NATIONAL QUALITY FORUM

\+ + + + +

MEASURE APPLICATIONS PARTNERSHIP
DUAL ELIGIBLE BENEFICIARIES WORKGROUP

\+ + + + +

THURSDAY
MARCH 30, 2017

\+ + + + +

The Workgroup met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 9:00 a.m., Jennie Chin Hansen and Michael Monson, Co-Chairs, presiding.

PRESENT:

JENNIE CHIN HANSEN, RN, MS, FAAN, Workgroup Co-Chair, Subject Matter Expert
MICHAEL MONSON, MPP, Substitute Workgroup Co-Chair, Centene Corporation
CHRISTINE AGUIAR LYNCH, MPH, Association for Community Affiliated Health Plans
JOE BAKER, JD, Medicare Rights Center
RICHARD BRINGEWATT, SNP Alliance
GWENDOLEN BUHR, MD, MHS, Med, CMD, American Medical Directors Association
ALISON CUELLAR, PhD, Subject Matter Expert
WENDY FOX-GRAGE, MS, MPA AARP Public Policy Institute
JOY HAMMEL, PhD, OTR/L, FAOTA, American Occupational Therapy Association
ALINE HOLMES, DNP, MSN, RN, New Jersey Hospital Association
K. CHARLIE LAKIN, PhD, Subject Matter Expert
ALICE LIND, BSN, MPH, National Association of Medicaid Directors*
THOMAS H. LUTZOW, PhD, MBA, iCare
STACEY LYTLE, MPH, CMS Medicare-Medicaid Coordination Office
D.E.B. POTTER, MS, Office of the Assistant Secretary for Planning and Evaluation

JENNIFER RAMONA, Homewatch CareGivers

KIMBERLY RASK, MD, PhD, Subject Matter Expert

E. CLARKE ROSS, DPA, Consortium for Citizens with Disabilities

JOAN LEVY ZLOTNIK, PhD, ACSW, National Association of Social Workers

NQF STAFF:

KATE BUCHANAN, MPH, Project Manager
MADISON JUNG, Project Analyst
DEBJANI MUKHERJEE, MPH, Senior Director
ELISA MUNTHALI, MPH, Vice President, Quality Measurement
ERIN O'ROURKE, Senior Director
RACHEL ROILAND, MS, PhD, Senior Project Manager
MARCIA WILSON, PhD, MBA, Senior Vice President, Quality Measurement

ALSO PRESENT:

BRIAN ABERY, MS, PhD, Coordinator of School-Age Services, Institute on Community Integration (ICI) & Adjunct Faculty, Institute on Child Development and School Psychology Programs, University of Minnesota*
ELIZABETH FRENTZEL, MPH, Principal Research Scientist, American Institutes for Research*
BETH JACKSON, PhD, Director, Truven Health Analytics*
KAREN JOYNT, MD, MPH, Senior Advisor to the Deputy Assistant Secretary for Planning and Evaluation, Office of the Assistant Secretary for Planning and Evaluation*
KERRY LIDA, PhD, Health Insurance Specialist, Center for Medicare and Medicaid Services, HHS*
CORETTA MALLERY, PhD, Principal Researcher, American Institutes for Research*
SUSAN RAETZMAN, MSPH, Senior Research Leader, Behavioral Health and Quality Research, Analytic Consulting and Research Services, Truven Health Analytics*
RENATA TICHA, PhD, Research Associate; Director of Global Resource Center for Inclusive Education, Institute on Community Integration, University of Minnesota*

ALSO PRESENT:

ROBIN YABROFF, MBA, PhD, Analyst, Office of Health Policy, Office of the Assistant Secretary for Planning and Evaluation*
RACHAEL ZUCKERMAN, PhD, Office of the Assistant Secretary for Planning and Evaluation*

* present by teleconference

C O N T E N T S

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

(9:06 a.m.)

CO-CHAIR MONSON: Thank you, everybody. So, we are going to get going. I think the goal is to try to start early so that we will start on -- well, actually we're starting late. The goal is to end early so people can catch flights and things of that nature.

You know, as I said yesterday, we had a really productive day yesterday. And again appreciate everybody's continued engagement as we come down to this being our last formal meeting, although I'm sure that there will be some new incarnation of this group. That was to put pressure on Marcia to make sure that happens.

So, but we've got actually a lot of presentations today. I think these are things that we've all wanted to see and we've talked about. We're going to see Charlie's team from Minnesota. We're going to see the HCBS experience of care survey, the new CAHPS survey. We're going to hear about social determinants of health.

So, without further ado, I think we're going to hand it over to Kate.

MS. BUCHANAN: Hi, everyone. My name is Kate Buchanan. I am a project manager here with NQF. And I just wanted to review the changes that we made to the family of measures.

So, the workgroup voted to remove four measures: the 0043, pneumococcal vaccine status for older adults; 0682, the percent of residents or patients assessed and appropriately given the pneumococcal vaccine, short stay.

And the workgroup noted that they were concerned about the removal of these measures because of the importance of the vaccines, measuring the vaccines. But they discussed that there may be changing in the standards of what is appropriate care, so that they will be on the look out for measures, new measures dealing with pneumococcal vaccination.

The other two measures that the workgroup voted to remove are 0558, HBIPS-7, the

Post-Discharge Continuing Care Transmitted to the Next Level of Care Provider Upon Discharge; and 0557, HBIPS-6, Post-Discharge Continuing Care Plan Created.

And so the workgroup discussed that originally these measures had been designed specifically for a psychiatric population, but have since been expanded. Additionally, we discussed the fact that there are additional post-discharge measures within the family, so we decided not to replace any of them.

So, the workgroup voted to include four new measures into the family. And there were some overall comments in the workgroup on these measures. They noted while many of these measures are not tested specifically in a duals population, there was some concern over this. Additionally, there was some concern about the risk stratification of each measure. But the workgroup said, even though these measures are not perfectly suited for the population, they can improve the quality of care that dually eligible individuals receive and so they should be included in the family,.

So, with that in mind, the workgroup decided to include NQF 3086, Population Level HIV Viral Load Suppression. And the workgroup noted that this intermediate outcome measure addresses an important aspect of care for individuals with HIV, a significant portion of which are dually eligible.

The workgroup also included NQF 2858, Discharge to Community. And the workgroup said that this measure focus, keeping people in the community, aligns precisely with the workgroup's charge.

The other two measures that the workgroup voted to include were NQF 2775, Functional Change: Change in Motor Score for Skilled Nursing Facilities. And the workgroup members really emphasized the importance of functional measures and decided to include it in the family.

The last measure, NQF 2776, Functional

Change: Change in Motor Skill in the Long Term Acute Facilities. So, there was some discussion around the ability to measure the short-term rehabilitation in patients, but the workgroup discussed the importance of functional change in any care setting, including LTACs, and decided to include it in the family.

There were three new endorsed measures the workgroup decided not to include. And those are NQF 2614, CoreQ: Short Stay Discharge Measure; NQF 2615, CoreQ: Long-Stay Resident Measure; and NQF 2616, CoreQ: Long-Stay Family Measure.

And the workgroup wanted to really emphasize that PROs and family-reported outcomes are incredibly important to include in the family, and that there are few opportunities for individuals and their family members to provide feedback on the institutional care they receive. Although, with this in mind, there was incredible concern about the validity of the responses. They were concerned that the measures will be skewed towards a positive outcome for the facility.

And with regards to NQF 2615 and NQF 2616, workgroup members were not sure that the questionnaires were accessible to individuals with literacy, health literacy, and cognitive function impairments, but that they really wanted to note in the report the importance of PRO in family-reported outcome measures.

So, with that, I want to see if there are any questions or comments?

(No response.)

MS. BUCHANAN: Seeing none, I will turn it over to our first presentation. And I want to briefly introduce our presenters.

Karen Joynt is a cardiologist at Brigham and Women's Hospital. Rachael Zuckerman and Robin Yabroff are analysts at ASPE.

Dr. Joynt led the ASPE team that completed the first report to Congress on social risk factors and Medicare payment policy. And Dr. Yabroff is leading the team currently working on the second report.

And I just want to make sure that we are able to hear Karen, Rachael, and Robin?

DR. JOYNT: This is Karen. Can you hear me?

MS. BUCHANAN: Yes.

DR. YABROFF: This is Robin.

MS. BUCHANAN: Great.

DR. ZUCKERMAN: And this is Rachael.

MS. BUCHANAN: Wonderful. And if you wouldn't mind, just when it's time to move to the next slide, just say "next" and we'll move the slides for you.

And with that, I'll hand it off.

DR. JOYNT: Great. Thank you so much for having us this morning. We're really excited to get to talk with you. And from even the little bit of your meeting that we just got to sit in on, I think we're going to learn a lot from what you can tell us about emerging measures as we get through our presentation here. So we're going to try to keep the presentation pretty brief, and then we'll very much look forward to your feedback.

I'm going to go through a brief overview of our first report. And then Robin will take the baton and go through our plans for the second part of the study. And so, hopefully, we'll get a chance to update you on where we have been and where we're going, and get your feedback on where we should be headed. So we really appreciate your time.

Next slide, please.

Okay. So, as you all know, we worked on this report because of the IMPACT Act, the Improving Medicare Post-Acute Care Transformation Act. ASPE was charged with addressing the issues of social risk factors which, as you certainly all know, play a major role in health. And as higher levels of provider accountability move across nearly all Medicare settings, the issue of social risk and value-based payment really have started to intersect.

And so the IMPACT Act mandated

essentially four pieces of work. And we'll talk about almost all of them today.

So, first, a study of the impact of socioeconomic status, which we've reframed as social risk, on quality and resource use in Medicare using existing socioeconomic data, which is what I'll be talking about first.

Second, a study of the impact of socioeconomic status on quality and resource use in Medicare using measures from new data sources, like education, health literacy, et cetera. And that's what Robin's going to be talking about.

Third, a qualitative analysis of data sources and context around defining SES. And for that we had the National Academy of Medicine do a wonderful set of reports sort of doing a deeper dive into the background and context around a lot of these issues which we'd refer you to as well.

And then a final report with recommendations due in 2019.

Next slide, please. So, the way that the report is set up -- I don't know how many of you had a chance to look through it -- but the intent was to set it up such that once you sort of knew what one chapter was going to look like, the rest would follow a similar setup. So I'll talk through sort of a generic setup here.

We selected a set of social risk factors. The predominant one ended up being dual enrollment in Medicare and Medicaid, which is convenient for this committee. But it wasn't selected just randomly. It was selected because in our analyses it was really consistently the most powerful social risk factor that we examined. So many of these other factors: residents in low income areas, black race, Hispanic ethnicity, rural residents, and disability were important factors, but fairly consistently the dual enrollment dominated. So that's what we'll really focus on.

We looked across the Medicare payment programs that are currently in place that incorporate resource use and quality measures. And if your screen is as small as mine, I suspect those are very, very small boxes. But essentially there's three value-based payment programs in the hospital setting.

In the ambulatory setting we looked at Medicare Advantage, MSSP, and the value-based payment modifier which will sunset and turn into MIPS in a few years.

And then facility-based, so dialysis facilities, nursing facilities, and home health agencies. Those last two we really just did some exploratory analyses because those programs are still in the sort of getting-up-and-running stage.

Next slide, please.

