we would then ask developers, it would probably, I think, serve everyone well.

MEMBER OGBOLU: Yes, I agree with Lisa's comment yesterday, sometimes by the time we get around, most of my questions have been answered.

I really wanted to piggyback, initially, on what Phillip was saying about the community-level variables and the use of ZIP Code and how that's -- presents an additional challenge, and I think that the way we're measuring community-level variables is not the best.

And then, I had some questions about the process, which was, how do we -- I guess, when I first saw the initial slide where it said we need to boost conceptual evidence as well as empirical evidence, my mind already started thinking, let me just wait until you have the second presentation from Karen to see where we're going with this.

But we all know that there is a lot -- there is much gap. There is a lot gap between, kind of, the conceptual frameworks and conceptual theories that we have and the empirical analysis.

So I think Susannah had suggested having the option of taking a look at some of the forms. I think that might help. And then, also, I think, if we think about special conditions that might really be able to show both -- show all the systems

involved, looking at the health system, looking at the community factors, an example like asthma that was brought up, can be particularly useful.

And some people could probably speak to some other ones, but I think there are certain conditions that might be really helpful for us to look at, like asthma.

So I think, we probably will have to dig into this a lot more as a group to really better understand the gap between kind of the conceptual frameworks and the empirical data.

CO-CHAIR PONCE: Great, thank you. Lisa Iezzoni and then, Eduardo.

MEMBER IEZZONI: Yes, I --

CO-CHAIR PONCE: And Tom.

MEMBER IEZZONI: -- I've been listening intently, about -- just to see whether there was anything intelligent that I could actually say, but I think people have done a great job. You've really, really done a great job surfacing a lot of these issues that are really, really complicated, and I really congratulate people on the efforts that

you've taken.

The question that -- I have one question and then just a comment. The question I have is, I really think that we need to know more clearly, and this has been asked before, so I'm just repeating something that's already been asked, what do you expect this Committee to do?

Because, I really am wondering, given some of the responses, whether we have the right makeup of the Committee, whether the right skill sets are around the table for some of the, kind of, answers that I've heard about what you expect this Committee to do.

I just think that this could be a huge amount of effort, and it would be really important to make sure that everybody around the table feels equally engaged in the effort, rather than feeling like they're kind of wasting time because this isn't their area of expertise, and so it's not a good use of their time.

But my comment was going to be some of this reminds me of the early risk adjustment days

and looking at hospital mortality where -- oh my gosh, and the work that we did, we tried to risk adjust mortality every which way that we possibly could.

And at the end of the day, the raw rankings looked very kind of just similar to the unadjusted rank to the adjusted rankings.

There might be one or two providers that would pop out, and you'd go, oh gosh, there's a difference; their ranking changes a lot.

But, I think that you should not -- you should not be surprised that you might not find much change in what's going on. I really just don't think you should find that you should be surprised, at all, about that.

I think it probably has to do with underlying statistical assumptions and the quality of the data and a lot of things that just make this a very messy enterprise.

But at the end of the day, you know, we're finding this at the MGH, that we're trying to adjust physician payments for our primary care

doctors, to some extent, not the entire payment, but part of the payment, based on some of the risk factors relating to their patients, like whether English is spoken by the patient or whether they need to have interpreters, et cetera, et cetera.

And, again, at the end of the day, it might not change the numbers that much, but it has a lot to do with the credibility of the data to the end user.

And so even though you might not find that much statistically, is this going to allow you to have more informed policy discussions with the people whose behaviors you want to influence?

So I think that that's what I, kind of, frankly, expect to see is that you're going to spend a lot of effort on this. You might see that much, in terms of impact, statistically, but it might allow you to move forward some discussions that you might otherwise not be able to move forward.

DR. BURSTIN: Let's come back to what the Committee will do when we hit the Evaluation Plan. That is absolutely right, and I had lots of

thoughts. I want to grab these two over lunch, perhaps, just quickly, and see if we can strategies a bit.

CO-CHAIR PONCE: I said Eduardo, but, actually, I think Tom was first, so Tom then Eduardo.

MEMBER SEQUIST: My comment may follow some of what Lisa is saying. Maybe a little bit of a different angle on my question, though, is, is when you find something -- or someone presents a measure, and we think just conceptually, how could it not matter what your social factors are, or you know, the different variables we're talking about.

And then you do the quantitative analysis, and it doesn't show anything, and it sounds like that's happening a lot, does there need to be -- given that it's, we feel like it's happening a lot, does there need to be some sort of structured process to then, sort of, push the question of saying it's not that the conceptual model's wrong; it's that the analytic model is wrong?

And the analytic model's not wrong; it's

correct, for what you have. It's just that you're not doing it right. And so if it's going to happen -- I was just struck by the numbers that you're describing in. And if it's going to happen a lot, should we -- is there a role here to sort of figure out, okay, well what do we do next, rather than say, okay, we have a measure that's mortality, and it doesn't matter if you're poor, or rich, black, white, and we all conceptually say how could -- that doesn't make any sense at all. Do we need to have, like another process built in, I guess, was my question.

DR. BURSTIN: And that's another issue we'll come back to. We've actually got that on the Evaluation Plan. But we can, for example, and we've already talked about that with some of those measures.

For example, saying, you know, in one year, there's an expectation you give us an update as to where you are with getting better SES variables.

Because that's, oftentimes, I think,

what we're seeing that it's not so much -- the analytic work, to Lisa's point, isn't so much an issue; it's just that they're not -- they don't have access or currently can't get some of the better variables we think might actually show in the empiric analysis, although, we don't even know that either, or again, we're assuming that.

CO-CHAIR PONCE: All right, Eduardo and then, Susannah, Yolanda, Ron, and Traci.

MEMBER SANCHEZ: So I'm sorry. I stepped out, Karen. I really appreciated the presentation, and it sounds like NQF is trying to figure out how to move forward as an organization.

And I think I've heard some comments that get to this. One is kind of the receptivity in these four areas, were they prepared in advance for what was coming?

Is there, has there, and should there be more socialization of this idea, so that folks maybe are in a different mode when this is brought to them?

Because it's just not -- it wasn't clear to me whether this is just, no we don't want to do

this, or no, we've never, we've never thought about this; we need a little bit more time.

And, I think, some of the conversation I'm hearing kind of gets at that. And then, it speaks to is one of the charges of our standing committee also to begin to figure out how to socialize this, if indeed that's one of the issues inside of NQF, that folks aren't necessarily thinking about this as much as we would like to believe they do or should?

DR. BURSTIN: Those are all good -- those are all great questions, and it sounds like you were here for the whole time. I think more socialization would help.

I do think -- just to be blunt in this room, I think there is still a debate. And I think there are still some developers who, begrudgingly, are submitting conceptual basis because they have to but don't agree with it.

So I think the more clarity this Committee can help give to those discussions, and I think over time, also, this all happened and

developers had about, what, six months or so to prepare?

MS. JOHNSON: Six months.

DR. BURSTIN: But, if you don't, if you can't get access to the data -- for those of us who tried to get access to the data, six months may seem like a long time, but it's not a long time.

MEMBER SANCHEZ: Right.

DR. BURSTIN: So I think some of this, I think, is evolving and, hopefully, over time, we'll see, with your input, people are better able to address the challenges.

CO-CHAIR PONCE: Susannah.

MEMBER BERNHEIM: So two quick things, and I'm trying to be very aware that in this piece of the discussion, I really wear two hats, and I will admit, I'm wearing both of them.

I think that it is important, there's a vibe in this room of, sort of, why aren't more measures coming forward, risk adjusted, and therefore, the process isn't working.

I think that we -- our job is to learn

from this process, and so we need a lot of information, and I think what we want to see developers do, and what I am trying to get my team to do is to take this very seriously, right?

So what you want to see, what I want us to ask is, did people put forward a conceptual model that makes sense? Does this Committee want to review some of those conceptual models? Did the empiric data support the decisions they made? Did the Committee take this seriously?

But, I think, one of the misconceptions that comes up a lot around quality measures that I think is just worth naming is they are not predictive models, right?

What, the goal for our mortality measures is not to do the best job we can predicting people's outcomes. That's a different process, and SDS and race can be very strong predictors.

When you put them into models that are designed to separate out and think about quality when you've accounted for many, many other comorbidities, and you're only trying to account

for those things that are present at the start of care, sometimes they don't make a big difference.

And so it's not that it doesn't matter, or people don't care. I think when we see things we don't expect, I would ask this Committee to come to that with curiosity and interest as to what that means, as opposed to a sense that, people aren't taking this seriously or not, right? So I just, I just want to put that out there.

CO-CHAIR PONCE: Okay, Yolanda, your card was up, and now it's down. Okay. It's still up? Oh yes, from before, sorry. So --

MEMBER OGBOLU: Or --

CO-CHAIR PONCE: -- it's going to --- go on.

MEMBER OGBOLU: Yes, but I think we're still having that conversation. One other thing I wanted to say earlier, a couple of people said that some of the variance was trivial around some of those factors, and again, looking at the forms might help. But those are the kinds of comments, you know, how do we quantify what's trivial?

CO-CHAIR PONCE: Oh, okay, noted. So then, Ron, Traci, David, Nancy, Kevin.

MEMBER COPELAND: I just want to make an observation, really building on what Lisa said earlier, around who -- was the intended use of this data risk adjusted or not, and whose behavior are we trying to impact with this recommendation that they move disparities work forward?

And I was waiting to see in the report whether you were going to give us any insight around the themes or the insights that came from the comment period that we commented on earlier. You said you had 800 something feedback.

And what wasn't clear to me is whether -- what was the composition of the feedback, and particularly, from the medical community, physician practicing community, medical associations, were they part of the group you solicited feedback from, and if so, did these things come back or come in at all?

Because, I think the concern about this risk adjustment piece, particularly for physicians

who just proportionately care for folks in the quote unquote safety net?

Because when you talk about risk adjustment, their two cares are how is it going to impact reimbursement, and how is it going to impact resource allocation to provide the care, and if those critical decisions are going to be based on some risk adjustment methodology, then, obviously, they have a care in it, care about how it's done and getting the data accurate and so forth.

If that's not the case, then there may be a different level of importance attached to it, other than just the analytics of it. So I'm wondering if the physician community gave you feedback.

And, the physician community, and it's very diverse, so some sitting in very large integrated systems, some sitting in academia, and some in the front lines, dealing with day-to-day issues, were their voices part of the feedback on what you were looking at?

DR. BURSTIN: Yes, just very briefly,

we got a ton of feedback from the physician and the provider community, I would say overwhelmingly supporting adjustment.

So probably the largest group of commenters were, probably, providers broadly: clinicians and health systems, et cetera. By far the largest group, yes, and almost universally in support of adjustment.

MEMBER COPELAND: Probably for the reasons I've mentioned.

DR. BURSTIN: Exactly.

MEMBER COPELAND: Yes.

CO-CHAIR PONCE: Traci, then David.

MEMBER FERGUSON: Yes, I would just suggest that even those that decide not to risk adjust and ask the developers during the sort of evaluation phase, that they continue to collect data or certain variables.

Because it may not be, and I don't know how they could do this, but it may not be, like, at that time when they're doing the evaluation that there was a difference, but maybe there's another

factor or variable that they didn't expect, but they collected that, later on, they can see that there's a difference.

I don't want us to -- I mean, it seems like it boils down to the fact that we don't have enough data or the factors in order to collect it, but we have to do something.

And, I think if we could just say, please just collect it, and maybe we can put different -- you know, different sample sizes together and then collectively get information that -- because, right now, it's -- I think, it's going to boil down to not having enough data and then, two years will be passed, and we still don't have enough data.

CO-CHAIR PONCE: Okay. David, then Nancy, and then, Kevin.

MEMBER NERENZ: Sure, if I could just be amateur statistician for a minute, and I'll be corrected here by those who are more real than I am.

This is on the issue, and I think I'm going to tell you that Tom, Lisa, and Susannah, both, about why would you not see an effective

something in adjustment the way that we try to do it?

Well, part of my thinking about this is, if -- let's take an example of a social factor that we think must matter, and let's just take, for example, income or poverty.

And let's say, for example that, we just, we know from some study that poor people are ten percent more likely than rich people to have some bad outcome, readmission, mortality, pick whatever one you want, ten percent difference.

Now, the direct translation of that into quality measurement would be as if -- and now we've got certain entities that we're measuring: doctor, hospital, whatever.

The direct translation would be, as if the measured entities had pure rich or pure poor, and if that were the case, you might expect a ten percent difference. But that's not the real world. That's never how it works; the entities always have a blend of proportions of this risk factor.

And so two doctors, two hospitals might

have, maybe, say just a ten percent difference in the proportions. Now, my arithmetic says, well your ten percent gap just went down to one percent.

And now, if your measurement of poverty is imprecise, you've added some measurement noise around that, and it comes to the point, it's kind of a miracle you can find an effect of anything when you recognize that you're never dealing with, sort of, these pure effects.

So if, if that -- now, then I just want to say that the same kind of logic, I think, applies to the process outcome relationships, that a certain process step, if that's known in a clinical trial, say to reduce a bad outcome by ten percent, you're only going to see that ten percent if you compare entities who always do it against those who never do it, but that's not how it works, either.

So it's -- and I'm glad to see Susannah nodding. I always get worried if I say something, and she frowns, but the -- Lisa as well. But it, but if this general idea holds, it means that, that this whole enterprise has not failed or somehow gone

off on the wrong track if we don't see robust effects of these adjustment factors.

And I -- so I guess, in some ways, I'm just elaborating and trying to understand why, you know, in the -- in the underlying churning of the numbers, why the -- that effect you described might occur.