So, we really had two main findings. There's lots and lots of little findings. But I think one thing that was impressed upon us when we put it all together was really that there were patterns across the different settings. And so we found both patient and provider effects. So I'll talk through each of those.

So, first, we found that beneficiaries

with social risk factors had worse outcomes of quality measures, regardless of the providers they saw. And dual enrollment status was the most powerful predictor of poor outcomes.

So, even when we ran models in which we were only comparing folks within the same practice or hospital or contract, there really was a significant and pretty consistent effect, especially of dual status, in terms of performance on quality measures across the board, not only normally process measures but outcome measures and resource use measures.

However, of course, the story is more complicated than that. And our second finding was that there was also a provider effect. So, providers that disproportionately served beneficiaries with social risk factors tended to have worse performance on quality measures, even after accounting for their beneficiary mix. And we can talk about why that might be the case, but it was essentially, instead of a simple finding, we ended up with a more complicated one, that we see both a beneficiary effect and a provider effect.

And the consequence of these effects is that under all five value-based purchasing programs in which penalties are currently assessed, these providers had somewhat higher penalties than the providers serving fewer beneficiaries with social risk factors.

Next slide, please.

So, here's an example. And the details here don't really matter, but I'm just going to try to walk you through the picture to sort of get a visual for the type of findings.

So this is looking at quality measures within the Medicare Advantage program, which are a range of measures from process to outcome. And you can see here all the red bars indicate where dual enrollees had lower odds of meeting the measure. So, for diabetes, for example, blood sugar control, up at the top there, they had 32 percent lower odds of meeting the measure. As you go down, those effects get smaller. So, very small differences, for example, in kidney disease, the other end of the red bars, at only 7 percent.

And as you get down to measuring physical activity and BMI assessment there was no difference. And for whatever reason, reducing risk of falling, that quality measure was much more often met in dually enrolled individuals. But, essentially, we saw a reasonably consistent pattern of a sort of a small to moderate dual effect across a wide range of measures.

Next slide, please. What that translated into was that when we had Medicare Advantage contracts with a high proportion of duals, they were much, much less likely to meet the four-star bonus threshold, which is once you get four stars you're eligible for a 5 percent bonus. So there's sort of a natural cut-off at four stars built into the program.

And you can see here that, on the left side of the screen, 72 percent of contracts with the lowest proportion of dually enrolled individuals were able to meet four stars, compared with only 26 percent of those in the highest quintile of dual enrollees.

Next slide, please. Now, of course, this group is probably one for which this slide will be almost self-evident, saying that there are differences doesn't answer the whys of why there are differences. And, certainly, one solution will not address all causes.

So I think we realized in doing this work that it actually isn't just about whether or not you adjust the measures, because so many things are feeding into these patterns that we see that we wanted to sort of take the opportunity and expand the discussion to think through what these factors are and how we might sort of start a conversation about the ways that measurement and quality programs actually influence these outcomes.

Certainly, on the bottom left, quality of care, you know, we did find that providers serving a high proportion of duals seem to provide lower quality.

We also know that duals have higher medical risk that's measured, which is incorporated into some quality measures and not into others. We know that they have higher medical risk that's unmeasured, things like functional status. And so I loved seeing the conversation right before we started about how to measure things like that.

We also know that duals have lower levels of social support and tend to live in more areas with higher levels of neighborhood deprivation, for example.

We know that there are issues with medication compliance and lifestyle, things like tobacco use among low income beneficiaries, that may contribute to outcome. And we certainly know that in life in general, many individuals with social risk factors face bias. And all of these things can feed into the ways that worse outcomes manifest in these groups.

Next slide, please. So we came up

with three sort of strategic goals in thinking about how Medicare and HHS might really start to think about how social risk interacts with the value-based payment programs. And these are all considerations for policymakers and clinical leaders and other folks to sort of get started on.

So, the first big strategy is to measure and report quality for beneficiaries with social risk factors. The first bullet point under that is essentially saying that we need to have adequate data collection and statistical techniques to allow us to measure and report performance for these groups about which we're particularly worried, so that we can see and monitor and track progress into reducing disparities.

The second bullet point talks about introducing health equity measures or domains into existing payment programs. And we don't know what those would look like, and we certainly look forward to your thoughts about that. But not only measuring performance but actually measuring equity, and publicizing those measurements through programs, would be one way to bring this more front and center.

And, third, prospectively monitoring the Medicare payment program, and as new payment programs roll out, understanding their impact on the providers that disproportionately serve beneficiaries with social risk factors feels like an important thing as well.

Next slide, please. The second big strategic bucket is to set high, fair quality standards for all beneficiaries, which is sort of a shorthand for saying that we don't have one recommendation about whether or not measures should be adjusted or not. Each measure is different. And so each measure should be examined to determine if adjustment for social risk factors is appropriate, which is exactly what the NQF has been undertaking during its trial period.

And really, I think, for us at HHS,

and also for folks who have been involved in that trial period, it's been very illustrative to see just how complex and varied these relationships are.

The second part here is also germane to the conversation you all were just having, which is that we would encourage the measure development community to continue to study measures to determine whether differences in health status may underlie some of these observed relationships. So, things like functional status, which are hard to measure, but we know are important. Things like disability or frailty. Ways that could better pick up some of the differences between patients with social risk factors and those without, that we might be able to pick up using, for example, claims data, functional status data, or patient-reported outcome measures.

Next slide, please.

And, finally, the third strategic heading here is to reward and support better outcomes for beneficiaries with social risk factors. And so the first bullet talks about creating targeted financial incentives within value-based purchasing programs to reward achievement of high quality and good outcomes or significant improvement among beneficiaries with social risk factors. And this would be a way to try to offset some of the perceived risk of caring for these folks under value-based payment programs.

Bullets two and three really talk about targeted support, technical assistance, and the potential for demonstrations or models that really focus on the providers that serve beneficiaries who are at risk, as well as the beneficiaries themselves. How can we really innovate in the ways that we know we should be in terms of things like integration of behavioral health, or care coordination, or linking with social services, or whatever people think are the best ways to innovate around this population. Are there ways that we could perhaps encourage that innovation to really start to change disparities?

And then, finally, additional research on the cost of providing -- or the cost of achieving good outcomes for beneficiaries with social risk factors, to think about what sort of resources it might take and how those could be deployed to try to address these issues.

So that's a very fast, high level overview of a big report. But hopefully it gives you a flavor for sort of the themes that we've tried to hit. And we can either pause there or we can move on to Study B and do questions at the end. Do you all have a preference?

CO-CHAIR MONSON: I think we have some questions now. So, if that's okay, we'll start with that.

MEMBER CUELLAR: Yeah, if you don't mind, could you give us an example of Number 1, your targeted financial incentive? How would you envision that?

DR. JOYNT: Sure. Yeah, I can give

two examples, actually, of existing things in Medicare.

So, the first is what the Medicare Advantage program implemented on a trial basis this year, which is essentially a small adjustment given to contracts based on their proportion of dual and disabled individuals. So they did a fairly complex modeling approach, found what the differences were on performance, and then turned that back into a bit of a bonus for contracts with a high proportion of dual and disabled individuals that fed into the Star system.

Another example, one that we talk about more in the report, is actually a current bonus opportunity in the physician value-based payment modifier program. And that program has only been in -- I guess it's paying in its second year now, so it's pretty new. But in the first year the setup was such that if you are a low-cost/high quality or average cost/high quality performer you're eligible for a bonus. And the flip side, if you're high cost and low quality you're eligible for a penalty or a negative payment adjustment.

But there's also an additional bonus opportunity, which is that if you earn a bonus, you can get a double bonus if you happen to have a very sick patient population. So, if your patients are in the highest quartile of medical risk, you have an additional bonus opportunity. That's not social risk, but it was an example within the current Medicare programs where there has been creative thinking about how, you know, bonuses or adjustments could be used based on the patient population that folks are treating.

MEMBER CUELLAR: So, following up, did you also discuss stratifying in terms of the reporting of these outcomes so that one could at least see where these disparities are?

DR. JOYNT: Yeah. So that was our, it's the first bullet under Number 1. And we don't use the word "stratifying" because it caused a lot of confusion actually.

There's a couple ways you could stratify. You could stratify by practices or by hospital. So, you know, the 21st Century Cures Act includes language stratifying hospitals under the Hospital Readmissions Reduction Program.

What we were really talking about was I think what you're bringing up, which is almost like subgroup reporting. It's stratifying the patients and saying how can we develop the data capabilities and statistical capabilities to be able to say, "Here's how our dually enrolled population is performing on cancer screening. Here's how our Hispanic population is performing on diabetes control." Whatever the issue may be.

A lot of the measures right now, certainly not some of the big ones like admissions or readmissions, but many of the quality measures that are a sample of patients, the sample sizes are very small when you get down to subgroups within practices or contracts or hospitals. So it's not even feasible with current data to sort of across-the-board make that happen.

So it will require some thought both on data collection, burden of data collection, of course, and also on the statistical techniques to be able to do that. But, yeah, our intent was to sort of spark that conversation with that bullet point.

MEMBER ROSS: Hi. This is Clarke Ross. In 2013 and '14, this workgroup spent two years looking at four major subpopulations in the duals population. And maybe D.E.B. or the staff or Jennie or somebody could remind me precisely what the four subgroups were. But my question was, if you examine those reports, you considered stratifying your analysis based on these four different population dynamics. Because at the time we thought it was really important to differentiate 30-year-old people with severe intellectual disabilities, 72-year-old so-called frail, elderly folks, to get a better picture of the duals population.

So, my question is, were you aware of

those, of our effort to look at those four populations, and what were your thoughts in building them into your analysis?

DR. JOYNT: Yeah, that's a great point. I cannot speak to whether or not we reviewed your work or not. I know we looked at a lot of stuff. But I will certainly look back and see.

I guess the fairest answer is "sort of." So, I suspect what you're talking about is the dual, non-aged, dual-aged, and then the folks with comorbid disability as the initial -- original reason for Medicare entitlement. And we did some of that.

In some of the measures, and some of the programs, that disability variable was actually built in. And in many of the measures, the under-65s are excluded. So in our report, it's a bit of a grab bag, actually, exactly who's in what, because of the ways that the different measures are built.

You know, part of something that we

touch on in the report, though I didn't put it in the slides here, is that there's really not probably enough understanding of the under-65s, in part because they're excluded from many of the quality measures because they're thought to be a different population. But that's probably where a lot of risk lies.

Another thing we looked at, and I'm not sure if this is one of the groups that you were looking at, but it's full versus partial duals. And we found that the partials looked much more like the fulls than they looked like the nons, if that makes sense. So when we looked at performance on quality measures, the partial duals were perhaps a little better than the full duals, but they were not -- they looked more similar to the full duals than they did to the non-duals.

And I think that's because when you're talking about full or partial duals you're still talking about pretty low income. And so we did lump those two groups together after looking at those analyses.

But if you could forward along any of that prior information we'd love to review it as we think through the second part of the study.

CO-CHAIR CHIN HANSEN: Right. Just to fill in the other two groups, other than what Clarke mentioned, were people who were duals and substance abuse category, as well as severe mental illness category. So, those comprised the four clusters that we looked at for a couple of years as a backdrop.

DR. JOYNT: Got it. We'd love to see that and think about it as we move into the second part of the study. That would be great if someone could send that along or direct us to where it is.

CO-CHAIR MONSON: Tom.

MEMBER LUTZOW: Yeah, this is Tom Lutzow with iCare. We're a managed health plan.

The IMPACT Act is quite a breakthrough, I think. But, you know, just looking back, I wish Part C plans had somehow been included strategically, because it is a measure of interest to plans. It features interconnectivity, encourages electronic interconnectivity between key providers, all joined around a common end. And it would have been helpful if somehow Part C plans were included. But that's water over the dam.

The observation that we have, we're spending a lot of time on transitions of care and readmissions with our key systems. And of course we're concerned about 30-day readmits. But if you go out to 60 and 90, with some segments of the population we're seeing, you know, 24, even 30 percent readmission within 30 days. But if we go out 90 days, it's more like 50, 60 percent.

And so, you know, I think a longer perspective is probably important from a recovery, quality of life, savings perspective, all the Triple Aim stuff. And it probably starts, in our case it does start with a readmission risk assessment in the hospital. Those that score 11 or higher on the scale are followed to home, if they let us follow to home. And a good percentage don't want us in the house. So, you know, there's that difficultly too.

But, you know, just an observation that the real impact here may be at the 60-90 day level than at the 30-day level.

DR. JOYNT: I think that's a terrific thought. And I love your comment about the home visits as well. This is, you know, I think in the hospital setting people worry about how long you get out that's away from sort of control of the hospital.

In the contract or the ACO or the primary care setting maybe it's the opposite, maybe we think about admissions instead of readmissions as sort of the, I don't know, locus of control or something. So your point is very well taken. Thank you.

MEMBER BRINGEWATT: Yes. This is Rich Bringewatt with the SNP Alliance.

As you probably know, this is a particularly important issue for our organization, and a particularly important issue for this workgroup. So, you know, really appreciate all you've done in this area.

A couple of questions relative to the research you did do, thinking about next steps in particular. There's some discussion as to whether, in terms of performance reporting, the best alternative is to look at sub-populations to compare apples to apples, so that if there is a reporting public comparison of how plans perform on Stars, you look at certain kind of SNPs versus other certain kinds of SNPs, as opposed to SNPs versus MMPs as a reporting process.

Another kind of general approach is to do risk adjustment on the different performance measurement. And the assumption is, if you do risk adjustment, if you adjust for social determinants of health on those performance measures, then you can do a proper comparison before, because it's adjusted for.

Seems to me like that probably is going to take a while in order to really get at that, as well as it's dependent upon kind of findings of the stewards of the measures. And so a second part of that question of comparison groups versus risk adjustment is, do you have any recommendations to stewards of measures for how they account for social risk factors, and other factors, you know, such as care complexity and its interventions?

But the methodology for addressing social risk factors, it seems to me, is as important as doing it. And if you don't fully take into account, particularly, as an example, neighborhoods and going down to 9-digit ZIP level, stay at a higher level where you have a mix of different income groups, you get different results.

So, comment on those two questions.

DR. JOYNT: I wish we could hand over part of our second report to you to write, because those are all the things that we've been grappling with, too. I don't think there is a perfect solution, unfortunately. There's pros and cons to adjusting, there's pros and cons to stratification. And so part of the next couple years is going to be working with various groups to try to figure out in what situations is each approach best suited.

And, you know, actually if you look in the Medicare Advantage chapter in the report, it looks like, if you look at the relationship between proportion, dual, and performance, it looked a little bit like a Nike swoosh in that the contracts that actually had the highest, the very highest proportion of duals tended to do a little bit better for the duals. So there, you know, there is something to be said for a group that looks more alike in which folks are actually, I think probably -- I don't know this for sure -- but probably investing in the types of things that can really start to make a difference in those populations.

So there's, you know, pockets of innovation and things that we need to learn from. So I don't think that adjusting everything or stratifying everything is actually the right approach across the board. And that's the tricky part that we, and, frankly, CMS who has to think about these things in a much more implementation-friendly standpoint than we do, will really think through a lot of in the next couple years.

MEMBER BRINGEWATT: If I could still do a brief follow-up on that, and that is guidelines as it relates to stewards or developers of measures in looking at the effects of social determinants of health or social risk factors on their measure. And that's been an important role of the National Quality Forum here in telling performance measure stewards that they need to make this, you know, they need to do the analysis.

But there's still a question as to how they do the analysis and in terms of what the results might be. And wondered whether you have any comments relative to that.

DR. JOYNT: You know, in terms of the measure development, we have largely left that to

NQF. It's been a pretty spectacular effort to corral all of that information and get it into this trial period. So we have not been -- we have not given particular guidance on how that should look.

I think as that trial period continues and concludes and decisions are made and all that sort of thing, we will be following that very closely. And it's terrific work, and we would certainly be remiss if we didn't take a very close look at the results of all of that work that was done when we think through where we should be headed.

MEMBER BRINGEWATT: Are these slides available?

DR. JOYNT: They're not, but the report is. So all of this comes pretty directly from the executive summary of the report, which is the first, I don't know, 15 pages or something. And then the last maybe 17 pages have a similar thing but with all the detail in terms of there's tables with the detailed findings for each program in the last chapter. So, the first and the last chapter will look very, very similar to these slides.

MEMBER BRINGEWATT: It's just a nice composite of information as the PowerPoint is presented, and was wondering whether that was available.

DR. JOYNT: Not currently.

MEMBER PARKER: I don't have a question for the report exactly, but I wondered if someone would just clarify again for me the categories, the subgroups that NQF developed in the MAP group. Because, Clarke, when you said a 72-year-old frail elderly, that's a real odd duck these days. You know, that's a real severe thing, because they're not so much that.

And so I think in terms of the old, the older old, which, you know, average is over age 85 usually. And then, you know, then there's the under-65 people with disabilities. And then some of those have IDD, and some of them have SPMI and substance abuse, usually. So think of those four groups.

So I don't know what the groups were that you guys were thinking of, or that you guys came up with. It would be helpful to know that just to put this in context of, like, what Rich is talking about, you know, and what this report is talking about.

MEMBER POTTER: I just had a follow-up on Pam's. In my memory, I thought one of the groups was the cognitively impaired. So, maybe we should go back and dig up the report and remind us all.

MEMBER PARKER: Yeah, it would be helpful to have that.

CO-CHAIR MONSON: So, do you have a burning question, Joy, or can we move on to the next presentation? Because they have another full presentation we've got thirty minutes left. But you get to go first, then, after the presentation.

All right, we're turning it back to our friends from ASPE to continue.

DR. JOYNT: Great. Robin, do you want to run with Study B?

DR. YABROFF: Sure. Thanks, Karen. And also thanks, everyone, for these really great questions.

I also want to reiterate before I start that we are very, very interested in your input and feedback. And what I'm going to be presenting is where we are to-date. We are in the middle of moving a lot of the projects forward related to what we call Study B. So your input's going to be really helpful.

So if we could have the next slide, please.

So, just to recap, as Karen said earlier, across a broad set of measures and programs we found that beneficiaries with social risk factors tend to have poorer health outcomes regardless of the providers that they see. And that also that providers serving as beneficiaries tend to have poorer performance, regardless of the patients they serve. And they are also more likely to be penalized.

Again, as Karen said, we don't really know why these patterns exist. It could be that beneficiaries have poor outcomes because they have high levels of medical risk. They could have worse living environments, greater challenges in adherence and lifestyle, or they may experience bias and discrimination.

Importantly, providers may also have poorer performance because they have fewer resources. They may also have more challenging clinical workloads, lower levels of community support, or worse quality. And many of these factors with both beneficiaries and providers are not easily measured with currently available data.

So this is what really where we're going to try and dig down a big for Study B.
If I could have the next slide, please.

So, just to give you a quick overview of our plans for Study B, we plan to build on the first report to Congress framework, which Karen described, where we evaluate risk factors and their effects on beneficiary outcomes and provider performance by program. We're going to use a conceptual framework and recommendations from a series of National Academy's reports. And this picture here is the summary report: Accounting for Social Risk in Medicare Payment.

We're going to be exploring new measures of social risk and evaluate both medical and social risk factors that are prevalent in dually eligible beneficiaries. And then, finally, examine program impacts and policy solutions.

If I could have the next slide, please. So, this slide shows the cover of the first report that was produced by the National Academy: Social Risk Factors in Medicare Payment.

And they identified five main social risk factors, the first being socioeconomic position: things like poverty, education, health insurance; race, ethnicity, cultural context; gender and gender identity; social relationships, things like marital and partnership status, living alone, instrumental social support; residential and community context, and that's where things like neighborhood deprivation or high levels of inequality come into play.

They also identified a related factor: health literacy. And it was mentioned in the request for a report, but they didn't identify it as a social risk factor per se.

And then, finally, the National Academy identified disability more as a product of social risk and health conditions, but they did not define it as a social risk per se.

So if I can move on to the next slide. So, the fourth report in the National Academy's reports had recommendations for new data sources that we could evaluate for evaluating and measuring social risk. Some of those data sources identified include surveys that can be linked to Medicare claims data. The Medicare Current Beneficiary Survey, the MCBS, is an annual panel survey and includes about 15,000

Medicare beneficiaries. And then there's also the American Community Survey.

And our plan is to link the American Community Survey at the individual level so we will have different levels of aggregation. Typically, ACS data are used at the area like the census block or census block group level, but we'll also have the individual level through this linkage. That has about 600,000 Medicare beneficiaries per year.

So, although both types of surveys have very rich data on social risk, the sample sizes in these surveys are going to limit the scope of the analyses that we can do.

Next slide, please.

So, we are going to be exploring the measures of social risk, as I said, using survey data-based projects with the MCBS and the American Community Survey. And our goals in these analyses will be to assess which social risk factors are the strongest predictors of poor outcomes, to explore interrelationships between individual and community measures of social risk, and also beneficiary outcomes, and then determine how these different risk factors influence provider performance.

The next slide, please.

We also have other projects that we plan which will evaluate medical or social risk factors that are more prevalent in dually eligible beneficiaries. And here we're thinking about things like frailty and disability, using claims data-based projects, with a much larger sample, to look at existing claims-based measures, such as readmissions, admissions, and costs.

And the goals here will be to identify and validate new measures of medical risk factors, assess relationships with social risk factors, assess relationships with beneficiary outcomes, and then also evaluate the influence of these risk factors on provider performance.

Next slide, please.

So, with that, I would like to open it

up again to both discussions and questions. I've listed, or we've listed our mailbox, if you have any ideas that you would like to share or if the discussion doesn't allow a full -- you know, we don't have an opportunity to ask additional questions. This is a really good way to contact us. So it's aspeimpactstudy@hhs.gov.

So, with that, if anyone has questions or comments, we're really interested in a broader discussion.

MEMBER HAMMEL: Hi. This is Joy Hammel. I had a question on how you're assessing social support and relationships. And, specifically, are you doing anything impact on that for dual eligibles and analysis of it?

DR. YABROFF: So we have a number of measures in the Medicare Current Beneficiary Survey related to activities of daily living and IADL, instrumental activities of daily living. We also have information about marital status and housing and living structure. So, in some of our survey-based analyses we will have some of those data.

CO-CHAIR CHIN HANSEN: Hi. This is Jennie Chin Hansen. And I have my previous background with the original PACE program.