CO-CHAIR PONCE: Okay. Nancy, then Kevin.

MEMBER GARRETT: I just wanted to pick up on Traci's point about the data collection. I, I mean, when I hear Helen say there's some, perhaps, push back from developers, I mean, I suspect some of that comes from resources.

They're being asked to do additional steps and, maybe they're getting the same funding to do the work, and so how do we start to align resources to really focus attention on the work?

And the developers are also not the ones who are responsible for what kind of data they have. But that's where, I mean, on our SDS Committee we struggle a lot, because at first we thought, well,

if we don't have the data, how can we do this?

But then we sort of said well, isn't this where quality measurement was ten years ago? We didn't have any outcomes data; all we had was process measures, and the way we got there was because we had enough stakeholders saying this is really important, and then resources aligned around it.

And so it's, kind of, that transition period where we really feel like we need this committee, if everyone agrees with this direction, to support it and to start to really line up those resources from a -- you know, if payers start to understand, well in order to get measures endorsed, we're going to have to measure sociodemographic factors, so that we have accurate risk adjustment, then that starts to create the platform to put resources behind data collection, and if we can get alignment around standardized collection of that data very quickly and that concept of screening for the SDS risk factors so that we can provide better clinical care and better understand risk, those

things can start to align, so it's -- I think that this Committee can play a really important role in setting that direction.

CO-CHAIR PONCE: Thank you. Sarah, let's -- oh. And so Kevin is done, right? Your card's still up, so -- no, Sarah's going now.

MEMBER FISCELLA: Oh, okay.

MEMBER SCHOLLE: I think Helen mentioned those begrudging conceptual discussions. I -- and I just want to pick up, it's exactly the -- this is, this is a huge new activity for the measure development enterprise to include this, and the data to guide it are really crummy.

And then, our ability to actually do what the -- what previous research might suggest we ought to do with the data that we have that are crummy makes this feel like it's an exercise to -- in futility, okay.

And so that -- that is, kind of, that if you -- if there is a sense that this is going through the motions, it's not really that; it's a sense of, if I know it's not going to change anything in the

end, based on all my experience and everybody else's experience, then that is a frustrating thing, and it becomes one more requirement.

And those forms are long, and we're asked to do a lot of things when we -- when we're presenting measures. So not to justify any of that idea, but that -- but it does become resources, huge amount of resources in in trying to present the measures.

So I like this idea -- and the idea that the one way you can handle this problem is to risk adjust it, without saying, well there might be other ways to handle it, depending on how the measures are used in different applications.

So in that way, it feels like there's one box we're trying to work towards instead of a whole panoply of options and a whole variety of individuals that could really help to make this easier.

So if providers are the ones that overwhelmingly said, we need risk adjustment, well then that means the -- one of the themes from this

trial period is to go back to the providers and say, you want risk adjustment? You better start collecting data in a way that it's going to support risk adjustment, and what would that look like?

Because it, you can't have -- you can't have your cake and eat it, too; you have to be part of the, the process of how it comes to understand what that is.

And I think, if we could really say -- that's part of what the road map is, is saying, we're going to have to need better -- we're going to have to get better data, and we're not going to get it by just using claims data and making estimates from a five-digit zip code or even a nine-digit zip code.

So I think that's part of the concern here is that -- and I -- and, really, as I think you did mention, NCQA is one of the groups that commented against this policy, and it's, primarily, because the risk adjustment feels like jumping to a conclusion, without looking at alternatives and without offering people that are using measures alternative ways to solve the problem, and that's

where the payment approaches come through.

You may not want to risk adjust the measure, but you could certainly change the way -- your payment policy, based on the results of the measure. So that's where -- that's part of the concern that you might be getting.

And, as we think about the road map and how to interpret the results of this initial pilot, we should be thinking about, well, how do we go back to the stakeholders and ask them, you know, what would work better; what are the alternatives for other people to help other stakeholders just support this enterprise more effectively?

CO-CHAIR PONCE: Okay, thank you. So I'm going to ask Kevin to have the last comment. Because this can continue. We want to give Jose and Christie's time to -- oh, Christie's not -- so is, well, will Jose still present?

CO-CHAIR PONCE: Okay, good. And then, I think, some of these comments would be relevant to the next presentation.

CO-CHAIR PONCE: Kevin.

MEMBER FISCELLA: So again, I would just echo the previous comments on the need for data. I think I completely agree with Lisa about the buy-in, and what Sarah just said of thinking that the buy-in to providers collecting the data, I think that makes a lot of sense.

You know, I think, I think one of the reasons why we may not be seeing this, beyond the obvious, of really poor data, is just that when you throw so many comorbidity and other measures into this model, you are capturing much of the impact of social disadvantage, both in terms of the cumulative effect on health and poor access and all of these other things that are already in there.

And so that, at the end of the day when you have this giant model with all of this, this incredibly rich set of ICD-9 and soon, even more nuance, ICD-10 codes, the amount of variance for many of these measures, particularly, when they reflect, you know, health outcomes, is probably going to be less than we anticipated.

CO-CHAIR PONCE: Thank you. Jose, are you still on?

(No response.)

CO-CHAIR PONCE: Jose, are you ready to --

MEMBER ESCARCE: Yes.

CO-CHAIR PONCE: -- make your presentation?

MEMBER ESCARCE: Yes I am. I'm on.

CO-CHAIR PONCE: Great, thank you.

MS. JOHNSON: Okay, in this section of today's program, we wanted to talk about two different things. We wanted to talk about some of the challenges, and I think you guys have hit a lot of these, so when we get to that section, we've already talked a lot about those things, but we also wanted to share some of the approaches that are happening outside NQF.

So we told you what we were seeing, so far, in-house. Go to the next slide, please. So we had a couple of things teed up for you, one, from Christie, talking about some approaches that

they're looking at, with their medication adherence measures.

And we're, since Christie couldn't be with us today, we're going to shift that, so Helen and I, together, will probably try to do a five-minute, not very good, discussion of Christie's slides. But, instead --

(Off microphone comment.)

MS. JOHNSON: We have heard it twice and it's fascinating. And we'll try to figure out some way that, that Christie can, maybe, be available to answer questions, et cetera.

But, Jose is going to tell us about what is going on with the National Academy of Medicine. I still want to call it the IOM. But that report just came out, what, last week, and Jose's going to tell us about what's going on there. So, Jose.

MEMBER ESCARCE: So I don't know if you guys are looking at my slides, but I'm still looking at the Pediatrics' measures on my screen. I can pull up my slides from my own computer, but the, the, what's shown on the, on the webinar isn't tracking

what you guys are saying, right now.

MR. PHEULPIN: We'll pull them up in, in one minute.

MEMBER ESCARCE: Okay. All right, so I'll just start by saying that I was asked to give a short summary of the work of the IOM Committee. Can you go back to my first slide? Well I, I guess -- that's fine, I know the -- so the name of the Committee is Accounting for Social Risk Factors, or SES in Medicare Payment Programs. And we can go to the next slide.

So I'm going to give a brief summary of, of the work that we're doing and this work was motivated, or triggered, by the impact of Act of 2014, which has been talked about already this morning.

And, in particular, this Act required the Secretary of Health and Human Services to report to Congress by next October 2016. On the impact of SES, so Socioeconomic Status, on quality and resources in Medicare, using measures such as poverty and rurality from existing Medicare data.

And this is language taken, abridged, but taken from the Act.

This, the Act also required to report to Congress, by three years from next October, October 2019, on the impact of SES on quality and resource Medicare using other measures from other data sources. And then, it required a qualitative analysis of, you know, where it would be, we'd be able to get these data.

And the agency that took these responsibilities on was the Assistant Secretary for Planning and Evaluation, or ASPE, and ASPE came to the Institute of Medicine to request advice and assistance on this issue. Could you go to the next slide?

So the folk, the staff at the Institute of Medicine, in thinking about the task and working with ASPE, figured out that it would be broken up into five different reports, and I'm going to talk, just give a very brief outline of each report.

The first report, which is the one that came out the other day, is to define SES the purpose

of application to quality, resources, some other measures that are using Medicare payment programs.

And then to, sort of, identify, so in an attempt to identifying it, to identify the SES factors and other social factors, and that's an important addition, shown to impact health outcomes and Medicare beneficiaries.

Report Number 2 is to identify the best practices of high performing hospitals, health plans, and other providers that show a disproportion of the higher shares of those socioeconomic effect of disadvantaged population. So this report is a little bit different from Report 1, of course, because it's, sort of, an identification of best practices. Go to the next slide.

Report Number 3 is to, I specified the criteria for determining whether a particular variable should, or should not, be accounted for in these quality and resources used, so forth. Medicare payment program that are used in Medicare payment programs, to identify those factors that

could be used.

So it's a specified criteria, identify the factors, presumably, that meet those criteria, and then, identify the methods that could be used in these adjustments.

Report Number 4 is, for each of these factors, to recommend where to get the information. So either data available, or do you require new data sources, or new strategies for data collection, and so forth.

And then, Number 5, is supposed to bring the whole thing together, the first four reports, with comprehensive timings, conclusions, and recommendations.

So I'll just pause for a second to say that, so this is very, a lot of work for the staff and for the members, as well, of the Committee.

It's, it's a very short period of time. I've been on IOM committees and, typically, they last longer than this one. We had our first meeting, I think, September, and the last meeting will be next August, so that's a one year duration,

or less than one year duration, for a whole mess of work and a whole bunch of reports. So I'll just say that, because it, it's important to understand, kind of, the way we're proceeding with the work.

So Report Number 1 came out, already, and we are now working on Reports 2 and 3, and, and then, obviously, we will shift to the last two reports.

And staff's, you know, remarkably talented all of this in doing a great amount of work, but, but actually Committee members, certainly, are doing as well. Could you go to the next slide?

Okay. So I'm just going to talk about the highlights from Report 1, because it's the one that's been released, and therefore that I'm free to talk about.

So one of the things that we did in our meeting is to develop a framework, if you will, for thinking about what are the categories of social risk factors that we wanted to incorporate, what were the outcomes that we, kind of had to look at, I mean, if these are outcomes that are used in

Medicare payment programs, so that's, that was the constraint. And, and this, kind of, represents what we came up with.

So there's a list of social risk factors that include socioeconomic decision, they include race, ethnicity, and cultural content, and I'll get back a little bit more detail on each of these, gender, what we called social relationships, and then residential and community context.

Now, the IMPACT Act and the charge from ASPE to us, specifically, asked us to assess health literacy. So health literacy is there, I mean, it might have belonged there, under any circumstance, totally under perspective, but it, certainly, had to be there.

And so with, rather than put it in, in sort of this thoughts of social risk factors, we put it into something that is, perhaps, influenced, to some degree, by them, and so you can see that at the bottom.

But these six things, on the left-hand column of this graph are, kind of, the things we're

thinking about when we are thinking about variables that we might suggest that would be appropriate for adjustment.

And then there's stuff that goes on in, in the black box that you're going to access and how they affect clinical and behavior risk factors.

And then we get to the right side of the picture, and the right side of the picture has the outcomes that we need to think about, or because they are in Medicare payment programs and they include some processes of care, some utilization measures, those sort of, kind of, labeled health care use.

Then, there are the health care outcomes and, basically, health outcomes, as well as patient safety outcomes, and then, patient's experience, and then, at the bottom, you see what we call resource use outcomes, and that's, really, made the, you know, the cost of care, Medicare cost, so you --

So this is the framework that we used to try to organize and that we, well, we came up with

and that we will use to organize the rest of the work of the Committee. Next slide.

So to give a little bit more detail on these things, socioeconomic decision, we are referring to indicators of a person's decision in society that captures access to material and social resources, as well as relevant practice when, when operation lines, we're referring here to the standard measures that belong in this box, so they would be income, they would be wealth, they would be educational payment, in theory, it would be occupation input.

Race, ethnicity, and cultural context is, really, race, ethnicity, language, and nativity. Gender, in theory, again, that's different from what happens in practice, but in theory, we're talking about people were conceptualizing the social dimensions of this beyond just, sort of, the, the, the biological classification.

For social relationships that's broken down into marital status, whether or not you live

alone, and then the degree of social support. And then, residential and community context views supposed to capture, as broadly defined characteristics of the environments where people live, their neighborhoods, so to speak, that may be relevant to health.

And, and just to touch on a previous discussion, in this framework, like in most frameworks, this is separated from the individual characteristics of race and of socioeconomic decision and so forth, conceptually.

And, of course, on our committee we had discussions about what happens if you were missing data on individual variables, but you have accessible data on what's going on at the census tract-level, or at the zip code level, and then, you'd be picking up a bunch of facts, right. In a sense you'd be picking up conceptual effect, but, of course, you might be picking up something about the person, because, actually, people vary in any given small area, but they don't vary, as much as we do, across the entire, the entire area. And

then, health literacy is the, you know, the variable in this.

So we'll go to the next slide, which I think is the framework, again. Only because I wanted to remind you about the outcomes, right, so I'm going to show a slide on, with a little bit more detail about the outcomes. Go back. I mean, next slide.

So these are the outcomes that we need to look at, and there is just a listing of what was on the, on the conceptual framework, so Medicare payment programs involve utilization measures. They involve clinical processes of care. They involve health outcomes, patient safety outcomes, and then, patient experience outcomes. They also involve measures of Medicare expenditures, or cost.

So again, these were the ones that we were constrained to thinking about when we were preparing our potential framework and when we were doing the, sort of, literature review that we did, or scan of the literature. Next slide.

So we called them literature retrieval,

because it was really, for a very specific purpose. Remember the charge to the committee and remember the goals of Report Number 1, it was to identify, sort of, factors. And we ended up calling all of these social factors, by the way, so you'll probably see the term social factors all over the first report and will be in subsequent reports.

That was the phrase, or the term that we decided on to capture, both, SES and all the other stuff that I showed you in the picture to the, I think, the legislation, or the IMPACT Act, specifically, referred to SES, but we were also told that they wanted us to go beyond that, and so, or at least that's the phrase. So that's why we have used the term social factors, to cover everything and have socioeconomic position, as one of the terms of ---

(Telephonic interference.)

MEMBER ESCARCE: Anyway, this retrieval was conducted by a professional librarian and was submitted to studies on U.S. patients, fairly recent ones in fact, and it was focused on

the social risk factors on health care utilization outcomes, such as those that are used in Medicare payment programs.

And then what we did, really, was to just find articles and describe them, generally, and look at them, generally. This was in no way a systematic literature review that tried to assess the quality of individual studies, or either tried to, to assess the bias that can be adhering to them, because of either the population studied, or the method used.

There was no attempt formal, or informal, to integrate the data, because the purpose here is simply to, to be able to make the statement that there was evidence that the factors, or at least some of the factors, listed in our conceptual framework, might affect outcomes of, you know, the outcomes that I talked about.

So I think that, if the, I'm really stressing this, because -- it's very likely that they'll go, oh what kind of literature was this, you know, I mean, some of these articles are terrible

and you missed these two really good articles, and, and that wasn't the purpose, and so they were, sort of, like, I guess, you might call them existence boost, so can you find evidence that there is a relationship between variable A and outcome B and that was the goal here.

And so the last bullet, basically, reiterates that, which was to say that the discussion in the report should not be mistaken for a systematic review that uses any form of system for waiting and describing the evidence, or integrating it, and so forth.

And then, the last slide. So this won't come, the last slide is, there. So these are, this is how basic and fundamental our findings are and getting these without doing this work, but, of course, we had to do the work, in order to be able to write them.

So the first finding is that all other things being equal, the performance of the given health care system, in terms of its quality and outcomes of its cause can be effected by the social

composition of the population that it serves. And that's, sort of, the, the conceptual big picture statement.

And then the other finding is that it's clear, at this point, is that health literacy and the social risk factors that we considered in the framework have been shown, so that's, kind of, of the existence proof that it shows, for instance, health care use cost and health care outcomes of Medicare beneficiaries.

And that is where the first report leaves off and there's an awful lot more pretty interesting work to be done and that's where we're engaged in now. I think that's the last one.

MS. JOHNSON: Thank you, Jose. Appreciate that, very much. We have some questions in the room, so --

CO-CHAIR PONCE: I think, Lisa and then, Nancy. Lisa Iezzoni.

MEMBER IEZZONI: Jose, hi. It's Lisa Iezzoni. Thank you for that report. I, as you know, was one of the people, who reviewed the

report, for the NLM.

And I am a little bit concerned that disability is not mentioned, as a risk factor, as a social factor, and I think that, you know, I looked at the membership of the committee and I wasn't surprised, because there really are no experts on disability on the committee, but I do think it's a very important thing to include.

There are 57 million Americans with disabilities. They are a very heterogeneous population. So it's not a small number of people for whom this is an important issue and I really hope that the committee, for their future work, could consider including disability as one of your social risk factors.

MEMBER ESCARCE: I really appreciate that comment and, parenthetically, I didn't know you had been one of the reviewers. But, but I do appreciate your comment. And I don't know if you've had direct contact with the staff at the IOM, but --

MEMBER IEZZONI: Well, actually, --

MEMBER ESCARCE: I --

MEMBER IEZZONI: -- you know, I, they actually asked me to be on the committee, but then they looked at my conflict of interest profile, and --

MEMBER ESCARCE: Oh.

MEMBER IEZZONI: -- and I'm on the Board of Commonwealth Care Alliance, which is a private nonprofit insurer and health, health care delivery system that serves dually eligible people, and so they --

MEMBER ESCARCE: Yes.

MEMBER IEZZONI: -- decided that because of that, I couldn't be on the committee.

MEMBER ESCARCE: Sure. Well, I'll certainly relay that message, if you haven't done so already, but I'll do it, anyway. Thank you very much for the comment.

CO-CHAIR PONCE: We're glad you're on this committee, Lisa. Nancy.

MEMBER GARRETT: Jose, this is Nancy Garrett. I just had a question about the report.

I was interested that you chose to use socioeconomic position, instead of socioeconomic status, as your, kind of, concept that you were measuring, can you talk a little bit more about that, I'm not sure I really understood the distinction?

MEMBER ESCARCE: Yes. I think that, first of all, socioeconomic position is a term that's much more popular and commonly used in certain demographics and sociologic circles, whereas, we tend to use, and by we, I mean, you know, health services, researchers, medical doctors, even economists, are more commonly the, the term that slips out is socioeconomic status.

I think the word, position, is intended to illustrate that it's not only a measure of your command of material resources, for example, through your financial means and stuff, but that it actually is part of a hierarchy in this society.

That that hierarchy is, has a number of dimensions, one of them is precisely these things like income and wealth, which, you know, measures, or, or, at least, gives some sense of your ability

to, to command material resources, but also, you know, your educational status, or your occupational status, and so you have more power, you have more autonomy, you have more influence on all sorts of processes in society, et cetera.

So that's what it's intended to capture. So status doesn't, I mean, it could do that, but it doesn't, normally, and, but socioeconomic position does that, or it could, so it's a favored term for demographers and sociologists.

MEMBER GARRETT: Thank you.

CO-CHAIR PONCE: Susannah then, Eduardo.

MEMBER BERNHEIM: Hi, just quickly. I was wondering where and how you guys talked about this issue of, sort of, how these factors overlapped with quality, how that came out in the discussions, or will come out in future reports? Because it doesn't, exactly, show up in the figure.

MEMBER ESCARCE: I'm not sure what you mean, by overlap, can you clarify?

MEMBER BERNHEIM: Well --

MEMBER ESCARCE: Because the model, the model, the figure is intended, so the figure, it, you know, it's like any other of these conceptual models with boxes and arrows and stuff.

It's intended to give a stylized and simplified view of how different things can affect other different things, and so the idea is that all of these factors can affect quality measures.

Each measure, the effect on each, I mean, for some of them, there may be no effect and we've talked about, or, some of those today, and others, it may be different factors that affect them.

But, but they also was intended to capture the idea that, in a, sort of, causal pathway, some of these factors may lead to differences in quality measures and the quality of care that people get. So that's why I don't understand the word overlap.

MEMBER BERNHEIM: Okay. No, so I think, actually, I think this comes back to one of, or maybe it was Phillip's comments earlier, but, no,

it was Tom talking about, sort of, quality versus outcome.

So when we use outcomes, as a measure of health care quality, then, sometimes, we're, sort of, it's confusing when we're talking about, so it's, you know, it's sort of an unadjusted outcome versus an outcome that has been, where you feel like you've accounted for things, aside from health care quality, and you're trying to illuminate quality, and so the outcomes can be, kind of, considered both things, depending on how you're talking about them, and I think that that causes confusion.

So, so I'm just going to say, actually, what I heard and then, which is that, the concept behind that figure is that all of these things can influence quality, as measured by current measures that Medicare uses in payment programs. Is that a fair restatement of what you were saying?

MEMBER ESCARCE: They can influence, yes, they can influence the measures that's right. And, and like I said, of the, which of these variables influence the measures, both from a

conceptual standpoint and empirically, can differ across measures supports.

And the mechanisms could differ, or the pathways, if you will. But, but it, you know, it's just one picture to, sort of, try to capture all, how these things can happen.

MEMBER BERNHEIM: Right. Yes, I think it was the mechanism piece I was, I was wondering about. And then, I don't know that you can talk about this, at all, but it seems like the third report, oh, the slide's not up anymore, but, as I'm not going to be able to quote it, quickly, but that has to do with, sort of, how the decision should be made. I think that's what, I'm going up to look at it on my --

MEMBER ESCARCE: Yes, so the third report is recommend criteria that the government should use, recommend measures that, at least, in our eyes, meet those criteria, and then, recommend something about methods of adjustment.

And you are correct that, well actually, well, let me just say that we've only started the

work, so there isn't much to say. But, but I can say what we have done, so far, because the IOM that's the way the IOM works --

MEMBER BERNHEIM: Of course. And it's only looking at the measures. The report's title and some of the conversations we've been having, have been around, should you incorporate socioeconomic consideration into how you pay providers, aside from how they do on measures, but literally, should that effect direct payment --

MEMBER ESCARCE: So actually --

MEMBER BERNHEIM: -- and that --

MEMBER ESCARCE: I mean, that's a really --

MEMBER BERNHEIM: -- some of this --

MEMBER ESCARCE: That's a really good question, because, I mean, the task of the Committee and the IMPACT Act talks about measures that are used in Medicare payment programs, and it talks about measures, specifically, but the Committee has, within its purview and within its, I think, within its mandate, to look at, actually, to comment

on, on payment methods, themselves.

So yes, the Act is aimed at, what do you do with measures, do you adjust them, how do you adjust them, what do you adjust them with, and these are measures that are used in payment programs. But I, my guess is that the Committee will also comment on payment, per se.

MEMBER BERNHEIM: All right, so it may be worth thinking about the benefit of incorporating these factors into payment directly versus the quality measures, themselves.

MEMBER ESCARCE: Yes, I understand. And we've had a lot of conversations about that.

CO-CHAIR PONCE: Anymore, thanks, Susannah, any more questions from the group? I have a question, Jose, and your primary, your Bullet 1 on your finding slide where all, all other things being equal, the performance of a given health care system, in terms of quality, outcomes and cost can, undoubtedly, be effected by the social composition of the population it serves.

So actually, I have two questions.

One, is that finding relevant across all the patients, or is it that, is that specific to the Medicare population?

MEMBER ESCARCE: I don't know that, I mean, I, my assumption, personally, is that that's relevant to everybody, but I'm not sure that it was intended --

CO-CHAIR PONCE: Okay, so --

MEMBER ESCARCE: -- to be, to be limited to the Medicare population, in that statement, although, although, of course, our charge is to look at the Medicare population.

CO-CHAIR PONCE: No, I think that's useful, because then we could, certainly, you know, borrow from your, reviewing your findings for our, for our committee.

And then the, the social composition part, so that, for me, as a researcher, that means it's the, kind of, community-level neighborhood, you know, proportion of minorities, or even getting at residential segregation-types of measures, what's under, what's behind that social

composition?

MEMBER ESCARCE: Yes, so I'm trying to look at, I'm calling up the slides on my computer, so I can look at this thing. So I think, you know, so, so terms are written and sentences and paragraphs are written in reports and you, actually, read them pretty carefully, or, at least, in a lot of cases, in this case, I think many of the committee members did.

And then, you look at a paragraph again, like I'm doing now, look, and you're pointing out and you, and you wonder, yes, I wonder why we said it that way?

But, I think, this just means the composition of the, of the people, who, of the patients, or, I mean, it says the population, it doesn't say patients and it says social composition and, and it isn't very clear, as to what that means, but what this really means is, the health care takes care of a whole bunch of people and what those people look like, in terms of the factors in that picture matters. I think that's all it means.

CO-CHAIR PONCE: Great. Thank you.

MEMBER ESCARCE: Yes.

CO-CHAIR PONCE: Okay. I'm not seeing any cards up. Obviously, that was a very clear presentation. Jose, thank you. And we're going to move on with Helen and Karen, to present Christie's presentation.

DR. BURSTIN: Yes, as Tom was saying and, you know, clinical medicine, it's see one, do one, teach one, so I've now heard Christie's slides twice, so I feel quite confident.

But, I'm, actually, not going to go through it in great detail. You have all the details, it is really an exquisite, very detailed, statistical analysis.

Stats questions are for Karen, I'll give the high-level overview. But, essentially, I just want to focus on a couple of the key slides here. So next slide.

So one of the big issues that has come up, and interestingly, these measures are not required to be part of the trial period, this is,

these are measures already endorsed by NQF from the Pharmacy Quality Alliance, PQA, we've probably heard that acronym a few times in the last couple of days, which are about medication adherence.

And I'll say, from the Committee discussions, whenever we had it, you know, the SES discussions, medication adherence kept coming up, as a good example of measures for which you could logically, certainly, have a conceptual basis that some of these factors may be in play.

So Inovalon's been working with PQA. These measures are used for Medicare Part D. So these are used to assess the performance of the Pharmacy Benefit Plans associated with Part D.

And this is all involved, and Christie could give you lots on this and, actually, this may be where Traci, or others, could help, or, or certainly, Ron, around the plans, issues, or around understanding the issues with the stars.

But, essentially, what she's done, and I'll walk through here, is pulled in various different data on sociodemographics, and you can

see that listed on the bottom here, issue data on income, education, household size, poverty area, physician shortage area, to really begin to see, whether among the three medication adherence measures they have, which are hypertension, diabetes, and lipid control, how much of an effect are they seeing when they, in fact, adjust for SES, and what are the issues that they have encountered.

So next slide. And they did this across 44 Medicare Advantage plans, so it's a very, very large sample size, millions of patients when she's presented this before.

I won't get into this very much, other than to say that, they were able to pull in data from the ACS, the American Community Survey, and others, to begin to look at data points, at both the individual person characteristics, as well as, interestingly reflecting on the last conversation, the behaviors to the near neighborhood characteristics, as well.