As I look at the way of drilling down in terms of factors, which is obviously both appropriate and needed, have you done a corollary look at just the reverse? In other words, having the variables of SES, neighborhood and all, and having outcomes, frankly, that are good? So it's the reverse study to look at kind of cultural deviance of places that are possible so that it's a way of just taking a look at what's different about these communities or areas that allow a reasonable, or if not good, quality to show up.

And related to that was a question I was going to ask previously, but offer this: the national PACE programs actually number in 31 states now. I think their N is still small for what you're looking at, but it is 40,000 enrollees. And wondered if you had a chance to connect with their national association to ask about some of their data. Because I'd say 90, 97 percent of their enrollees are duals.

DR. YABROFF: So, I'm going to take a first crack. And then I will also ask Karen to talk a little bit about prior work.

But, first, in relation to the PACE program, it is something that we've started to look at because we noticed the PACE program includes a frailty adjustment in addition to medical risk adjustment. And so that's an interesting model in thinking about social risk, or at least medical risk that's more common in dually eligible beneficiaries.

And then your idea, or your comment about looking not only at those who fare poorly but also at the other side of the scale, beneficiaries and providers that do very well, is well taken. I think that's definitely something we want to be looking at, not to focus only on outcomes that are worse, but also to focus on outcomes that are better.

And my understanding -- and, Karen,

I'm going to push this back to you a bit -- but there were definitely practices that served large portions of dually eligible populations but had very good outcomes.

DR. JOYNT: Yeah, so, in the report there are a bunch of scatter plots. And I think those are fascinating for a couple of reasons. One is that, you know, we distill these relationships down to a single number, right? We say that, you know, hospitals with a high proportion of duals have 15 percent higher readmission rates, or something like that. But in reality, the variability is pretty impressive.

And I think your point is very well taken, that understanding that variability is probably pretty important. And I mentioned it, I alluded to it briefly when I said that we saw that the tail of the MA distribution sort of picks back up. And we did some preliminary qualitative work trying to understand what these programs may be doing that could explain the high proportion of duals and the high performance.

So, I think there are two possibilities. And one is a happier news story and one is a sadder news story. So, the happier possibility is that there are providers that have figured out how to do really good things for disadvantaged populations, and if we study them we will be able to figure out how we could potentially scale some of those interventions. I think that is almost certainly true.

The other piece that may also be true, which would be the sadder news story, is that -- and this comes a little bit out of the MedPAC work in MA that harkens back to the previous comment about different types of duals -- the MedPAC published some work demonstrating that having a lot of aged duals, so the over-65 duals group, was not nearly as correlated with performance as having under-65 duals in the MA program.

And so it may be that some of the positive deviance we see are actually just us not measuring very well the social complexity of the patients. And I am not a betting person, but if I were I would suspect that both of those things are going on, that there's noise in the data because of how poorly we can measure social risk, and there's real innovation and positive results happening.

So, part of what we will hope to do is to suss out both of those things a little bit better.

MEMBER ROSS: Hi. This is Clarke Ross with the Consortium for Citizens with Disabilities, which is a Washington, D.C., public policy coalition of 113 national disability groups.

And I wanted to suggest complicating your analysis to make it more relevant to the disability community. And these are around the dimensions of social relationships and residential and community context. And I'll suggest four resources for you.

So, for 50 years the disability community have been striving to promote community integration and inclusion to fight segregated living situations and to overcome isolation, which is associated with segregated living situations. So, you can look at the Olmstead Supreme Court decision, which operationalized the Americans With Disabilities Act in terms of Medicaid funding and public funding of residential programs, to get citations on this.

CMS a year-and-a-half ago published final rules on Medicaid home and community-based service settings rules. These are required dimensions to be called a home- and community-based setting.

Three, the National Quality Forum had a committee on home and community-based services and quality measures that has citations of the literature.

And, last, in the disability area we have two quality measure programs that are each operated over 20 years: the National Core Indicators and the Personal Outcome Measures.

So, to make this relevant to the

disability community, community integration, inclusion, dealing with segregation and isolation are essential around the social relationship and residential and community context parts of your analysis.

Thank you.

DR. YABROFF: Thank you so much. That's really helpful.

DR. JOYNT: I was just going to say the same thing. Are any of those measures currently integrated into payment programs? And do you know if like in the home health setting or, you know, any of those settings where there might be some overlap of that? Or is this a separate set of quality measures collected through that group?

MEMBER ROSS: So, the National Core Indicators and the Personal Outcome Measures are both used by the publicly-operated state intellectual and developmental disability authorities. And Medicaid is the predominant financing source and state general revenue is matched for Medicaid and state DD systems.

These measures are also used for other co-occurring populations. But the arena for the application of the measures are the state developmental disability authorities that use Medicaid.

But I'll send to this email address that you've given us summaries and contacts and resource information on all of these things that I've mentioned. And then there are experts in each of these domains that can provide further detail, if you want them.

DR. YABROFF: Wonderful. And this is Robin again. I just want to thank you for these resources. And also ask if we might follow up at a later date for additional information?

MEMBER ROSS: Sure. I have, again, access to 113 national disability groups, a whole spectrum and variety of groups. And there are 20 or so who are quite involved in these particular issues. And happy to refer you to all of them, and each has their own niche or specialty interest and expertise.

DR. YABROFF: Great. Thank you so much.

MEMBER AGUIAR LYNCH: Hello. Good morning, this is Christine Aguiar Lynch with the Association for Community Affiliated Plans.

I was wondering if, as part of your Study B, if you all anticipate recommending to CMS how they could collect some of this data much more broadly? So, taking frailty, for example, that's something that we've been asking CMS to consider adding that to their Medicare risk adjustment payment. But a limitation with that is that they don't collect that information widely. And as you guys recognize, MCBS and the ACS are really small sample sizes.

And then if it's not possible to collect the social risk factors data broadly for all Medicare beneficiaries, do you anticipate recommending a proxy measure that CMS could use instead?

DR. YABROFF: That is an excellent

question. One of the things we're planning on doing is evaluating claims-based measures, which would, of course, be available on all fee-for-service beneficiaries, of things like frailty, that might be possible to use.

As part of the National Academy's reports, the committee had a list of different types of data, and also their current availability and where additional research might be needed. They also identified different criteria for thinking about data and how making data better available and when. You know, some measures are fixed or fairly fixed, and that might be collected at entry into the Medicare program, whereas other measures might change over time. And so really the big concern is the burden of data collection to beneficiaries and providers.

So I think it's definitely a balancing act in terms of having sufficient information to better understand health disparities in outcomes and performance. But also, you know, having those data, additional data would be wonderful. But there is a burden, of course, in collecting those data.

And I don't know, Karen, did you want to add anything to that?

DR. JOYNT: You know, I think in the long term, if we put our, like, futuristic caps on for a moment, one could imagine a lot of these data could come from electronic health records and from other data sources that would not be as burdensome to report.

Obviously, patient-reported outcomes would be the other end of the spectrum where they're highly burdensome to both report and collect, but extremely important. And so finding a balance of ways that we can enhance data collection, harmonize across different programs and measures and private/public, and, you know, from a big picture standpoint there's a lot I think we could do. By "we," I mean a group much, much bigger than those of us on the phone. But to try to get at some of those things. And at the same time, we will need the nitty-gritty, patient-reported stuff to also be growing.

So, we are not in a position to make specific recommendations around that right now. But you've sort of zeroed in on one of the key things that would have to happen in order for a lot of these things to move forward.

DR. ZUCKERMAN: And this is Rachael Zuckerman from ASPE. I just want to add one more thing sort of in terms of what we're tasked to do and how we're tasked to do it in the law itself.

So we are asked to make recommendations. And they don't necessarily need to be in the report. So we're, as Robin said, we're going to start looking at these things as we do the second report. But it's possible that we will sort of continue thinking about it and the recommendations will come after the second report.

So, don't get disappointed if it's not right there right away.

MEMBER CUELLAR: I'm interested to

hear what you think the state of research is on measuring income, both at the individual level and at the area level? It looks like some, you're deferring some work to your Study B, and what might that work look like?

It looks like you may have looked at the ZIP Code median income. Have you thought about using, leveraging some of the work of Raj Chetty and looking at the disparities in income in the community as opposed to just a measure of the central tendency?

So those are two different questions. What do you have in the hopper in terms of looking at income and data sources for that? And then thinking about income, area income measures.

DR. YABROFF: Those are both great questions and a little bit complicated, especially for the Medicare beneficiary population. We are looking, with MCBS we are looking at data on assets, which may be more informative when thinking about socioeconomic position than actual income. Although we will have some data on income as well.

And we do plan on looking at area level median income, things like that. But also we are exploring other neighborhood sort of deprivation measures which would encompass things like disparity in income within an area.

So, I think it's --- we are in process, I think, on all of those factors. But, you know, income can be complicated to measure, in part because many people are uncomfortable reporting it.

MEMBER CUELLAR: So if you don't mind my following up, if you're doing the analyses using surveys like MCBS, that would imply that to implement it we'd have to have a data source like a survey, as opposed to relying on something like a census?

DR. YABROFF: Yeah. One of the advantages of linking claims data to the American Community Survey at the individual level is that we can actually explore that in more detail. And it may be that for some measures like income, collecting it at the area level is sufficient.

But there may be other measures, like comparing individual income to income disparities in an area, that are really quite different. So, I think it's going to be one of those it depends on what the measure is and what we find.

MEMBER CUELLAR: Thank you.

DR. YABROFF: Thank you.

MEMBER LUTZOW: Yeah, this is Tom Lutzow again.

Certainly, you know, you do want to look at claims data and so on. But there's, the real story about social risks is missing from that data.

There are ZIP Codes in downtown Chicago where Medicare-funded skilled nursing visits to the home don't occur without an armed guard present to the nurse. And I'm not sure where armed guards show up in the claims stream. And a study where you were to take those ZIP Codes and look at Medicare-funded home visits in those ZIP Codes and compare them to a suburban

Washington location and ask the question how many of those suburban Washington skilled home visits were accompanied by an armed guard, you would begin to appreciate or be able to tell the story about how social risk really does affect the delivery of healthcare.

If you look at PCP follow-up visits, post-inpatient stays in those Chicago ZIP Codes you would find, I think, that they are noticeably depressed because those folks are so fearful they lock themselves in their homes. They don't get out for fear of their lives. And that, again, affects their interaction with the healthcare system.

So, that kind of information doesn't show up in the claims stream. And it doesn't -- therefore, the claims stream really can't tell the story. And you have to get out, I think, and, you know, ask those kinds of deeper questions.

DR. YABROFF: Yeah. And I agree with you that those data are not readily available in claims. This is also an area of exploration where we can look at residential and community context factors related to things like crime and things like there are a number of, increasingly a number of measures of things like social cohesion that occur at the community level. And so that is definitely something we are exploring using data from sources, other sources that have information that are important for social risk.

But I, I think it's a really important point that without understanding the context of where people live, it's difficult to understand what's happening in terms of their outcome.

MEMBER BRINGEWATT: Yes. This is Rich Bringewatt again from the SNP Alliance.

Two things. One is, each year the SNP Alliance does an annual survey of its members that includes both quantitative data and qualitative data. And we're, we just pulled together the report from the end of last year. And part of the qualitative report focused on questions about social risk factors.

One of the questions included among the members, and they all have high concentrations of plans serving duals. But the top social risk factors observed by their care managers, and in order of priority, frequency, it was, number one was low health literacy, difficulty understanding health information.

Two was low income, poverty status.

Three was lack of mental health-related supports.

Four was living alone, few social supports.

Fifth was a tie between housing instability and transportation barriers.

Now, that's qualitative information. But I think it's probably useful information as it relates to care managers who serve these populations in terms of what they see as the primary risk factors. And we'd be happy to share more about that survey and some of those findings. And would appreciate your input in terms of what we might do next with some of that.

The second thing I wanted to note is simply, is a question as it relates to teasing out relationships particularly, I think, between social risk factors and care complexity. And it, you know, it seems to me like a very old study, Linda Fried in I'm going to say maybe 2004 or 5, but there was an untangling of chronic conditions looking at measurements for frailty, comorbidity, and different types of disability.

And it seems to me like that construct might be a useful place to look in doing some looking at the relationship between social risk factors and those, those particular factors in order to know. You know, it may be that some of these things are influencing other things in the mix. And just a thought about what might be observed there.

DR. YABROFF: Thank you. As to the first comment, I think we would definitely be interested in seeing the results of your survey. And so I'm going to ask you to use the email address, that way we all, all will have a chance to look it.

And then in terms of teasing out different relationships, we're definitely familiar with Linda Fried's work. And we'll circle back on this particular study you're suggesting.

I also want to make the point that increasingly data are available on things like housing insecurity and food insecurity, not always at the individual level, but those sorts of measures we can really start to think about a neighborhood and social context, in new and different ways. So we are definitely going to draw on expertise from our colleagues who have done work in that area.

MEMBER PARKER: Hi. This is Pam Parker.

Maybe I missed this because I had to step out. But in terms of the data sources that you looked at or you're thinking of looking at, have you looked at -- I think the biggest source of detailed functional status and, you know, home living and all of that is state Medicaid data on home community-based service clients. And there's just, you know, a huge amount of data that's collected in the assessments for that group.

So anybody that's dual and frail or disabled and meets the criteria for, you know, home community-based type services, states have huge repositories of data in incredible detail on that. And not all the states perhaps have that data automated, but I think a lot of them do. Some like, for instance, where I'm from in Minnesota they use it for risk adjustment.

So there, that's probably the richest data source I can think of for some of the things that you're talking about.

DR. YABROFF: Yes. Thank you for that suggestion.

I will say that the process of linking Medicare claims data with Medicaid data can be complicated. But I do agree that it's potentially a very rich data source.

MEMBER PARKER: Yeah, Minnesota does have -- has had linked data. But I agree with you that it's underestimated how complex that is. I've seen people think they can do it and it hasn't ---

And the other obvious thing is, of course, you know when you're talking about all these other sources of data for income and asset status on dual eligibility, that's the one thing that is a good proxy, to use the dual status indicator because that's what it's about, is income level first of all. And so, you don't really have to go that much further when you're talking about the duals I don't think.

CO-CHAIR MONSON: So this is Michael Monson. I'm with Centene.

Just I want to build on Pam's point. I would agree that the data that sits in those comprehensive assessments is very robust. And one way -- Stacey, don't look at me askance when I say this -- but one way that you might want to think about getting at that is that the MMPs have all linked that data. Right? So, and I would say all the FIDE SNPs have also done that.

So, that would be a place to potentially look to for information because there's been a ton of work by states, plans, and MMCO to make sure that that -- at enrollment that we know who's who. And it is not easy. But that is all linked. So that is a place, that is a potential starting point in order to find some of that information.

The other piece that I wanted to just advocate for was to think about the setting of where the person lives. And that might, in fact, have an implication for how you think about social risk. And we've got this population, dual eligibles, there's a disproportion of dual eligibles who live in institutions. I think we can't ignore that when you look at that as a risk factor because institutions have all sorts of problems.

And I know you'll say, you'll have trouble figuring it out, but again that data is actually available in the Medicaid data because the rate cells are all linked to institution versus home- and community-based services.

And then the final thing I would just say on that point is that you probably want to think about crossing the -- people who are on the waiver versus people who are not. Because if you're accessing waiver services, I think we would all likely think that people who have waiver services probably have a better health outcome, or hopefully have a better health outcome than those who do not have waiver services.

And, I mean that could be, that could explain your -- that tick on the MA plans at the backside. Those could all be FIDE SNPs that are actually accessing waiver services. So I will leave it at that.

Did anybody else have any other comments or questions?

MEMBER PARKER: Could I just add to that waiver or personal care.

CO-CHAIR MONSON: Yes.

MEMBER PARKER: It might be a state plan option.

CO-CHAIR MONSON: Yes, state plan. Yes.

MEMBER PARKER: So not just the waiver but also the personal plan.

CO-CHAIR MONSON: Good point.

DR. JOYNT: Yeah, and to make it even broader, you know, we've thought a lot about what -- you know, we are essentially tasked with and looking at Medicare. But all these folks operate within a context of their state-based services, their community-based services, the communities they're in. And, you know, not just from a health standpoint but from really a community, social support, and social organization standpoint. And that sort of stuff is almost certainly influencing people's outcomes.

And it's really tricky to think about how we might access and really understand some of that context. That's a great point.

CO-CHAIR MONSON: Joe.

MEMBER BAKER: This is Joe Baker from the Medicare Rights Center.

So, we've spent a lot of time talking about beneficiaries. So I guess a broad question is, you know, we haven't really talked about the quality of the providers. So, what are you doing to take a look at the actual provider community that folks are accessing, and who they are, and, you know, whether it really is resources, whether it is something else? And how are you, you know, looking to approach that, even drilling down to, you know, individual providers if that's possible or if that's something that you feel is, you know, you want to do?

DR. JOYNT: We have not thus far drilled down to individual providers. In some of the work that the National Academies did they did some best practices work in which they did look at examples of organizations that had been doing good work in this area. And I certainly think that's important.

We have, particularly for physician groups and things where sample sizes get pretty small, it's actually, it's kind of tough to sort out performance issues. So we've not really drilled down onto the individuals yet.

We may do some follow-up work qualitatively trying to understand what high performance looks like in different contexts. But I think we have a lot to learn about context before we feel confident that individuals -- that we can really assess some of those details in the context of social risk.

MEMBER BAKER: Great. Maybe I shouldn't have said individual providers because then we get all crazy about, like, oh my gosh, Dr. Smith is going to get slammed.

But I'm even thinking about systems or, you know, look, we all know there's a crappy hospital in every town. So, you know, I mean, -- and physicians there either don't have resources or, you know, they're not the best physicians or group of clinicians in that particular community.

And yet, that community is kind of relegated to that particular institution because of the fact that it's, you know, because of travel and other restrictions.

So, I guess the broader question is what are you doing to focus on, you know, because one of the lists of the things that was in your, you know, the two pieces or the provider piece was the quality of the provider, so what are you looking at in particular with regard to that to come up with either recommendations or indications of that, something that -- and the solution could be we just need better quality providers, right? Or we need different levels of providers and a different mix of providers for certain communities. So what might be the data that you're looking at to kind of get at the bottom of that?

DR. JOYNT: I'll give you an attempted response, but take with all attendant caveats, which is I think there's a lot of enthusiasm on the policymaker side that things like value-based payment programs, accountability, and measurement will improve quality even for the poorest performers. And I think there's some initial indications from the Readmissions Reduction Program that some programs have disproportionately led to improvement among the worst performers.

I think there's pretty strong feeling in the clinical community that there are providers for whom simply putting financial incentives in place is not going to adequately change anything because the problems go much, much deeper than a lack of financial incentive. And I think that's sort of the groups that you're talking about.

And I will just say we don't know how to solve that problem and have not, and will not, I think, pretend to know the answer to that one.

CO-CHAIR MONSON: Well, Karen, Robin and Rachael, thank you so much for taking the time this morning. This has been, I think this committee has found this to be extraordinarily helpful. And thank you for just doing this work. This is very important work. And it's not easy. But you are, you guys are pushing the ball in a very serious way forward on a very important topic.

So thank you for your service.

DR. JOYNT: Thank you for having us. And please, everyone, do feel free to reach out. We heard -- I have a bunch of scribbled notes with all the great stuff that we heard. But it'll be much, much more accurate if you all reach out if there's stuff that you want us to know, reports you want to send, please be in touch.

DR. YABROFF: Right. Right. And I will reiterate that I have also been taking scribbled notes. But I can't always read my writing. So thank you very much. We really appreciate the opportunity to speak to you today.

DR. ROILAND: And we can send you a transcript as well. That way you'll have your notes.

DR. YABROFF: Oh, that would be wonderful. Thank you.

DR. JOYNT: And we'll sign off. Thank you very much.

DR. ROILAND: You're welcome.

MS. MUKHERJEE: Hi, everybody. So this is Debjani again. What I'm going to do now is continue the discussion of risk adjustment and sociodemographic factors and just go over the homework that we had sent around.

And basically, the homework was asking all of you to think about five social risk factors most relevant for the duals population that HHS and CMS should keep in mind while sort of conducting and evolving their work, and for each of those social risk factors identified we wanted sort of some practical information of is that data available anywhere, is it available?

So, with that I'm going to take you through the first graph.

So, basically what we did was we provided one of the graphics from the ASPE report which listed all the social risk factors and which ones were easy to collect and not collect and things like that. And asked for you to sort of give us some feedback and add to that list.

And so the number one category was social support, loneliness, widowhood, and social capital.

Number two was sort of two groups, residential and community context, socioeconomic position, status, and income.

And some of these, because we didn't ask for a definition, some of these based on how you're defining could sort of overlap. But I just wanted to sort of let you all know that, that we realize that some of -- these are not completely sort of independent of each other.

The third group was race, ethnicity, language, gender, sexual orientation.

The fourth group was education, accessible safe housing.

The fifth was a large group, rural isolation, segregation, locus of control, valued/devalued status.

And finally, the last group, the sixth group was food insecurity and medical risk factors.

And what these groups did was they kind of built on the ASPE factors, risk factors. And the ASPE risk factors were income, education, dual eligibility, wealth, race and ethnicity, language, acculturation, gender identity, sexual orientation, partnership status, living alone, social support, neighborhood, geography, housing, and other environmental measures.

So they, even though it wasn't a one-to-one match, we kind of captured all the ASPE categories.

And then I think this was sort of something we really were looking forward to, sort of the part two, which is current and potential future data sources that could be and should be used to capture these factors.

Dual eligibility benefit level, full benefit was a proxy for income. Area deprivation indices were another one for neighborhood sort of geographic data.

Medicare Advantage plans for primary language; Medicaid agencies for sort of race, marital status; county infrastructure for community context, geographical data, census tract, health record.

And some other ones were the University of Minnesota and University of California project that we're going to hear from later today, community context, isolation, poverty, medical risk factors.

The National Core Indicators -- so that should make Clarke very happy that it was sort of one of the things that came up. And that would not only talk about communication status, social capital, isolation, also giving some data on disability, LTSS/Medicare-Medicaid Plans for housing, social support, food insecurity, food security, access to transportation, the CAHPS, HCBS.

So, we got a good sort of

representation of where the data could be found. So we want to sort of provide that to CMS so that there's -- we're not only providing our thoughts on the important social risk factors but also where they could be found so that there's some practical guidance as well.