They were able to move beyond five-digit zip and had access to the nine-digit zip, sometimes

called five-digit plus four. And they also had data from the area health resource file on some of the community resources that were available. Next slide.

So quickly, this is a really broad overview slide of the way they laid it out, and all of those slides are, somewhat, similarly laid out here, that allows them to look at the unadjusted rank of the plans, before SES adjustment and then, when they're able to incorporate the SES adjustment in.

And this is a theme you'll see through most of this. Essentially, what they have found, which is, I thought, quite interesting, is when you adjust, the plans ranked best, tended to stay best.

That's, I think we had similar discussions, as part of the SES Panel, as well. Those ranked worst, continued to look the worst. Adjustment didn't change the top, or the bottom, but it had a fair amount of movement for those in the middle, for which some of those factors may really be in play.

And it's probably too small to see them, but she does, specifically, pull out, for example, one of the little triangles there is Plan B that moves from a rank of 22 to 28, but it ranks higher, if you don't consider the low income subsidy for the population they share.

In addition, Plan A moves down from 57th to 50th. And these numbers may not sound huge, but as you're thinking about it, in terms of stars, it changed the star ratings for those plans, which has, certainly, significant implications. And this is using low income subsidy. Next slide.

So this, she's going to, in this handout you've got, she goes through, in exquisite detail, what they added to the model, what they found in each of these.

She's, mainly, going to focus in on the MAH, which is the analysis, specifically, for adherence measures for patients with hypertension.

And these are all fully claims-based measures, so their measure of adherence is prescription fills, did the patient get the, next,

prescription? Next.

So these get a little bit small, you've got them, as well, but a couple of key things. They found low income subsidy, which is something LIS people use a fair amount of these analyses, was not significant, after they put Medicaid status in.

They didn't have language for most of the member-level data, for example, and the shortage area, some of the data they were able to get from the area resource file, also, was not significant. Next slide.

And then, when they started to begin looking at some of these other issues and, Lisa, you may find some of this interesting, they, specifically, also focused in on the issues of disability, not surprisingly, disability, they were able to say here, for example, disabled beneficiaries less likely to be adherent, younger disabled measures least likely to be adherent. And goes through some of the analyses on this page. Next slide.

Here, they went through some issues

around gender and race and, here, they also looked at collinearity here and interaction terms, but they did find lower odds, and again, this is African American males and hypertension, a significantly lower odds for taking, for being adherent to blood pressure meds. Next.

In this next one, she looks at dual eligibility, again, having data here and found in overall dual eligible patients were more likely to be adherent and that, there were some differences here that I won't get into, between partial and full dual beneficiaries, because it goes beyond my understanding.

But what's very interesting here is looking at, just extraordinarily, the number of different meds dual eligible patients are on. I mean, she's presented this, they are frequently finding patients on 16 plus meds.

So pulling out adherence to, you know, hypertension, lipids, and diabetes, probably, isn't, really, the story, either, it's the comment, I think, it was Traci made earlier about, you know,

it's not just adjusting the measures you have, but are they actually the right measures, and this might be an example of that.

They then looked at some of these -- I'm sorry, next slide. They then looked at some interesting issues, again, at the community factor levels, since we just had that discussion.

So they did look at a variable looking at people, numbers who lived in areas of high-level of home ownership, which was, sort of, a similar one to, to people talking earlier about some of the work out of Missouri and Steve Lipstein's work, where they looked at a proportion of houses, a proportion of housing, within a given area that was unoccupied, and so this is the flip of that.

If you lived in an area with high ownership and they were more likely to be adherent, poverty, less likely to be adherent, and education, as well having an effect. Next slide.

So again, this was their overall summative work, but I've already mentioned this, this key issue, but, essentially, what we're seeing

is that, the folks in the middle are having the most movement and, at times, it appears that some plans that are providing lower quality of care, based on the initial assessment, looked significantly better when some of these factors were considered. So last slide and I'll wrap it up.

This is the summary, and I didn't do it justice, but, essentially, they did find that when they were able to look at the different SES factors, stratification by the low income subsidy, alone, didn't actually have much of an, did not change the percentile rank for most plans, and that's what we keep hearing is readily available, variables like LES, and five-digit zip.

And, again, I've already mentioned the point about which ones change. They did find that non-duals, who are poor, have worse outcomes than dual-eligibles, who are poor. And she can, certainly, talk more about that, and that may have to do with the benefits associated with being dual status.

And they, their point here, is that it

underscores the importance of adjusting for income and poverty, beyond dual and/or LIS status.

And in the other work she's demonstrated, she has a consistency here of the key variables and I'm not sure if she has it in this set, but she's also found nine-digit zip code, in particular, to be a very robust indicator of poverty, much better than five-digit zip, explaining pretty significant differences in variation in the, in the overall work.

So income and education are significant predictors, even after controlling for dual status, age, disability, interactions and other variables and, particularly, the ones we tend to have access to, like, LIS and five-digit zip.

So she's not here to answer your questions, although she indicated a desire to follow-up and have further discussions. But I thought it would be helpful, just to hear a flavor of somebody actively out there trying to do this work and, kind of, struggling with some of the discussion.

CO-CHAIR PONCE: Emilio.

MEMBER CARILLO: Yes, maybe, Traci can --

(Off microphone comment.)

MEMBER CARILLO: What --

(Off microphone comment.)

MEMBER CARILLO: I want to ask now, what constitutes quality for the plan, what, what is the, this, what's behind that ranking?

MEMBER FERGUSON: For the Medicare Advantage Part C and Part D, it's the star ratings. So they put in a lot of the endorsed measures in NCQA, HEDIS, a lot of NQF measures that are already endorsed.

They have, every year they had changed their waiting, but that pulls into a five-star ranking, an so that's how they measure quality for a Medicare Advantage plan.

MEMBER CARILLO: So does that include the net worth, the density of the net worth, the --

MEMBER FERGUSON: Correct.

MEMBER CARILLO: -- the GEO access --

MEMBER FERGUSON: GEO access, a lot of, you know, --

MEMBER CARILLO: -- or location?

MEMBER FERGUSON: -- call to answer, if you get any complaints, some administrative, but a lot of these process and outcome measures, the CAHPS, the HOS surveys, play a point, a part in that.

CO-CHAIR PONCE: Bob. Bob and Traci.

MEMBER RAUNER: Speaking for someone, who just knows just enough about statistics to be dangerous, sometimes. Maybe a little bit, I think, David, you touched on this a little bit that, obviously, there's lots of P-values here that the difference between statistical significance and clinical significance, so definitely there's a difference.

But, from say, like, looking at a perspective of the family doc, taking care of a panel of patients, is it the difference between $100 and $150, or is it a difference between $100 and $101.22, which is just not enough that I'm going to change my practice much, or how I approach patients.

And it looks like some of these things, yes, if you had a bunch of these things together, or if you're poor, plus African American, plus this, plus this, may be enough that I might send you an extra reminder, maybe, but beyond that, I'm not, I was, how much of these are things that are actually going to get me to make a change in my workflow and my practice, because of this study? Maybe it's too big a question that --

CO-CHAIR PONCE: It depends on how many. So, Traci.

MEMBER FERGUSON: Well the, a lot of this, in terms of adherence, there's maybe a pay for performance that would be included, and we've done things to incentivize the providers, also incentivize the members to go to the providers, but it could be a bump. Again, it depends on how many you have.

But it is, you know, for our population, in terms of Medicaid, it's almost like a, paying for an extra visit, a Medicaid visit, so it's, it could be a substantial amount of money.

MEMBER RAUNER: So if I open up QHC, this could adjust me enough that I get put into a different, you know, PMPM tier, or maybe we'd get an additional $10,000 bonus, if I'm part of an ACO. That's real money. That might, actually, --

MEMBER FERGUSON: Correct, yes.

MEMBER RAUNER: -- pay our half of our care coordinator for our office, for example.

MEMBER FERGUSON: Correct.

MEMBER RAUNER: Okay.

MEMBER FERGUSON: Correct.

DR. BURSTIN: Just to clarify, these are actually plans, so these are not clinicians. And I don't know what these would be extrapolated to, if we looked at clinicians.

But, in particular, when I've talked to, I've talked about this a lot in the last year, last couple of years, probably, the topic I've talked about most, the special needs plans, in particular, the SNPS, like the one Lisa is on that was, ironically, excluded from the IOM Committee, have come off more characterized, as a group, you know,

as a, as a special needs plan that, specifically, focuses in on the most vulnerable patients.

And so for them, they find this work really key, because some of them are trying really hard to do the right thing and it's been hard, because the adjustment hasn't been there.

MEMBER RAUNER: Yes. And so to follow that, to make this very real, is that, although this is meds in our community work, exploring with our Medicaid Managed Care Plan, utilizing our FQHC, these things are, are just as applicable to what our Medicaid Managed Care Plan might do to us, or more helpfully, help do with us, not necessarily to us, but, you know, this is a very direct application to what everybody's, who's trying to do these value-based projects, is working, so.

CO-CHAIR PONCE: So just --

MEMBER ESCARCE: Can you put me in the, can you put me in the queue?

CO-CHAIR PONCE: Yes, in fact, I, I just wanted to check if your hands, if Lisa's hands went, was up, too, or -- okay. But, Jose, you're in the

queue. Traci, you had responded, but is this a new point?

MEMBER FERGUSON: Yes it's a new point. I wanted to see where this fits and is this part of the, sort of, the trial period, or how is this study fitting in with --

DR. BURSTIN: So we just wanted to give you an illustration of something that's out there. Those measures are all due for measure maintenance, which means they have to come back to SES, or --

MEMBER FERGUSON: Okay.

DR. BURSTIN: -- they have to come back to -- I can't even say the name of our organization, anymore. They have to come back to NQF and they will need to, you know, they're just, kind of, getting ahead of the curve of having this work done quickly.

MEMBER FERGUSON: And this is the, the way I was, again, thinking about either the cardiovascular measures and the other measures that have already gone through, is putting, again, a database, you know, they're looking at I don't know

how many plans that pulled that information to get to this difference that we see.

That, could that be, sort of, a means of, well let's put all the data, so then you'll have the individual data, individual, sort of, at the member level and all of these social factors and risk factors, and then, put it in a pot, then have someone do some analysis.

So that could be a way to get to this. It may not be through this trial period, but eventually, if we have enough people gathering that information, we could get to something where we can see a statistical difference.

CO-CHAIR PONCE: Thank you. So, Nancy, Sarah, and Jose.

MEMBER GARRETT: I just wanted to comment on the data that they used, so they used this data from ICM, a company that aggregates data from the census, but also, lots of other things, like, magazine subscriptions and credit card data and they know a lot about us.

So I just thought that that was an

interesting point, because we've been really struggling today with how do you get this data and I think that they would create a, I don't really know, I would ask, Christie.

How much, how many of the variables that were used were actually enhanced by that, or was this all stuff that was in the census? I'm not sure, but I do think that that is something that we should keep in mind that there, there are these other data sources.

I know, you know, various health plans and providers have experimented with, kind of, using this in health care, but I don't think it's very mainstream, yet, partly because you have to buy it, but it's interesting.

CO-CHAIR PONCE: Very good point. Sarah.

MEMBER SCHOLLE: Could, Traci, can you just say a little bit about the payment policy, because aren't the plans paid differentially to, I know Medicare stars, so another component here is, how do you handle this, do you handle it in the,

through a different star rating approach, or do you handle it through payment, or do you handle it through risk adjusting the measure? So there, there are multiple options and so CMS does consider some of this in the payment.

MEMBER FERGUSON: There is a bonus payment that gets doled out to Medicare Advantage plans that hit the, starting at the four star, so four and five, they get bonus payment, depending on their membership.

For those and then, for those who are three-star, well, less than three stars, there is a risk that you, if you're three-star, less than three stars three years in a row, that CMS could remove you from participating, so again that's loss of revenue there.

MEMBER SCHOLLE: But the base payment is based on the --

MEMBER FERGUSON: The stars one.

MEMBER SCHOLLE: No the base payment, not the bonus. The stars are part of the bonus, but the base payment that you get per beneficiary is

based on an HDC score --

MEMBER FERGUSON: Correct.

MEMBER SCHOLLE: -- and other stuff. So there is some complexity that's already handled in the payment mechanism. So one alternative to addressing this measure-by-measure sociodemographic risk adjustment, one approach would be to include sociodemographic risk in --

MEMBER FERGUSON: In the base payment, correct.

MEMBER SCHOLLE: -- the base payment, rather than in the measurement.

MEMBER FERGUSON: Yes.

MEMBER SCHOLLE: And then that would be encouraging plans to do, to really focus, not just on measure-by-measure, but to focus on that population, as a high risk. So there, just to illustrate an alternative approach to using this information, to encourage efforts on the part of the plan.

MEMBER FERGUSON: And then, I also think that, you know, even with gathering that

information, because, again, the providers may not have the resources to make sure that they have everything on the claims to get the correct risk, I mean, HCC score that that's where we see even some difference.

MEMBER SCHOLLE: Yes. And in fact, because it's based on HCC where, which is all coming from the claims data, if CMS were to think about, well can I take something about where this person lives and say, I know that's a community --

MEMBER FERGUSON: Right.

MEMBER SCHOLLE: -- that's harder, and then adjust the payment, based on that community and the individual risk factors, rather than just paying based on the clinical complexity, which is the HCC score that that might be a way to improve quality more generally, rather than focusing it on one measure at a time.

MEMBER FERGUSON: Correct.

CO-CHAIR PONCE: Jose, you're next, but I'm going to ask, Cara, to go first.

MS. JAMES: Yes, thank you. I just

wanted to provide a little bit of context, because you guys have presented some of the data that Inovalon has done and you presented some of what NQF has been working on.

This is also something CMS has been working on, and so if you haven't seen it, already, in the request for comments on the Star Ratings Program for 2017, you can, kind of, review some of what we've done, thus far, and where we are, and that's available, I can share that with you.

But that is, because we've looked at low income subsidies and disability, as well, and the impact that that has on star ratings and so where we are with some of the quality pieces for that, and so I would encourage you to review that, as you're thinking about this, in the broader context, as well. But that is more information to help provide context.

CO-CHAIR PONCE: Great. Thank you. Jose, you're on.

MEMBER ESCARCE: Yes, I had a question, and I'm worried this might have been presented and

I just spaced out and missed it. But what was the average value of this measure of adherence?

I mean, was it very, very high across the board, and what was the spread between these plans? Because, of course, somebody has to rank first and somebody has to rank last, but, but the spread in the values might have been, either, very big, or very small, I just wonder if, if we know that, or if you know that?

DR. BURSTIN: Jose, this is Helen. Since I'm playing Christie, I don't want to make it up, I don't know that. I don't --

MEMBER ESCARCE: So we don't know, or even like the mean value across the plans, or anything like that?

DR. BURSTIN: I don't recall it well enough to feel like I should give it in a meeting with a transcriptionist that's public.

MEMBER ESCARCE: Oh.

DR. BURSTIN: Yes, no.

MEMBER ESCARCE: Okay. All right. That's a good, honest answer.

DR. BURSTIN: Well, but we'll certainly ask Christie and we'll send an email around to some of these questions that came up.

MEMBER ESCARCE: Yes.

DR. BURSTIN: We'll make sure we get her, she just had to get on a plane.

MEMBER ESCARCE: No, but, I mean, obviously, it matters, right, if everybody's between 85 and 90 percent and these adjustments move you up by a fraction of a percentage point, it could easily flip your, your order and your rank. And, actually, it also touches on the question that was asked about clinical significance, you know?

DR. BURSTIN: Yes.

MEMBER ESCARCE: It's, if people tend to be in the 80s or 90s, even a notch ratio of .5, which seems awful, actually only changes your, your adherence by, you know, four or five percentage points. So it's important to, kind of, know where you are on the zero percent to 100 percent scale.

MEMBER FERGUSON: And, this is Traci. I don't know the exact value, but I can tell you that

it is wide and it's not at the 90, above 90.

MEMBER ESCARCE: Okay, that's good, yes.

MEMBER FERGUSON: Yes.

MEMBER ESCARCE: Yes, I mean, most of the time, people don't adhere at that rate, so I didn't anticipate that, but I am curious about what the total was.

MEMBER FERGUSON: And most, I mean, you'll look at sort of, the cut point, and a lot of this is, like, 50th percentile, so you think about, sort of, stars measure, there's some that could be, you know, will put you in, like, the, like, three stars and it'll be, the cut off is, like, 80, but most of these are, like, 50 is the cutoff, so between --

MEMBER ESCARCE: It's okay.

MEMBER FERGUSON: -- like, three, that will get you to a three so --

MEMBER ESCARCE: That's good.

MEMBER FERGUSON: -- that's why.

MEMBER ESCARCE: Thank you.

CO-CHAIR PONCE: Thank you. Any more questions? And just to make sure, Lisa didn't put her hand up this time?

MEMBER COOPER: I didn't. Thank you, it's been a very rich discussion and I'm definitely learning a lot. I love adherence, I think it's one of the, really, ones that, one, one of the ones that providers can actually wrap their heads around.

Again, but the adjustment is like so important, because of what we hear from providers about how, you know, no matter what they do, if patients are dealing with all these other issues, they feel like they get punished for things that, you know, they can't, they can't actually control, even though they're working very hard to address those issues.

CO-CHAIR PONCE: Yes, thank you. We're going to, to have public comment now, the NQF Member and public comment. Open that up, please.

OPERATOR: And, at this time, if you would like to make a comment, please press star, then the number one. There are no public comments

from the phone line.

CO-CHAIR PONCE: Thank you.

CO-CHAIR PONCE: The next session is the discussion of the trial period evaluation plan. So, I think it's very important, the concrete steps of how to evaluate this robust trial.

And, Karen is going to take us through some key questions of the session.

MS. JOHNSON: Thank you.

So, we are going to step back just a second, and we wanted to give you guys, in case you are not as familiar with our evaluation criteria, we wanted to make sure you see them. For those of you who are looking at our website, this is a late addition to our slide deck, so you won't have it, but you can look at the developer -- I'm sorry, the Steering Committee Guide Book page 29, I believe. And, that's available through the link in our webcast.

But, we talk about the disparities really in two major places in our criteria. The first, just to orient you a little bit, we have five

major criteria that we ask committees to use when they are looking at and evaluating the measures that come through.

The first is the Importance to Measure report, and under that there are, for the most part, two, but sometimes three depending on the type of measure, sub criteria. So, the sub criteria help answer the question of how you know if something is important to measure or report.

And, the other thing I do want to point out is, and I alluded to this earlier, the hierarchy of the criteria. So, we talk about importance to measure and report first. So, that is what we would call a must-pass criterion, and the sub criterion underneath it are also must- pass.

What that means, you know, in our actual meetings is if we are looking at a measure and it does not pass the evidence of criterion, that we just stop the session. We do not go on and continue through the rest of the criteria.

So, we look at evidence. Then we talk about performance gap, or opportunity for

improvement. And, that is one place that disparities conversations can definitely come up. To demonstrate that there is a performance gap, you could, potentially, say that, and show data, to indicate that pretty much everybody is doing a poor job, or you could show there's a lot of variation between entities that are being measured. You know, some are doing great, some not so great. Or, you could show disparities they have, so indicating that certain sub populations maybe the performance isn't as good for certain groups than for others.

So, sub criterion 1b is where we get a lot of information, when we get information on disparities.

Then we go on to scientific acceptability of measure properties, and the key ones there are reliability and validity. And, we, actually, talk about the specifications of the measure in both of those. Under reliability, we want to see very explicit, precise specifications, but then we go on to think about consistency of data collection and ability to distinguish differences

in performance. And then validity, that speaks to the accuracy of the data, or the correctness of the conclusions that can be made about quality of care.

And, it is under validity that we talk about specifications. So, we expect the measures to be constructed so that they conform to the evidence. We asked about testing, so we do expect testing to demonstrate validity. And then, we go into the section that we call Threats to Validity, and it's that section that we talk about things like exclusion, and missing data, but also risk adjustment. So, that is the place where the SDS discussions would come, so under scientific acceptability, specifically, the validity section.

We talk about feasibility. That has to do with really trying to minimize the burden of data collection and implementation of measurements. You'll notice that that one is not a must-pass criteria, but it does come into play. It's coming in to play even more, I think, as we head into the direction of e-measures, or ECQMs.

And then finally, the fourth line,

usability of use, that's where we, actually, want to see whether measures are being used, and what kinds of applications, whether there has been demonstrated improvement in the measure. So, are providers, actually, improving on what is being measured.

And then thirdly, benefits out to patients outweighing evidence of unintended negative consequences. So, we do want to make sure that actually doing that measure is not hurting patients in a way that would make it unwise to do it that way.

Finally, we look at comparison to related competing measures. That is getting to the idea, if a measure has made it through the four things that you see above, then we may go on to talk about comparison, comparing related or competing measures. The idea there is, there's a lot of measures out there, and if there are duplicative measures, or measures that are similar but different in the way that they are specified, is there -- number one, is there a need to have

duplicative measures, and sometimes there are, or if there are what we would call related measures, but specifications are different. So, for example, one measure defines diabetics one way, and another measure defines them differently, is there a reason to have that kind of variation in the measure. So, that's when that kind of discussion comes up.

So, I really just wanted to give you this to kind of give you some context about what we are asking our committees to think about. This gives you a flavor, and Susannah is going to send around -- okay, Michael has sent to the entire committee the submission form for one of Susannah's measures, I guess -- you sent a blank one. Okay. That will give you a flavor, that is the submission form. So, we have to ask a lot of questions so that we can get the information that we need to have committees be able to discuss these criteria.

And, you can imagine that folks who are new to NQF may, you know, sometimes include more or less information. So, we don't always get

information on disparities, or we may get it in a very rounded out way. We are not prescriptive, in terms of what -- other than, we ask for these kinds of information, but we don't say it has to be from this kind of a source or this kind of a set of literature, et cetera, et cetera.

So, let me pause there and see if anybody has any questions about our criteria, and it looks like we might have one.

Lisa.

Oh, okay, no questions.

CO-CHAIR PONCE: Nancy, I think that's an old data.

MS. JOHNSON: Yes. Okay.

Let's go on to slide 109, Michael, please.

Happy to answer any questions if they come to you later on. But again, we just wanted to put that in context, so, you know, the ways in which we discuss these things in our meetings.

MEMBER COOPER: Are the slides, actually, being projected right now, because all

I'm seeing is one slide.

MS. JOHNSON: Okay, here we go.

MEMBER COOPER: Thank you.

MS. JOHNSON: Yes. So, now we are back to our regular slide deck.

I wanted, before we get to see the evaluation plan exactly, I did want to give you just a little bit of input from stakeholders, and this is mainly coming out of our more recent costs of resources project. That's the one that's made it the further-est through the process. And, some of the input has been the limited availability of patient level data. We've talked about that very much. The nine-digit zip code, or Census block data not easily accessible. That's, actually, one of the reasons that we wanted to make sure that you saw some of Kristy's work, because again, they have bought that data from an external vendor, so that might be an option. Some developers might not even know that's an option.

So, concerns about factors that are selected and analyzed to date. Basically, the

concern, the variables that are being used, and I use the term proxies here, that might not be the best term there. But, if your conceptual rationale tells you that income may have something to do with your outcome and interests, is dual eligibility standards really getting to that construct that you are trying to show.

And again, some of Kristy's work suggests that maybe there's other things in additional to dual eligibility status that might need to be looked at.

There is some discomfort, I think, in using race in some of the modeling. The SDS panel did make it very clear that they thought that race should not be used as a proxy for SES, but they did not say you shouldn't look at race ever. But, that has come up.

And then, there is a call to some extent for a more prescriptive approach from NQF, even so much as saying, here's the five, or ten, or whatever variables that everybody should look at every time.

So, the next slide, please.

We keyed up these questions, and I'm not going to have you go through these questions right now, but these are things to kind of think about and we'll probably get to them some in our following conversation. But, would you guys have recommendations about the use of variables that are currently available? Should we take a more prescriptive approach to variable selection? And, I should probably also add in methodology approaches as well, that could be something that you may want to think about, and how can NQF help encourage the development and innovative approaches to SDS adjustments.

Dave.

MEMBER NERENZ: This is just going to be a broad observation, but in the -- to me it's to explore, and I think it could be within our purview to do some of the exploring, is the distinction between community level variables and individual level. I know much of what we've done so far, and much of the examples in front of us, involve individual level, including the concept of taking

a Census level variable and using it to impute an individual characteristic. In the end, what you've got is individual level analysis.

And, that's fine, and, you know, we should take that wherever it goes. But, I think also we need to recognize that there are factors that aren't truly community level factors, on the side that begs the question, what's a community, okay, technical issue.

But, you know, it seems like the available resources, local transportation, how good is the Meals on Wheels program, things like that, these are not characteristics of individual patients, other than, perhaps, to say the patient lives in a community with these characteristics. So, maybe the boundary.

But, I think as our work goes forward, and we have examples come forward, I'd certainly be interested in exploring, or seeing developers explore, explicitly community level variables, and see how that works. There are some published examples of these having, actually, fairly powerful

effects. But, we ought to take a look. I don't think there's anything on entire group recommendations that say that should not be done. I personally think it should be done.

MS. JOHNSON: And, I will say we are batting around some of the ideas for upcoming webinars that we've had, and I was kind of hoping that that might be something of interest to the committee. So, we'll get further input from you on that.

Let's go to the next slide. And, actually, in the next one, and let's, actually, talk about our evaluation plan.

So, there are limitations to our trial today. First of all, we've got two years to do something here, and we don't develop the measures, or implement them. We look at what comes in our door, basically.

We control what's required for submission. We control our criteria, and, mostly things like this get endorsed. But, historically, NQF has not been prescriptive about things like

methodology, sample sizes, thresholds for reliability, that sort of thing. In general, we are not prescriptive.

We don't, at least at this point, have additional funding to do special research. There are data limitations, and another, I think, really great point is that different developers have a range of expertise. Some may, I'm being a little hyperbolic, I don't know if that's a word, they might not even exactly know what we mean by "risk adjustment".

So, when we do this outreach and say, you know, include SDS factors in your risk adjustment approach, that might not be as clear to some people as it is to others.

And then, different developers maybe are or maybe are not the implementers of measures, so they may not have data to be able to show some of these things that we would like to see.

So, let's go to the next slide.

So, in ten years' time we need to evaluate the success of the trial. And, you know,

really our primary question is going to be the temporary change that we've instituted, should we make that a permanent change. That's our main question.

But, I think our secondary question is probably, you know, how -- what things can we learn, have we learned, and can we share, that would kind of push the field forward. So, we are kind of thinking in those ways.

And, with the two-year limitation, the kinds of things that we can answer, at least right now, or we think we can answer, is which measures had the conceptual rationale, and maybe we can go further and say, what was that conceptual rationale. What variables and data were available, and then which ones were analyzed. They are not always the same, right?