And with that, I will turn it over to the group for guidance, from measure developers specifically, practical guidance on how can they use these based on sort of the homework information and sort of looking at gaps that we have, and sort of input to CMS on how can all these social risk factors and sociodemographic factors be used and garnered to sort of fill those gaps.

And we might have talked a lot about this and sort of killed the topic, but I just want to get some more thoughts.

MEMBER ZLOTNIK: I've been thinking a lot about this when I was doing the homework, and then kind of really reiterating, thinking about it with the presentation earlier this morning. At some level this information is kind of like, you know, it's not about medical care.

And these issues are so complex and sort of confounded together. And so when I think about, you know, what would be helpful in terms of measure development I think of either sort of in urban or sort of very rural communities have many of the same issues. And people are working in, you know, low staff, high case load, high work load, complex needs.

So one piece is like how could you develop a measure that measures something that's really complex but measures it in an easy way so that people would actually use the measure?

I don't have the answer to that question. But that's kind of a little of what, you know, has been kind of going through my mind.

The other piece is, once again I don't have a suggestion as a measure developer, developed, but are really particularly sort of transgender populations and people who really are, you know, kind of isolated, are not well-served by the healthcare community at all. We have a doctoral fellow who is actually doing a dissertation on that this year in terms of really looking at, you know, do medical providers, will they serve trans folks or not.

So, those are really, I guess I want to add to the questions and dilemmas. I don't have any great suggestions at the moment in terms of measure development. But we have to sort of balance ways to tease out complex and intersecting information with, also, measures that are easy to use.

MEMBER CUELLAR: I agree with Joan, the National Academy report has this nice framework for -- not prioritizing but categorizing these indicators by whether there are data and how hard it would be to get the data. And I think that's an important consideration. There are lots of things here that we would love to see that would be very expensive to collect, even if we had the right measure.

MEMBER LUTZOW: Yes, I mean just picking social capital, that certainly -- I mean isolation is not helpful and creates greater risk.

Where -- I mean certainly you've got buddy programs within the IDD group where these programs that intend and actually do align someone with a disability typically with some other person who's in high school. And they have a buddy relationship and with the purpose of developing social capital.

You know, I think if in the assessment process it's identified as a problem for an individual, you know, there could be easily a measure developed as to whether that goal or that need was addressed in the plan of care. And a listing of possible solutions, some kind of affiliation with a relative, affiliation with a friend, affiliation with a volunteer, a social engagement in some way could all be solutions to that, to that need.

I'm not sure why it wouldn't be an

acceptable measure if it's identified in the risk assessment as a need that has risk attached to it.

So, we try to do it. And when we see it as a need, certainly it has value in the area of medication adherence, reminders for, you know, to keep an appointment with a primary care resource. It has transportation value. This kind of, you know, building social capital around somebody who's isolated, it has value.

And it's -- now is it outside of the medical treatment area? Yes, it is. But it has an impact on the success of a medical plan.

So, are we afraid to do this? Is it a question that it would be gamed? Maybe. You know, here's where I think we have to say that this population has got some unique needs that strictly medical measures can't get at.

MEMBER BRINGEWATT: The SNP Alliance has a position paper on this topic in terms of guidelines for measure developers for dealing with social determinants of health. I had every intention to bring copies of that here today, and forgot them. So, I don't have them here today.

But I do want to highlight what the recommendations are that's contained in that report that was put together through our performance evaluation leadership group and other consultants that have been dealing with social determinants of health.

The overall recommendation is that CMS has a minimum set of standards for measure developers and stewards to consistently test the measures that what, you know, if everybody kind of chooses their own approach some may be really good and others may not, and it might be that some measures really aren't adequately tested when they're found, that there isn't an effect of social risk factors. And it may be because of the methodology used.

So in terms of suggested requirements, there's a half a dozen different categories that we had recommendations on. One relates to the sampling, the importance of including a minimum percentage of younger adults with disabled -- with disabilities, as well as older adults with multiple conditions, that if you have, you know, the sampling mix of subgroups needs to appropriately represent the broader problem in relation to the unit of analysis.

Use the smallest geographical area as unit of analysis. Studies show that variances are masked when a 5-digit ZIP Code data source is used. They really need to dig down to a 9-digit data ZIP Code as the focus. And, you know, there's also some neighborhood data level that is shown to be highly predictive of individual health outcomes.

Variables tested, you know, it seems like it would be important to establish a minimum set of factors. As the starting point, we suggest that it would include dual status, disability status, factors with significant effect on outcomes such as living in a poor neighborhood, single person household size, limited social supports, low education level, and limited education proficiency. Those are factors that are known to have a relationship, so it ought to be tested. Whatever testing is done should include those.

So just that accommodations and sound methods of administration are addressed to ensure that the methodology accommodates whatever they use, their survey methods accommodates low income, diverse, non-English speaking beneficiaries, shouldn't require beneficiaries to use cell phones, computers or internet. There's some other things like that, recommendations.

Two more. Transparency: if there is measure development work, if there's analysis work done it should be public so that other measure developers can look at it and make some suggestions and make observations as to whether they think that it has been adequately tested. And then dissemination in terms of, you know, wider public distribution of the kind of analysis that was done.

We're not saying this is the only way

it should be done. Our intent here is to really create a kind of a talking point paper and perhaps stimulate other thoughts that might generate other recommendations. But we think that without some sort of minimum standards and guidance that we're really not going to get the kind of evaluation that everybody is looking for.

MEMBER HAMMEL: Just an addition to your instruments, too, in addition to the NCI and some of the ones coming out of the disability world. I've been looking at the PROMIS initiatives through NIH, right, the patient-reported outcomes. Because there's been so much more validation of those with people with disabilities, people under 65, specific different groups there that they've been doing both qualitative and quantitative validation.

And they have a whole set of lovely social supports. They're really spent a lot of time on instrumental support, emotional support, intimate, you know, support. Different scales for each of those but also short forms of all of them. So you can do only, like, a 3-item version that's as reliable and as valid as the longer version.

They also have a new one out on societal stigma and the culture of your community and neighborhood on how you would rate it as a person with a disability or a person with any kind of diversity, kind of things in it.

So, just to note, they spent like -- it's a group of rigorous researchers, so even as I was looking at some of the stuff earlier yesterday, some of those surveys felt really kind of uninformed to me, like, from a rigor standpoint. You know, like not even having the right scale, right, you know, to even look at things. Whereas PROMIS, you know, started as a research initiative.

And now it's available as a computer-adapted testing. Easy available, anybody can join it. It's free. And it has done some really nice validation of the social and system level kinds of things that are affecting people in a patient-reported outcome measure.

MEMBER CUELLAR: Is there some communication going on to have maybe someone shepherd a measure through the NQF process?

MS. MUNTHALI: Actually, there is. We have been talking with the PROMIS folks through our Measures Applications Partnership. And that group took an early look at the survey and gave some feedback on potential measures that can come out of the survey. So we're hoping that those come through our endorsement process at some point.

MEMBER CUELLAR: Any idea what that timeline might look like?

MS. MUNTHALI: I don't know the timeline. I would say maybe in the next couple of years.

MEMBER CUELLAR: Okay.

MEMBER POTTER: I would love for that to come about, but there still has to be a data infrastructure. So somebody has to collect the measure. So that means it's either collected at the clinician's office or it's collected by the health plan. And I think that's a limitation of some of these things.

And I'm back to the point that Joan made about how to operationalize any of this in terms of infrastructure. So, I'm always coming back to data that are collected across the board, i.e., the census data and all of the things that hang off of the census data, as just one approach.

One thing that's happening in the survey research world is there's always been an issue around how do you adjust for non-response. And so there's follow-up surveys that interviewers fill out when they attempt to collect survey data about the neighborhood that the person was in, and things like that. And those kinds of information for a while now have been used to help adjust for non-response. It's called paradata. It's not always available publicly on the survey.

But beyond that, now people are using

things like Google Earth to look at neighborhoods and to attach measures to it. So if there was a database that used something like Google Earth and it was at the 9 level ZIP Code level, and it was a national database, then that's something that could be used by measure developers and others in operationalizing.

CO-CHAIR MONSON: I'm just going to pause for just a moment because Jennie, unfortunately, has to leave. But I wanted to make sure we took an opportunity, because Jennie has been shepherding this committee for several years now, and thank her for her work as co-chair. Because I don't think we would have gotten to where we've gotten today and all the great work without her.

(Applause.)

CO-CHAIR CHIN HANSEN: Thank you. That's way too generous. As many of you know, it's, frankly, a lot of the substantive contribution and the fact that this is so outwardly focused back to the consumers. So, you know, let's hope that this is a pause. And I thank you all for your contributions and your teamwork and review in this process.

As you can tell, I'm coughing way too much and causing enough of a disruption that I'm just going to head home and take an earlier flight. And so I want to say thank you to all the NQF staff. Some of you've been here a long time, some of you more recent, but, you know, talk about quality, the quality of the folks here all around are fabulous.

So, thank you very much. And needless to say, I think we're all going to dog this topic. So, we'll still be adding things. So, thank you for everyone's heartfelt and substantive contributions.

And, Stacey, I know this is your first meeting. You've been really able to manage, you know, this duality that you play. And you know now that there's a lot of people here supporting the work that you're trying to do at CMS.

So thank you very, very much,

everybody. I will leave you in peace. And also mitigate some further exposure here.

CO-CHAIR MONSON: Feel better, Jennie.

MS. MUKHERJEE: And, Jennie, definitely thank you for the staff. You've been sort of a great support, especially through the transitions and your knowledge and sharing that and being there to support us when we need you even if it's in a pinch, so.

CO-CHAIR CHIN HANSEN: Thank you.

And I don't know if Alice Lind is on the line. For those of us who've been here for a long time, Alice, your leadership has really helped us move as far as we have. So, thank you very much for your leadership in this work.

MEMBER LIND: Thank you, Jennie.

(Applause.)

CO-CHAIR MONSON: So, sorry to interrupt the flow.

So, we were having this conversation kind of around the instruments. So why don't we just continue -- does anyone want to continue that? I know, Clarke, you had your tent up. Is it on this topic? Okay, go for it.

MEMBER ROSS: So, a research question. We have these existing and functioning quality measurement systems with narrow purpose. And how do we get the resources and target the resources to expand those systems to other populations in other settings?

So, the National Quality Forum has had several presentations I've been at on PROMIS. And I took PROMIS to the National Health Council. The National Health Council is a multi-sector coalition that's been around since 1920. And there are 52 voluntary health agencies: heart, cancer, lung, MS, the big ones. And they were very skeptical of the existing PROMIS because they said, well, PROMIS is in the hospital, the clinical setting, has predominantly been used with orthopedic and cardiology. And what do they know about X, Y, and Z?

And that's our typical human response to most ideas to use something is where it's currently used, and, well, look at all the inadequacies. Because of the National Quality Forum, this work group and the Patient Reported Outcome Committee, so broken record on the National Core Indicators. The National Core Indicators were designed for people with intellectual and developmental disabilities and slowly has gotten into people with co-occurring IDD and mental illness, because that's the prevalence of the population.

Because of the recommendations of two National Quality Forum workgroups, the Administration on Community Living developed an investment into the National Core Indicators with state aging and disability agencies to see -- to pilot, to adapt and then pilot how it would work with physically disabled, non-elderly folks, and aging folks.

So, these are the -- rather than recreate the wheel and just criticize existing approaches, National Core Indicators and Personal Outcome Measures each have been functioning for over 20 years in multiple states. So, how do we convince the measure developers to get out of their comfort zone but invest money in, and let people have some running room in order to develop, expand and adapt, because you have to adapt for different populations and settings.

And how do we -- and is that a major recommendation? Some of the National Quality Forum reports have said we need to invest more in measure development. And so I'd like to carry that theme and here PROMIS is an example, and National Core Indicators and Personal Outcome Measures are examples, that have done a little bit but need to do a lot more.

Thank you.

MEMBER PARKER: As I've been sitting here listening to this, we're in some ways rightfully, but overlooking one of the major data sources that CMS uses now in both the PACE frailty adjustment and FIDE SNP, and for other purposes. And Stars is the HOS, Health Outcome Survey.

And now I can't in some ways believe that I'm bringing this up as a source of data because we've got so many problems with it. But, for instance, it's not well available in language, in terms of language and ethnicity groups. It doesn't have a good proxy methodology. It says nothing about proxy. It's a year -- you've got to be in it two years. And half the elderly are dead, you know, in the second year.

So it doesn't, it isn't appropriate for people with cognitive impairments. It's a terribly, terribly small sample. And it's not done at the planned benefit package level, it's done at the broader health plan level, so it's all mixed up, you know. So if you're trying to get at dual measures and solve some of the problems that you're talking about in the ASPE report, it's a problem.

However, it's done all the time. It's something that CMS, you know, administers and is invested in it. And it needs to be fixed. At the same time it has a lot of the data elements that we're talking about in it. It has -- you know, there's income, sex, age, functional status stuff, widowhood, single, education level, some of the other, you know, it could have the 9-code ZIP Code maybe, you know. It is -- you know, so if you increase the sample size, fix these other problems that should be fixed anyway.

One of the things that we did in Minnesota was gathered all -- there's a little profile that goes along, each health plan gets a profile, each Medicare plan gets a little profile. We took the profiles of the plans that had big Medicare broad plans versus the ones that just served duals, and so we had a bunch of D-SNPs. And we compared the two profiles. So we compiled them and compared them, you know.

And, you know, it was startling, the differences. I mean, you could see the difference in the income level and the education level and the, you know, where the -- the widowhood and the, you know. So all kinds of proxies in there for some of these things we've been talking about you could --- that was very startling.

So somehow if that instrument were fixed and utilized in a better way -- I'm just saying, you know, it should be anyway -- it could be a source of data that's collected at a, you know, a larger level. And it is already relied on. And it wouldn't be wasted effort like it is right now because it's such a poor instrument.

MEMBER LAKIN: I sort of feel like my job here is to support Clarke in everything he says. But, you know, I have been struck with a lot of this discussion and this discussion of neighborhood factors sort of brought it back to me.

You know, we're so often looking for the easy variables that are already there that we can draw on and pump them in and see whether we can explain some of the variance in what we pay or what we achieve.

And, you know, I'm struck. I've done

a lot of work with the NCI, merging data from many different states into a large data set so we can begin to look at a lot of individual factors that are predictive of outcomes for people. And, you know, in this whole area of neighborhood I've just been struck with how persistent a single set of variables is to the outcomes that people experience, whether they're loneliness, whether they're inclusion, kind of whatever they are. And that is whether people say they're afraid in their home and they're afraid in their neighborhood.

And, you know, it seems to me that, you know, we run around grabbing census track data and 7-digit or 9-digit ZIP Code data trying to get a proxy for a neighborhood when, really, asking people what their interaction with their neighborhood is and the extent to which it is comforting or neutral or fear-producing is really the element we ought to be looking at.

So, you know, and then I think, you know, that we're stuck with how these elements here are predictably more or less evident within people we call duals or non-duals, which is really, you know, it's an element of policy and procedure rather than a personal characteristic. So, you know, I think we've really got to attend to the factors that really make this difference between duals and non-duals. And those are, those are personal characteristics that people carry with them, and environmental characteristics that are imposed on them.

And I just think that all this commitment -- and I hear it again with the ASPE stuff, and I'm not at all opposed to it -- but it just, it works with existing variables. Large, you know, large data sets that are already out there that are rough proxies of these things that are really important to people as people.

And I just think there needs to be a lot of investment -- not in this top down, what can we throw in the equation and see what it predicts -- the working in projects that start at the individual level that really understand people's needs, characteristics, and elements even of personality that make these outcomes all important.

We know that those are the things that really drive medical outcomes. And yet we seem to be content to go back to the census data to see what they can tell us. And they really just don't tell us much about people.

So I think, you know, I think Clarke's been saying this in, you know, 100 different ways, but I just don't think we can overlook that. That we spend our -- we give our hearts to the person but we spend our money on these large data sets that really don't, really don't describe the person well enough to understand what we're doing.

MEMBER BRINGEWATT: I want to agree with both Charlie and Ross here. It seems to me like this is reflective of the dilemma that this group has had from the beginning in the sense that we're constantly struggling with these are the measures that are available. Of the measures that are available, which ones do we think are going to be most useful? And then we have this huge gap question where I think most of us around the room would say what's meaningful isn't necessarily what's measurable.

And so how do we get at and incent and help improve quality for duals with a limited set of data that's not as meaningful as we would like it to be, and at the same time know that there's other things that are really important and that we can't ignore? And I think part of our job here is to find some balance, you know, between those two issues.

And so I would suggest that while, you know, I sincerely agree with you, Charlie, in terms of what's most important, if we're working with what's available, and from a research standpoint there has to be some, you know, we have to maximize validity and reliability, and has to, you know, deal with sample size and data that's available on a national basis, and it has all of its criteria, you know, of the research that has been done over the last couple of years on social risk factors, the neighborhood one just keeps popping up, particularly at that 9-digit ZIP Code level as having the greatest relationship to, you know, the effects of, outcome effects of social risk factors.

So while it's limited, and I fully share your perspective on that, I think, you know, we have to run with the data that we have available that's the best of the data that we have available, recognizing its limitations.

MEMBER RAMONA: In looking at some of this information, and certainly asking about security -- and I'm not talking about armed security but how we feel safe in our -- how our patients feel safe in their community or home -- I kind of go back to that loneliness aspect. And so I was wondering if anyone has experience or knowledge of the short scale for measurement on loneliness, three questions? And if there's any way that we feel that would get to the questions that are at hand or the concerns that are at hand?

MEMBER ROSS: Well, both the National Core Indicators and the Personal Outcome Measures ask the individual of their perception of loneliness, isolation, connectiveness, this important domain of a series of questions. And as Charlie's been working with the team that's been revising those for over 20 years to get the questions as correct as one can get them.

So that's one example in the area of intellectual developmental disability, a little bit of mental illness now in ageing and physical disability, to get to that area that you've asked for. So, again, this is, maybe ASPE can do this. This, you know, this requires a huge amount of research to crosswalk these multiple existing measurement systems and try -- and everybody's selling their measurement system -- to try to figure out what's the best way of asking that question and the most cost-efficient way so people will actually use it.

MEMBER RAMONA: Yeah, this measure

does indicate that it's been tested both for in-person and telephonic proxy, some proxy information. And getting to not just the idea of physical isolation but actual social aspects of it.

CO-CHAIR MONSON: The only thing that I would add to this conversation is it does feel like there's a point at which we're not collecting all the data we need to collect, not necessarily measures but some data elements, right. There's some key data elements, we could probably debate exactly what they are, but my guess is we could probably come up with five key data -- key data questions or forms of questions we should be asking.

And some of that problem is that we don't want to have data burden on physicians or other actors in the system. And then survey data is great but it's survey data, so it's not 100 percent.

And so I do wonder, I mean as kind of back to the measure developers, duals by definition all have regular touch points with the system. Right? With the Medicaid office. I mean there's a Medicaid office. They have to get their eligibility re-upped every year, or on a regular basis, each state's a little different.

So, I mean there is, I mean it's just something for us to think about that there are some questions that could be -- some data that could be collected at that point at the system level. And so I think that we haven't talked about that at all, that there's information that could be gleaned outside of the medical system, right, and these regular touch points that we have, so that would just be the one thing that I would contribute.

Yes?

MEMBER ROSS: So, the National Core Indicators and the Personal Outcome Measures, the whole focus is extended interview of each individual person and their family. And years ago as a result of this process I took, I introduced them to the NCQA. And NCQA said, we're not going to do this. Our plans aren't going to do this, they're not going to invest the money to do the kind of time-intensive person-to-person engagement in order to learn about the real individual, and then build a system around them.

And so that's the reality is NCQA was clear, we do what the health plans want us to do. We try to improve it, but that's what we do. And we're not going to advocate time-intensive.

And so we, we can get information, a lot of ways of doing that, but who's going to pay --- we're not willing to pay for that yet.

CO-CHAIR MONSON: Well, that's what, I think that was where I was going is that there's multiple pathways. So, first of all, we would say, I think we think the NCA -- NCI -- national indicators both for individuals with intellectual disabilities and the AD one for ageing disabilities, lots of states are starting to use it. We're very supportive of that.

And it is, it is an expensive venture,

right? CAHPS is going to be expensive too -- we'll talk about that in a little bit -- to do in-person interviews.

So from a health plan perspective that would be, that's what you've heard me say it, beat that drum too, which is these comprehensive assessments. We spend a lot of money as a society, it's not just health plans, but we're getting paid to do it, to collect this information. Now, it's not going -- we can't, it's not going to be -- we can't ask every question that NCI-AD would ask because some of them it's not appropriate for us to ask and we're not going to get the right answer. But we can ask a lot of them, right? We do ask a lot of them, right? We do already.

And we all have responsibilities around person-centered care planning, which is the same idea. So this is where we're not using the data we already collect. But not everyone is in a health plan either.

So that was where I was going.

There's other places that we want to catch it. But I think that we just need to be smart about how we collect the information. I mean a part of what's happened is that there has not been -- there are no standards on the Medicaid side, right, that require common questions to be asked. Each state's a little different. Some states have a common form that plans have to use, some states don't.

And so without some common -- and then even in the states with the common assessments across their state they could be different from state to state. Now, some states are using NRI a lot more. Right? I mean there's movement but there is an opportunity to say, for measure developers to say, look, states, you know, if you want to use these types of measures then make sure these questions are asked, right. And these are the questions.

So I do think we could, I think we could tap that system. And it shouldn't cost us any more than we have today.

I think Alison was first, and then Rich, and then Charlie.

MEMBER CUELLAR: I guess looking at our homework response list I'm encouraged that the follow-on report that they mentioned for ASPE will be addressing residential community context, and socioeconomic position and income. So the report they did was a horse race against the 9-digit ZIP Code versus the duals flag. And the dual flag performs better. And now they're going to use survey data to, it sounds like do a similar exercise. You know, what about more fine-grained, individual area income, kind of what, which metric would we need to invest in to improve on the dual flag or the 9-digit ZIP Codes.

I'm encouraged by that. I didn't hear much that would address, number one, the social support, loneliness. I mean they talked about functioning and trying to go someplace with claims on that. So that, that's a little disappointing. But I am encouraged. We'll just have to see -- I mean the studies so far have only used 9-digit ZIP Codes as they were constrained, just like the first ASPE reports. And now are they going to make a pretty large investment and dig, dig deeper.