If data are not available, was there a pathway forward? So, for example, with cost and resource sheets, and we find the speaking, but the committee knew that you couldn't do the nine-digit zip look, and kind of gave you -- they said don't

spend time fooling with the five digit, to be honest.

But, I don't know that they came back and said, you know, in a year, or two years, or three years, or whatever, we really would like to see the nine digit. But, in some cases that might be an option to do that sort of thing.

MEMBER BERNHEIM: No, I think that was implied. But the other thing, just as you reminded to go through this, measures come back for a full look again in three years, and every year we bring them back to see if there is any changes or updates. So, it's not like once it's endorsed it's out in the world forever. I mean, there's a very consistent re-evaluation of measures. So, there are opportunities.

MS. JOHNSON: We can certainly tell you, as we've done today, the numbers and types of measures that have been submitted with adjustments, and what happened with those evaluations. And, we can also speak, to some extent, about were the specifications for stratification also included,

because that's how you can be looking at disparities.

We've also considered soliciting feedback from stakeholders on the impact of the trial period.

Next slide, please.

But, basically, some of the things that, the longer-term questions, and maybe the ones that are most interesting to folks, are things that we don't think we can do. We can talk about the availability of the data of SDS variables and quality of that data a little bit now. I'm not sure how much we can do that now, and we think, or at least we are hoping, that that will change over time. There might not be a lot of change in two years. I don't know.

How do the healthcare entities react to the SDS adjusted scores and the stratified data in terms of their improvement in these, or how do purchasers and payers use these different scores in their programs. And finally, does this adjustment, actually, have any impact on

disparities.

So again, these are questions that we are not sure if or how we could answer these questions. We think they are out of our control, but we would like to have you help us think through some of these things, and maybe get some insight into these long-term questions.

CO-CHAIR PONCE: Karen?

MS. JOHNSON: Yes.

CO-CHAIR PONCE: Michelle had a comment.

MEMBER CABRERA: Thank you. I hope it's okay if I just jump in before you are done.

One question I have, based on the earlier presentation you gave about what it's been like to date. I think of this body as sort of the stewards of what came before. And, you know, at first I think, in terms of the things that we were supposed to be doing, the monitoring of the trial period was the one that seemed most daunting to me, and the one where I figured, oh, I probably won't have much, you know, engagement in that.

But today's presentation really changed

that course of thinking for me. So, one of the things that I want to just suggest or throw out is that, I think in order for there to be a good buy in and, you know, for us to do our job in such a short time frame, I mean, that's the other thing, obviously, you are bringing him and stressing now, is, you know, I think we need to, some of us, not everybody, certainly, but some of us, and I don't think it's just the advocates, I think there might be others of us, might need to have kind of a little bit of a supplemental, optional webinar training to go into some of why some of these things are landing the way they are right now.

So, I want to be brought along to help me understand better why some of these early conclusions, and to have the knowledge base to be able to either accept it and explain it to other people, or to push back on it.

And so, I just wanted to say, again, I don't think everybody needs it. I think there are some people here who can help with it, but if NQF would indulge, you know, an additional allocation

of your staff resources, and if some of our colleagues on the committee would also indulge us, I would be really appreciative.

DR. BURSTIN: Yes.

CO-CHAIR PONCE: Helen just said yes we can do that.

MS. JOHNSON: Let's go to the next slide.

And, actually, your question does kind of segue us into the discussion questions that I hope you can kind of bat around.

What would lead you guys to recommend putting our prohibition back. Is there anything that would make you think that we should rescind our policy? And, if so, what would that be, and what kind of information would we need to be collecting as we go to help you make that decision. What information, that's the second one.

And, are there additional questions that we should be able to expect to answer, and what data should we collect. So again, the things that I mentioned already, one that I think you asked earlier, it would be nice to see the little grid of

things, and that's, actually, what we are doing. As we are going through, we are filling out these grids so that we kind of know what's going on. But, there are probably other things that we could be doing, we just haven't thought of them.

So, I'll turn it back over.

CO-CHAIR PONCE: Okay. That's provocative.

Bob.

MEMBER RAUNER: I guess going through that list, part of my problem is I think a lot of my suggestions are in the realm of belief, not knowledge, because there's just not enough science there yet. It seems to me I think I'm guessing that that's probably true for a lot of people in the room. Some people like Kristy, actually, did that.

I really hope that SDS stays on and doesn't rescind this, unless there is overwhelming proof that it's causing some unintended consequence. So, I think personally it was a mistake to do that in the first place, but yet maybe, based on the level of understanding at that time it

was the right decision. I think we do, definitely, need to be studying this area and looking into it, because are potential unintended consequences. But, I think a lot of our recommendations, frankly, we need to start gathering the science on this, and synthesizing it somehow. And that might be the role of this standing committee, is to keep monitoring thing as things go, come up with the next layer recommendation, because IOM seems like it's pretty constrained. I don't know if they can make a full recommendation in their time frame.

Of course, I guess, when you look back some of those are due like in 2017, 2019, so I think they are safe. This is going to evolve quite a bit over the next couple years, too, so what is our charge in the next couple years as well.

CO-CHAIR PONCE: Sarah.

MEMBER SCHOLLE: So, I think there's a huge amount we can learn from the submissions. And, I think -- and, I, actually, think learning from what the IOM is doing as well.

So, I would suggest that since the IOM

has this nice framework, about which social risk factors are there, it would be great to have a summary of, to what extent, and the kinds of measures that are coming through, could we use that to help us.

Now, it doesn't include disability. We may want to change it in other ways, but I think trying to put the experience in the context of the IOM recommendations would be very helpful, because I'm curious to know, I, actually, know that we submitted some measures recently, in that pediatric call, but I think we only submitted process measures that we probably said something like this in the process measures that we now think.

So, it may have gotten in under that, I don't, actually, remember how we handled it. But, I think it would be good to really look at it and separate out the process measures from the outcome measures, and the point at which -- to what extent risk adjustment, actually, even came into the conversation, because I know I was listening to part of that pediatric thing. So, there was a lot of

other problems before they got down to that level.

And so, but I think trying to build this into the IOM framework and the variable would be good. The data sources, I'm very curious to understand how people use the data, what data they had, and to what extent that influenced their ability to identify factors or used factors, or even measure it.

And then, you know, what kind of risk adjustment methods are they using? What's the range of methods? So, it feels like, you know, kind of a deep dive into those applications so far would really be helpful for this group to understand, and putting it in the context of what the IOM committee is recommending, would help us to be able to say that's what the recommendations were from the IOM, should NQF follow those or diverge.

CO-CHAIR PONCE: Okay, thank you.

Nancy.

MEMBER GARRETT: So, I think this is a question, perhaps, for Helen, but this questions leads me to ask another question, which is, what is

it that the board needs to, you know, observe, consider whether to make -- lift the trial period and have this be part of the way NQF does its work, because if I recall, and Kevin can correct me, but our committee recommended that NQF change its practice and allow SDS risk adjustment where appropriate. And, the board were the ones who said, time out, we are not comfortable. We are worried about masking disparities. We are willing to do this as a trial period.

And so, is that the main thing that we need to investigate, is whether there have been unintended consequences from disparities, is that going to satisfy the board. Would they then be willing to say we are going to go forward with this.

DR. BURSTIN: It's a great question.

So, first of all, the board has vested in this committee that evaluation, and they will take a recommendation from this committee. So, that's why we think it's so important we work it through with all of you to determine what the right evaluation is.

I think we are unlikely to see positive and/or negative effects of adjusting measures in the real world between now and two years. It's just so highly unlikely that measures will get into use and we'll see their impact.

That's still a critical part of this work, regardless of, you know, whether we need that two-year window or not. So, part of it is, and I like what Sarah just laid out for us, there are clearly some things we can do by being incredibly transparent, cataloguing what we are learning, kind of posing back to this committee the key questions that come out, like what do you do about community factors that Dave teed up before he left.

But, you know, we just need to go back to the board with a sense of based on what we've seen so far is there anything to suggest that just allowing them to come forward and including the stratification is something we should retract. They did not give us more guidance than that.

I do think that's one of the key questions going forward as well, is we made this

recommendation, but we said in our adjusting you've got to also have stratification for that transparency, and I think that's something we want to go back to this group about as well. Is that something that persists, or is it something that we'll get more comfortable, we simply allow adjustments.

CO-CHAIR PONCE: Lisa, and then -- I'm sorry, and then Philip.

MEMBER IEZZONI: Thank you, Nancy, for reminding us of what the board said, because I think maybe it was Kevin, in your presentation, maybe it was you, or maybe it was David who said, oh, the statisticians tells us this masking isn't going to really be an issue, because that's not really a big problem.

There are, however, mathematical exercises that you can go through that would show you that you couldn't de-mask, you could, actually, cover up the fact that one group is not performing as well as it should be.

And so, when you said that I wondered

whether the jury is still out on that, or whether the committee process that you guys went through, you know, which Nancy was on, whether the masking issue, where the masking issue really stands, because I think that if one of the things that our committee needs to deal with is the masking issues, it gets to the data that Karen is going to have to give us. It makes it very, very different, the kind of information that we are going to want to have.

MEMBER FISCELLA: What the statisticians weighed in was that with appropriate statistical models, generally, there would not be, you know, depending on the distribution of the data, and how it looked. That's why there wasn't a blanket statement that, no, it would never happen.

MEMBER GARRETT: And I would also say, I think another remedy to the masking issue that we talked about a lot was that conceptual basis. So, if conceptually there is an institutional history of racism in the way healthcare is delivered for a particular measure, then you might choose not to have race be a risk adjustment variable, even if you

think that there may be some other reasons to do it with unmeasured factors.

So, it's that careful, really careful look at that conceptual basis that helps you decide what you can do with and what makes sense measure by measure, which again is why we didn't prescribe an approach, because you really have to be cycled back on the statistical issue, that's more of a conceptual one. And so, that's one way to make sure you are addressing it.

CO-CHAIR PONCE: Philip.

MEMBER ALBERTI: I wonder if there's any value or opportunity, in addition to a deeper dive into the actual submissions themselves, is there any way to reach back out to the developers and understand their process, and how they arrived at what they did, and the quality?

DR. BURSTIN: We have two of the best people, that helps. But, no, actually, that might be an interesting webinar, is to have a couple of the developers who have been thinking about this and kind of working it through sort of like we heard what

Kristy was doing. I think it would kind of give a little bit more of a thought processes, the barriers that are data related, the barriers that are literature related, et cetera.

MEMBER ALBERTI: Because I'm pointing to a role that NQF could take in terms of, you know, if this is new to you here are some of the kinds of variables for which data are available.

CO-CHAIR PONCE: Sarah.

MEMBER SCHOLLE: Likewise, this is where the work from the IOM could also be something that would be useful for developers, I think, in saying, here, take a look at this, this is what we mean. And, here's the range, and here's some recommended approaches, just kind of best practices. That would be nice.

DR. BURSTIN: We usually look at our old applications, as we've doing in work with some of the others that submitted pediatric things, here's what we -- here's our experience of what the committee is looking for in that thing.

And so, there may be a way to, actually,

provide more direction, or support, to people who are looking at it and saying, really?

DR. BURSTIN: We could do a monthly measure of a webinar with all of us, maybe 70-80 different developers on the call. So, that would be great to, actually, present some of the IOM work to them.

CO-CHAIR CHIN: I wonder if we partly have a problem of like you join where the light shines, you know, like this is going to be like the vast majority of all of us over the last couple years, people using the readily available SDS variables claims data which is going to be limited, and what we see so far is that, is really data that may not have much of an effect.

You know, there's the area of like what I would call, say, cutting edge research, where you have something like the Avalere project which shows an important detailed data set that you can do more.

So, whether or not it gives us a combination of reviewing what's been done, and just getting some more details with a black box, what

happened with those, the 24 percent we have to worry about through, so that you've got it there.

But, the literature review, I guess, or the current status review, but I wonder what, and I'd like to revisit this, but, you know at the end of the day, I mean, I can imagine like if this is like Marshall said, in two years it's going to be, well, you know, we really don't know the answer. It's on existing, available crude claims data, you know, marginal, whether, you know, it really is helpful or not. But, conceptually, we still have to wash the well, and if we really have the barrels to capture the sub source to get it, you may be at least able to show that at the edge, yes, it's important.

DR. BURSTIN: It's, just real quick here, this might be for Marshall, it's a great comment. I mean, at the end of the day, we are not, necessarily, going to be able to say in two years this was really effective. I think, essentially, the question is, are we going to be able to say, boy, there's a lot more to learn here, and we should keep

being in this vain and trying to learn, or will there be anything, I don't know, that the policy context changes a lot, tomorrow they can start paying safety net providers. There are many things that could happen that may change the perspective or evidence may emerge about different approaches, as Sarah noted. I don't know.

But, you know, I think as we try to build, work backwards to what we need to do, we can easily pull that information forward, we'll get a good robust table for you guys as to what happened to each of those measures.

But, you know, what would make this -- well, what information would this committee need to say, you know, hard stop, this is just not the right thing to continue to let this go forward.

CO-CHAIR CHIN: So, this is a binary question.

DR. BURSTIN: Yes.

CO-CHAIR CHIN: This is an important one. And, there's also Karen's continuous variable question of like, should we, yes or no, or, you know,

how be proactive of saying, well, you know, we want to influence the deal by encouraging or somehow, you know, adding more detailed SDS to the question. The rest of the questions that Karen raised that we haven't really addressed yet, so it will be the binary question, and the answer is probably going to be what Bob said, that, yes, at the end of the day we're probably going to say, well, we don't know enough, and that's fine from the trial period, or for permanently.

But, we still haven't addressed the fundamental question of, well then, what is the right way to do it.

DR. BURSTIN: Right, but I feel like in some ways the answers to those other questions are given, that's within the context of the work of this committee. We absolutely want you to help provide the information to help do this better.

But, in terms of the formal evaluation, the binary question is one of the ones we want to make sure we'll have enough information for you at some level, surely not including unintended

consequences or positive consequences, what would be what you would find helpful. And again, maybe it's something you need to kind of think about and provide to the table the details.

But, we just want to make sure that we are building into the process when lowered prospectively the right kind of information to help you make that decision with us.

CO-CHAIR PONCE: Kevin, Traci, Eduardo, and then Nancy. Oh, and sorry, and Lisa.

Kevin.

MEMBER FISCELLA: Yes. It's hard to imagine that in the next two years that things are going to advance so quickly that we are going to have, you know, that much better data. I mean, it would be great if we had the nine-digit zip codes in place. I don't know what it takes to get that data, what it entails, but it's hard to imagine that would happen within this time, and that we are going to have all of that.

So, within two years it's hard for me to imagine that we would have, you know, persuasive

evidence at this point that we don't need to continue. The committee is going to continue, and there will be continued monitoring. And, as new things emerge, there could always be a recommendation, look, we need to stop this. For X, Y and Z we are saying this.

You know, I do think in the interim it certainly makes sense to keep track of the extra time burden, and that needs to be part of the equation. On the other hand, we did have a clamor from the provider community that this is what they want, and it certainly increases, you know, credibility at this point.

So, it's -- I think it's very unlikely that there would be data points that would come up that would say, wow, we need to go back to -- or a list -- I don't know what that would be.

CO-CHAIR PONCE: Okay. Traci.

MEMBER FERGUSON: I think putting it in the framework of focusing on what can we do to make sure that we have a defined process of how we can identify where the disparities exist, and how we can

develop -- you know, have a conceptual and empirical framework to bring forth in looking at sort of a trial.

So, you know, it's our output over these two years, or what made some risk adjusted measures work, and what made some that didn't, and that we can give sort of a tool kit to developers and providers that this is the type of data, so that we can lay that sort of foundation for this process. And then, eventually, we'll get to where we can measure outcome.

But, we know it's a good thing in terms of best practices. This is what you are supposed to be doing, and so if you are saying this is what I'm supposed to do, I'm going to do it. And, eventually, I'll see the outcome.

So, maybe thinking of taking a step back and saying we are going to develop the best practice in terms of looking at disparities, how we can identify, how we can measure, how we can collect the data, and then continue on and eventually we will get to the outcome.

CO-CHAIR PONCE: Thank you.

Eduardo.

DR. SANCHEZ: I think I'm thinking exactly along the same lines. I figure that in two years we probably should ask ourselves, do disparities still exist or persist. I suspect the answer will be yes.

And, in an iterative QI approach, we have to ask ourselves why, just like we are asking now, how do we better understand the why, how do we reduce and eliminate. Hopefully, in two years we learned the how part, despite the fact that some of those will continue to persist, and then ask ourselves, have we done the structure process elements that need to be in place to achieve the desired outcome, I think repeating very much what was just said, starting with the conceptual empirical framework that allows us to iteratively move to better than where we were six months ago from a measurement perspective and the ability to inform what might change those disparities.

CO-CHAIR PONCE: Thank you.

Nancy, and then Lisa on the phone after.

MEMBER GARRETT: So, I agree with Kevin, it's hard for to imagine the data that we'd have in two years that would tell us, oh, this is a really bad idea, we need to go back to the old ways.

But, one idea that I have for something we might be able to look for as a positive result of this is influence. It's a hard thing to measure, but I see, the SDS report cited all the time. And, I think it's inspiring a lot of other efforts. I know that the legislation I shared in Minnesota, you know, we were sharing the results of the work of that committee with the legislators and its influence, you know, to see this is what's going on on a national level.

And, that's going to cause us to start collecting sociodemographic data statewide. I mean that's influence that's going to make -- you know, have the ability to make a positive change eventually.

So, you know, if there's some way for us to kind of survey the environment and understand

what the impact has been in a positive way, short of those goals over reducing disparities but steps along the way I think that is something we should also look at.

CO-CHAIR PONCE: Thank you.

Lisa.

MEMBER COOPER: I think again I waited so long that pretty much what I wanted to say has already been said. So, I won't say anything else.

Thanks.

CO-CHAIR CHIN: I was, actually, a little confused with the past three or four comments, when people are talking about, well, disparities exist, figure out the root causes then marching ahead.

Are we talking generally, or were we talking specifically about the risk stratification and the sociodemographic. Maybe like if Traci and Eduardo you can clarify what you meant.

DR. SANCHEZ: Yes. Relevant to the conversation we are having, yes.

MEMBER FERGUSON: Yes, these particular measures, so the measures that are not, I guess,

creating the sort of new measures that haven't been thought of or brought before NQF, I'm talking specifically, looking at what we can do to identify those measures that have gone through the process, through the trial period, and those that are NQF endorsed.

CO-CHAIR CHIN: And, to both Eduardo and Traci, I still don't understand. If you could be a little more specific.

DR. SANCHEZ: Sure.

As it relates to the SDS factors and even risk adjustment, it's looking where we are in two years, being able to look backwards. I suspect that the rationale for what we will have done over the prior two years will still exist. We will have made progress, but there will still be opportunity, so it's about refining the work that we are embarking upon now.

And then, I mean, that's the way I'm reading that question about the recommendation. Is that a recommendation tomorrow, or is that a recommendation down the road, the first bullet.

CO-CHAIR CHIN: I was thinking two years from now.

DR. SANCHEZ: Yes, so I figure two years from now, and two years from now, maybe two years from now we say, woo hoo, we don't need to be doing this at all. There's no reason to do this, either because we've analyzed it enough that we don't believe there's a difference, or, I don't know, nirvana happens.

But, not to be flip, I think that in two years what we want to do is make a recommendation about what are the steps forward as opposed to up or down, based on what we've gleaned over the last couple of years, both from an experience perspective, and from a what's happened in the environment around this perspective as well, because there will be some change that happens, hopefully, positive change.

MEMBER FERGUSON: And, I think for me, so for those measure developers who had a very strong or convincing conceptual framework that there was some disparity difference, that in two years we

would have a very defined process of making sure that even if it doesn't have at that time empirical evidence to support it that they know exactly what we are going to have, these data measures, we are going to continue to look at more information. And then, in a year or two years maybe come back and by that time have the conceptual and the empirical evidence, but give them sort of a means to continue on in that effort, so you don't just drop the conceptual and just say, okay, it's over.

CO-CHAIR CHIN: Got it. Thanks very much.

CO-CHAIR PONCE: Yolanda.

DR. OGBULU: I just wanted to clarify the two year point. I think in the beginning of your presentation you said April, and so it's really like a year and two months. Is that correct?

CO-CHAIR PONCE: Two years after the report.

MS. JOHNSON: Yes, it really is less than two years today.

DR. OGBULU: Okay.

MS. JOHNSON: Yes.

DR. OGBULU: Okay, so I would agree with everything that I have heard.

CO-CHAIR PONCE: With the IOM report, or the National Academy Report was referenced a lot, Jose, if you are still on, would you remind us when the recommendations will be made?

MS. JOHNSON: It's October this year.

CO-CHAIR PONCE: Okay. There's five reports, there are five --

(Off microphone comment.)

CO-CHAIR PONCE: Oh, so all five, so the 5th one which synthesizes all the four previous reports, everything will be due in October.

(Off microphone comment.)

CO-CHAIR PONCE: They are going to bring them out as they are ready. Yes. Okay.

Susannah.

And, sorry, Nancy, is yours still up?

MEMBER BERNHEIM: I just want to come back Marshall, I was a little bit confused the linking of sort of if disparities still exist then

we would still want to risk adjust, right? That's lumping a lot, and I want to make sure I understand.

So, today if I told you these measures, not my own, that there are disparities in rate of catheter-associated infections, the committee felt largely that that's not a case where we would run a risk.

So, we said, right, that there's not a conceptual basis. So, the distance of disparities as the rationale for risk adjustment, I mean, I don't think you really meant that, I think that's what everyone was saying, but I think that you want to be careful about that, right?

So, that's my thought.

CO-CHAIR PONCE: Kevin.

MEMBER FISCELLA: Just to echo Nancy's comment about influence, I think that should be part of what we are doing here. And, I think that the committee's recommendation certainly helped to put discussion around SDS factors out on the forefront and get people talking about it.

And, perhaps, we'll accelerate actions

towards collecting these data, which have, certainly, as we all know, has uses way beyond just risk adjusting quality measures. There are lots of other clinically relevant, quality improvement, disparity reduction, on and on, ways of designing services somewhere that incredibly important.

So, I think that that should be part of the equation. You know, I also think that it's unlikely that over the next few years that any of this adjustment is going to make that much difference in terms of payment, and that that needs to be a broader issue of how we address and discuss that. You know, just given what we've seen, and I -- you know, there may be measures that will show bigger effects like we saw today as they come in. But, by and large, I think the effects are going to -- are going to be relatively modest, and so I think we need to be thinking about other ways beyond thinking that this is really going to be the answer. It's, certainly, not going to be.

MS. JOHNSON: I guess the only other thing, and we don't have to go into it today, but

if you continue to think about guidance that you think you as a committee may want to put out to developers who are going through this process with us.

What we have heard from the field is anything that you've learned and can help us with we'd appreciate knowing. So, if you have things like that, we would like to know those as well.

CO-CHAIR PONCE: Great.

Next?

DR. BURSTIN: Just to thank everybody, that was really helpful. We will take that, and there are some great comments there. I think we have a much better sense of how to structure this.

This next section is really about how this committee will interact with some of our NQF Measure Endorsement and Selection.

I know many of you are having to leave in the next half hour or sooner, so I just want to say thank you for those of you who leave early and catch flights home.

And, this is an issue, we probably won't

get through all of it today, but I just want to at least tee it up so you can be thinking about it, and I'll turn to Elisa.

MS. MUNTHALI: Thanks, Helen. And, I wanted to first apologize to everyone. We were thinking about the sequence of slides, and where we should put all the different topic areas as we were trying to come up with the agenda. And, as we have been discussing the bigger picture issues yesterday, and the more concrete specifics today, we realize we probably should have started off with a discussion around our process, and how we have incorporated disparities, and not just in measure endorsement, but also in measure selection.

So, much of what I will go over, I probably will not go over, because we tried to pull out, in the interest of time, knowing folks needed to leave, pull out information about the measure evaluation criteria related to endorsement. That's what Karen went over, so that those people could hear it before they left.

So, next slide.

This is something you've seen before, but it's a slide that Helen went over yesterday. And, I wanted to just reiterate that we are doing quite a bit of work in many different areas, and throughout all of our work the number one thing that you will notice is that we are bringing multi stakeholders together, whether it's through recommending measures for endorsement, the 600 plus measures that we have in our portfolio of measures. They are not just clinical, the majority are clinical, but we do have measures in crosscutting areas, like person and family centered care and population help.

But, we are also recommending measures to Health and Human Services about CMS for inclusion in Federal programs, about 20 of those Federal programs, and also doing quite a bit of work with stakeholders who are particularly interested in safety and other key areas, to help move, you know, measurement in a way, in an advocacy way. And, I'm doing a lot of work around measurement times that Helen mentioned yesterday, a project that I know

many of you are very interested in, it's one we just picked up around attribution, and that will be a theme we'll be talking about throughout the next few years.

Next slide.

So, I just wanted to highlight a couple of things here, our evaluation criteria, which includes the five major criteria that Karen Johnson went over, are standardized.

So, when we talk about the evaluation criteria being standardized, we mean that it's transparent, it's open, developers know what's expected of them, but also the standing committees that are reviewing the evaluation, the measure submissions, know what to look for.

And, as Karen has mentioned before, it is evolving, our criteria. We are listening to stakeholders, you know, as the enterprise, measurement enterprise, evolves. We are trying to keep it as steady as possible, but we need to be responsive, especially, if science is changing.

Next slide.

So, we've laid out again here the evaluation criteria. I'm not going to go over this, just to reiterate that this is hierarchal, you know, the first two, importance to measure and report, that include performance staff, don't really want to measure -- we don't want to assess anything where there isn't an opportunity to improve.

So, you know, we look for that. We also want to make sure that there is an evidence base to the measures that go through NQF endorsement. So, importance to measure and report is must pass. If measures fail on this criterion, we do not look at them beyond that.

Reliability and validity testing is also a must pass, and also Karen mentioned the feasibility and usability in use, and how those are important to how we assess measures into our criteria. And, the importance of reducing the burden and making sure that we are, indeed, endorsing the best in class measures.

So, you can advance to the next slide,

please.

I also wanted to talk about our MAP work. This is Measures Applications Partnership. We are in what we call MAP season, this is a very intense period for us. We, as Helen has mentioned several times during this meeting, we'll be having the coordinating committee meeting next week in this room. It will be pretty well attended, but the coordinating committee will be looking at recommendations that are coming from three major work groups that are part of the MAP, one on post-acute care and long-term care, one that is looking at clinician-level measures, and another that's looking at hospital-based measures.

The MAPs recommendation, we have -- that's called the pre-rulemaking recommendation, as I mentioned before, there are about 20 Federal programs in that.

So, you can advance to the next slide, please. Oh, yes.

In addition to, and Helen just whispered to me, in addition to the three work groups and the

coordinating committee, we do have two work groups on child and adult Medicaid, and also a work group on dual eligible beneficiaries. They give significant input to both -- to all three of those work groups and also to the coordinating committee.

We have included here the measure selection criteria. And, how the MAP process works is, by December 1st we publish a list of measures that are under consideration for the different Federal programs. It's a pretty intensive period, we don't get that list well advance, often times just a few days before Thanksgiving, or on Thanksgiving, and our staff works very hard to make sure we put that list up.

We have gone through a number of improvement activities in which we have implemented an early commenting period. And so, before these measures are sent out to the various work groups I mentioned before, we give our members and public an opportunity to comment on the Federal Government's measures under consideration.

And, while there's the comment period is

open, our team is also looking at the measures to make recommendations to help the work group leads, and the different experts on the work group, make recommendations for what to include in the annual measures on their consideration list.

And, if you'd go to the next slide.

And so, in terms of disparities and cultural competency, it is hard wired into the MAP process, perhaps, not to the extent that it is in our endorsement process. But, what we wanted to show you here is the language and criterion we use for inclusion of disparities and cultural competency. I'm not going to go through it, because there's quite a bit on the screen, but we did want you to see some illustrative examples of how we are trying to advance on the elimination of health care, especially, disparities and health now more so, and cultural competencies in our major core work.

So, next slide.

And so, I don't think we are going to have an opportunity to dive into this question. I

did show you very briefly how we've incorporate disparities and cultural competency into our work as I mentioned before. But, a couple -- one major question we wanted you to go away with is to think about, not just about the NQF work that we talked about today, but think about the larger issues we talked about around data collection, the availability of data, and our ability to really push and have a stick for us to get in the measures that we want. But, how can we increase our focus on disparities and eliminating those disparities, not just in health care, but also in health, as we talk more broadly.

So, I don't know if we want to open it up today, but --

CO-CHAIR CHIN: I'd just be curious. What your impression is Elisa, as, you know, the top dog in terms of the quality measurement, but what -- what currently works in NQF, what do you see as the strengths and weaknesses, and where do you think that, you know, you personally think that would be the next steps of NQF getting more involved in

disparities?

MS. MUNTHALI: That's a great question.

I think one of the things I think that we have tried to do over the last couple of years is align our work. So, for many of you who don't know our processes so well, or are not as familiar with NQF, the two major processes around endorsement and selection were very siloed, we have integrated the two departments, they were separate departments, actually, dealing with this very important work. And so, we integrated those two departments. And so, what has helped now is that our mission and our vision around elimination of disparities is a lot more aligned. While the two processes are very separate, that is a primary goal around endorsement and selection.

I think also we have a very good relationship with developers, I think by and large. And, the questions that we ask for performance staff around disparities, and the information that we are trying to glean from then on SDS with validity, we have very open and honest conversations about the

limitations that they have in getting that to us.

But, I think that data, in terms of what we get from them, has improved, and, Helen, I welcome your thoughts on that as well.

DR. BURSTIN: I think, ditto, exactly what Elisa said. I do think of ways you could help us think through how the -- how we know, for example, which measures coming forward are especially important for disparities, kind of give us more guidance. We may want to read this at the disparities sensitivity criteria we've done before, just it didn't really work thinking about it prospectively, and so guidance from you as to really hone in on the measures that we want to make sure get looked at really closely for disparities I think would be helpful.

And also, the MAP process is incredibly influential, 20 different Federal programs. Marshall is, actually, the disparities subject matter expert on the coordinating committee, he'll be back, hopefully, next Tuesday, assuming we are not completely dug in for weeks here in Washington.

You know, are those criteria enough? Is there something else we could do to build into our recommendations on the Federal programs as these measures come forward. They do have latitude to talk about the programs themselves. CMS puts forward the program goals, as Karen certainly knows well, but, you know, are there opportunities to think about, wow, does this really have to only be in a team program, could this be something where there might be payment for the trajectory or percent improvement to help consider issues of disparities, always stratifying paying in that way.

So, you know, where are sort of the leverage points around, particularly, as measures come forward, and we evaluate them, and also the work, particularly, from MAP, which I think is so really potentially very high leverage.

CO-CHAIR CHIN: My impression, and, of course, you guys are steeped in it on a day-to-day basis, but my impression has been it's been sort of below the surface, the equity issues, I mean, sort of indirect. The discussions we've been having

over the past couple days are just very explicit and direct. And, I think the degree, for example, that many of the things we talked over the past couple days are explicitly discussed within all these committees.

I mean, these are great discussions.

DR. BURSTIN: That's why we will -- you've now seen the measures submission form, you can see the questions we routinely ask about disparities. We want to know are there differences across populations. We often find most of the developers, many of the developers can't give us that information back.

So, there's also a little bit of a chicken and egg that we can't really insert more into our process to make it more explicit, if we can't get the information from the developers. So, you know, we are all kind of commensal organisms in this, we've got to think about how we kind of best feed off, really, the whole measure, you know, identifying the gaps, prioritizing the gaps, developing the measures, bringing them in to us,

putting the measures in use. I mean, it's really a cycle, and we've got to think about where we can help insert disparities reduction to every part of that cycle.

CO-CHAIR PONCE: I'm just going to check in with Lisa, in case she wants to say something now.

MEMBER COOPER: Thank you. I don't, actually, have anything to add right now.

CO-CHAIR PONCE: Okay. Traci.

MEMBER FERGUSON: What's the possibility, I know that, you know, in terms of funding, in terms of with our disparities standing committee, but what opportunities for outside funding to really get into creating a way that we could give the developers, to get the information that we need to make a decision. And, I mean, I just don't know. I mean, is it possible that we could help assist with grant writing, I don't know, or something, just to get additional funding.

DR. BURSTIN: It's a great question. Traci, and there are two different issues here. One is we would, obviously, love to find support for

this work going forward. And, we've had some early discussions with the government, and we'd love to try to see if that's possible.

If you have suggestions of foundations or groups you think we should approach, we'd be all ears to see if we can get this funded. I mean, truly, the idea of saying we'll have another meeting next year is just -- we couldn't just throw this on here without a budget and without any resources to do it.

But, if we could get the resources, it's not bandwidth for us internally, it's really just trying to make sure we have it. So, your suggestions of who to approach would be very welcome.

On the developer side, you know, I'd love to, you know, certainly, Sarah may speak to this, but there are issues, much of their work as well is dependent on funding. So, you know, we can say, hey, you've got to have information stratified by race or ethnicity so we can see if there are differences, but if they can't gather the data they

can't, if they don't have the resources to do it.

CO-CHAIR PONCE: I'm going to go to Romana, because I think it's on this point.

MEMBER HASNIAN-WYNIA: Yes, it is.

So, PCORI has funded convening boards, I can't call them grants, because they are always contracts. But, we have funded convening awards.

So, what would help is if I got an email indicating what you would do, why it's important, how it would advance the disparities agenda. That's what I need.

I don't want to say too much, because of potential conflict given that I'm on this committee, but I think that's fine to send that, and then I can pass that along, and, hopefully, get a dialogue going.

CO-CHAIR PONCE: Okay. Philip.

MEMBER ALBERTI: I just have a question, and I'm not suggesting that this committee take this on is strategically possible, but we've heard that there are 600 plus endorsed measures already. So, has the NQF or anybody begun to just take a look at

what you've already endorsed to do, you know, quick lit reviews for the last ten years to say, here are the 150 for which there is an evidence base that these kind of disparities exist for these specific metrics. And then, could that offer developers guidance going forward when those measures come up for renewal, that there's an expectation, perhaps, that those are really the targets for developing a conceptual framework and a model and testing them.

DR. BURSTIN: That's certainly something we can consider. And again, like you, we have standing committees across most of the system, you know, cardiovascular, pulmonary, et cetera. So, one thought might be we could ask them as well to help guide that work.

It isn't often so much at the measure level, it's more, you know, the area level, it's asthma measures of course, it's HIV measures, of course, pregnancy measures.

But, it's an interesting idea, we'll follow up and see if we have any thoughts.

CO-CHAIR PONCE: Sarah.

MEMBER SCHOLLE: So, I wanted to kind of follow up on that idea, but how do you figure out which measures should be considered or evaluated for this. And, the steering committees have a lot of latitude in this respect, and it, actually, kind of feeds on the idea 600 measures, which ones would be high priority.

So, I think it would be helpful to think through with the committee, with each committee, what is their role in doing this. We often find that committees have different --

CO-CHAIR PONCE: Personalities?

MEMBER HASNIAN-WYNIA: -- yes.

So, and it's -- you know, some areas are going to be much harder. It's just going to be really a lot harder to try to address these issues than others.

So, I do think a discussion with the committee and, actually, with the measure developers who already have kind of different groups. So, as you think about what are best practices, it might be helpful to think about

bringing your standing committee people along with the developers into one conversation about what's possible, what's expected -- or, what would we really like to see versus what could we really get, and how much money would it cost to do what you really want us to do.

CO-CHAIR PONCE: Thank you.

Eduardo.

DR. SANCHEZ: So then to add to that, that sounds very, very wise, but I wonder, so sometimes those of us who are in the disparities/diversity world engage in conversations that say it's not enough for people to be thoughtful about things, sometimes you need folks with the perspectives at the table.

So, could we think about a strategy, in addition to what Sarah just said, to begin embedding, if you will, folks in these other standing committees who come in with a perspective that brings the notion of diversity and disparities to the table, and that maybe on a go forward basis one way to begin embedding this is to have as one

of the criteria for consideration something that speaks to health disparities and sociodemographic factors as not just what we are looking for in terms of what is the person's phenotype, but also in terms of how they are thinking as they come on board, the other standing committees.

DR. BURSTIN: That's a great suggestion, Eduardo, and we, actually, do, as the person who kind of does the final comment and decision-making on lots of those, we always look to see if we can find someone.

The issue is, you know, how do we find the right people. So, we also may need some help thinking about who we need an outreach to, to say how could I make you interested in cardiovascular measures and want you to come on this panel.

But, that was an easy one, by the way, but there are others.

CO-CHAIR PONCE: Cara.

MS. JAMES: Sorry. I just wanted to, actually, build on the point that Helen just raised, because it is also a challenge for us. And, I think

that one that the committee could also think about is getting the right people to give the feedback.

And so, for example, with the NPRMs, where we are soliciting comments, that's, you know, making sure we do hear back about what we are doing. And so, that's something as you guys are thinking about, making sure you are sharing throughout your network and being mindful that helping to identify potential experts in this that I think would be incredibly helpful.

CO-CHAIR PONCE: Okay. All these are concrete recommendations.

Speaking of building, shall we -- let's see, I think we are done, except right now we need public comments.

OPERATOR: To make a comment please press * then a number 1.

There are no comments from the phone lines.

CO-CHAIR PONCE: Thank you, and now Michael is going to take us through next steps and timelines.

MR. PHEULPIN: Yes, sure.

So, I think -- so as we heard --

CO-CHAIR PONCE: Sorry, but before you do that, there was a -- there was some side conversations in a volunteering spirit involved in this corner of the room, which I encourage to diffuse to other corners of the room.

But, Eduardo, Susannah, and Sharon have talked about doing the logic model, perhaps, the conceptual model. And, I think if you don't mind I want to also invite anyone else who would like to join them.

If, you know, spoken word isn't your thing at this point, you can write to us if you want to be part of that.

CO-CHAIR CHIN: And, if it turns out like -- again, this is still being formulated what the actual plan moving ahead is. But, just any of these topics that you are particularly excited about, and you are gung-ho, like Susannah sort of mobilized the forces in the corner to work on the conceptual model, and Eduardo, and to Sarah, all sitting there.

MEMBER COOPER: So, this is Lisa, I'm going to put myself out there and say that I'll help with the conceptual -- the logic model, I'm sorry, conceptual model.

CO-CHAIR PONCE: Wonderful, awesome, thank you.

CO-CHAIR CHIN: Thank you, Lisa.

So, if there are other topics that you think of, you know, on the ride home, and you know I would really love to work on this with a subgroup with a couple people, let us know. You know, the more the merrier.

Michael?

MR. PHEULPIN: Yes, okay.

So, with that as an immediate next step, we will just kind of look at the upcoming quarterly conference calls, so they are April 26th, July 21st, October 19th. And, they should be on your calendar, but if not we'll send another invite through.

And, you know how to contact us, and you know, of course, we'll use Share Point as a tool to

share documents.

CO-CHAIR CHIN: So, if I could say a couple things.

So, I asked Helen, I think we have the record at NQF, that was three hours, four hours and I think 45 minutes. It was three hours. Yes, and people were engaged, which means two things.

I mean, one is, the topic, you know, disparities is the topic we are all engaged in, but still it's the topic plus the people, the people, you can tell everyone here is mission driven. So, it really has been great, I'm really privileged to be on this committee working with you.

So, thanks so much for all your great inputs over the past couple days, it's a great start. I think we have a long climb, but I think we've got a good foundation here. And so, I'm looking forward to it.

CO-CHAIR PONCE: Thanks so much.

I also want to shout out to Mara who came back, but as we're adjourning it shows again this commitment.

I don't know if you heard it, but there's a group that's going to look into doing the logic model conceptual framework, and if you would like to join that, that group currently involves Eduardo, Susannah, Sarah, Kevin, Lisa Cooper, and if you would like to join.

So, thank you. This was really great.

Oh, Yolanda, see the longer we stay the more --

CO-CHAIR CHIN: Just one more shout for the staff, because as you all the people who run the committee, you know, are the staff here at NQF.

CO-CHAIR PONCE: Thank you. Thanks so much.

(Whereupon, the above-entitled matter was concluded at 2:29 p.m.)