And I, you know, we'll learn a lot from that. I mean it certainly -- and I don't know to what extent the group feels that the message has come across to say that subgroup reporting by duals is really important. Not just whether the measure is appropriate for dual population, but whatever that measure is they have measure upon measure upon measure, where just reporting it by subgroup is really important and to put that information out there.

MEMBER BRINGEWATT: Thinking about, you know, where does this discussion go and what are the next steps relative to this, and we go away after this meeting, so what is it that we recommend as it relates to this issue? You know, in representing the SNP Alliance, I can tell you that from a performance measurement standpoint our members see this as the number one priority. There's nothing in serving duals. 85 to 90 percent of special needs plans are duals. And we also represent Medicare-Medicaid plans.

And this is their number one performance issue. And there's two aspects of it that are important to them. One is improving quality care, regardless of what the measures are. You know, so there needs to be responsibility and accountability to address the influence of social risk factors on health and health outcomes. And, you know, we need to work at that.

At the same time, there needs to be more progress made on recognizing the effects and measuring the effects of risk factors on health and health outcomes.

And so I think kind of at the highest general strategic level of recommendation, you know, might be that, you know, indicate to CMS this is an important area, please elevate the importance of this within your overall performance measurement strategy to the Quality Forum itself. This is an important area for duals going forward. So if there isn't a Dual Workgroup, you know, there is a focus on these issues that are really important to duals. And so that becomes another place for addressing these factors.

A third might be that we recognize the interdependence of social risk factors with care complexity. And, you know, so things like frailty and disability, certain kinds of disability and comorbidity, et cetera, it seems to me like the National Quality Forum perhaps could look at a strategy where there's more focus on those kind of population-based performance measurement issues that are cousins, if you will, to dual social risk factors.

And that might be another place where we could have some recommendations without drilling down into some specific recommendations as to what the, you know, we think the measure developer should do.

We can have a discussion about those. I think the discussion itself is useful. But I think we would be remiss if we didn't at least have some kind of focused recommendation of the importance of this going forward, you know, since, particularly since Dual Workgroup is going to be going away.

MEMBER LAKIN: Well, I just want to follow up on what Rich says. You know, I think this is all about improving quality. And there are different directions to approach that.

I think we, we should be committed to measures that are not just good for assessing but are also good for improving quality. And I don't think we've talked a whole lot about how these measures can do that.

Just to kind of return to that loneliness thing, which over time I've just become really attracted to because I think it's sort of one of the most debilitating human emotions, just to feel like you're not connected to other people. And as we've looked at it, you know, I mentioned that being afraid in your home, being afraid in your neighborhood is really connected to loneliness. Contact with family and friends, things like that, you might predictably see.

But also, you know, the environment that one lives in, you know, whether you live alone, whether you live with a family member, paradoxically is less associated with loneliness than being put in a congregate care setting where loneliness is much higher, at least among the 15,000 people in the NCI data sets that we've had.

So, you know, it seems to me that people have learned from those things. In Kentucky they really went on a rampage to reduce the use of pharmaceuticals for people in their developmental disabilities program because they noted their rates were much higher than other states.

So these things can create learning environments I think that are terribly, terribly important. But it doesn't happen when it's one agency kind of doing it, you know, over the -- doing it to agencies, it only happens when it's integrated into what people are doing. And, again, that's why I think a commitment to how we do all this stuff that we do to an organic effort to begin with people and what's important to them, and what we know is important to them, to build measures that begin to include what we can of what's important to people, is really important.

And, you know, I hope, I hope as -- and I know it will -- as this world continues to evolve, with or without NQF, that's a movement that will not be suppressed. But it would be nice if we could find ways to support it through federal commitments.

MEMBER PARKER: Well, I certainly agree with Charlie and what Rich just said. But I wanted to go back to a technical point I think Alison was making about the ZIP Code versus the -- so this is like a minor thing compared to what you guys are talking about, right? -- ZIP Code versus dual status. My understanding is that the issue is that you put them together and they're much more powerful. So one enhances the other and so it's not an either/or, it's like mix them together, you know, and use them together, that's what really counts.

MEMBER CUELLAR: I think in the report the dual ones dominate.

MEMBER PARKER: Yeah. But there's been other research that's been done on the duals, using dual factors. But then adding the ZIP Code, the 9-digit ZIP Code, and that's shown more powerful results, is what I'm saying.

So I just wanted to make a point, but I don't want to disrupt the flow of this.

MEMBER FOX-GRAGE: So I just wanted to pick up on what so much of the discussion which is, okay, so after today where do we think this should go, where should this discussion go? And so for me, and both have been mentioned, but I just want to kind of come back to it.

So, I think tomorrow the two things that would need to happen is in order for this work to continue since we're going on hiatus for however long, is I do really like Charlie's opening statement that he made yesterday which is I think, Stacey, if you could go back and create kind of an interdepartmental, you know, --

MEMBER LYTLE: I'm on it.

MEMBER FOX-GRAGE: Yeah, on it? Okay, good.

It's been very powerful in the disability community when all those different agencies come together. And so, since we can't, I think you all can.

And then, also, I think for NQF, I not only think you should but I don't think you have a choice, I do think you are going to need to look for other funding sources. And so Jennie mentioned, actually, a very good foundation, SCAN Foundation. And there are many, many others: Robert Wood Johnson, this Commonwealth -- I mean, there are lots of them. And I just think we're in that reality. And it's kind of a bummer because now you've got to write grant proposals and all of that.

But I just think it's -- well, I'm the daughter of a CPA, and diversify has been what I've always been taught. So I think no matter what, it's smart, even if the federal funding, you know, does come back I think it's smart anyway to diversify your revenue sources.

So, anyway, that would kind of be my two tracks of where I think this probably needs to move on. Thanks.

MEMBER LYTLE: I jokingly said I'm on it. But I think you're exactly right. Deb mentioned yesterday how, you know, we talk about different agencies and then you get to CMS and there's CMCS and CM and MMCO, and I think we recognize the need for that coordination. Which is why MMCO exists, because for years there was, you know, there were the two things.

And I think over the past few years, several years, we have tried to make some forays into making that communication happen. And have tried to do it first within our agency and then beyond. So, you know, Deb and I actually have talked before, and didn't just meet today or yesterday, about various things. We probably should talk more. And I think we're still in HHS. But then beyond that, because there are other agencies that still influence care for this population.

And I think we, we are aware of that. And so it's not, it's not lost on us when we hear it again that it's definitely a necessity for actually improving quality that we can't, you know, sit in our offices and never talk to anyone else.

So, I was joking, but I was also serious.

CO-CHAIR MONSON: Tom, you can have the last word. No pressure.

MEMBER LUTZOW: You know, it sounds like we're sort of talking to ourselves and maybe trying to talk NQF into going to the other side.

And, you know, going to the other side is, I think that's what, that's -- we haven't gone to the other side. And that's the biggest source of dissatisfaction when we know these conditions have an impact on quality and cost on the medical side. Somebody's going to have to do this because this vacuum is going to get filled by somebody.

And it could be NQF or it could be somebody else. The danger, of course, and I have to believe there's a segment within the halls of CMS that's fearful of this because expectations, expectations created by measures create the demand for funding. And so expectations, you know, unfunded expectations tend to be a source of pressure.

That being said, that being said, NQF, or whoever does this at a national level is going to have a bully pulpit. And gradually migrate over time other funding sources, other than CMS funding sources on the social service side to recognize these as legitimate, holistic, integrated kinds of views. And I think that bully pulpit can over time work to join the resources together toward a common purpose.

So we shouldn't really be afraid of it. CMS shouldn't be afraid of it that now they're going to be dragged into funding a bunch of stuff that really has social service content instead of medical content. But, you know, unless you cross over you are not going to save money on the medical side. You have to do this in a responsible way, probably unfunded. But that doesn't mean you don't create the expectation that if loneliness shows up in the assessment as a need, it is in the plan of care and the expectation on the part of plans, ACOs as well as plans, is to deal with it in some way. You have to deal with it.

Now, we have to be careful, ACOs, plans, we can't sell for poverty, we can't sell for the high school dropout rate, you know, we can't sell for that. So understand, you know, Clint Eastwood was right, a man has to understand his limitations, so we do have to understand our limitations. But we need to begin to create the crossover expectation, I think, for us to be successful.

MS. MUNTHALI: I'll be quick.

Just wanted to thank you for this discussion. We are recognizing the need to diversify. I know a couple of people mentioned partnership with organizations like the SCAN Foundation. We actually did some work with SCAN. So on a small scale we're doing this.

I think it's important to recognize that while we are trying to diversify, make sure we get to the right measures. CMS is probably the largest developer, funder of development, measure development. And so we need to be cognizant of that. We are looking to other partners to make sure we get the richness of data, data elements, everything we talked about. But it's going to take all of us. We're thinking us at NQF, CMS as well, and these organizations and trying to find how we can find, you know, that middle ground where we come together.

And as much as we are a multi-stakeholder organization, we are committed to doing that. I have some ideas already on how we have the relationships with developers, and it might be good to kind of put something like a consortium together of a group like this to help measure development earlier.

When Stacey talked yesterday about development, I was like, well, why don't you come to this group to get technical assistance. This is what we do. We would love to inform the development of measures earlier on in the process so they don't come to you if you're looking at measures and these are not the right ones that we need.

So, we think there's opportunity. It might be different for all of us. But we're committed to rethinking the way we do business.

CO-CHAIR MONSON: All right. We are going to take -- Quickly, Alice Lind, you're on the phone; correct?

MEMBER LIND: Yes, I am.

CO-CHAIR MONSON: Do you have anything to disclose? I'm sorry. That's not a nefarious, your disclosures of conflict of interest.

MEMBER LIND: Nothing to disclose.

CO-CHAIR MONSON: Thank you.

All right, a 15-minute break. We'll be back here at 20 to 12:00 Eastern Time.

(Whereupon, the above-entitled matter went off the record at 11:24 a.m. and resumed at 11:42 a.m.)

CO-CHAIR MONSON: All right, Erin, you're on.

MS. O'ROURKE: Perfect. Thanks so much, everyone. I'm Erin O'Rourke. I'm one of the senior directors here at NQF, and I am supporting the work of our Disparities Standing Committee. And I wanted to give this group an update on our trial period for risk adjustment for socioeconomic and other demographic factors. We abbreviate that to SDS.

So, I can skip a few slides since this

background was covered better than I could by Karen and her team at ASPE.

But so just to give you some of the NQF-relevant details. So just about two years ago we began a trial period where NQF would allow measure developers to bring forward measures that included social risk factors in their risk adjustment models. Prior to that, our criterion policy prohibited the inclusion of such factors, and only allowed developers to include clinical factors that were present at the start of care in their models.

So, during the trial period we lifted that ban, if you will, and allowed developers to bring forward measures that were potentially adjusted. And we implemented the guidance of our panel, which was actually now back in 2014, the Risk Adjustment Expert Panel's recommendations related to the appropriate use of social risk factors.

So a little bit about how we're operationalizing this. Each measure must be assessed individually to determine if SDS adjustment is appropriate. Not all outcome measures should be adjusted for SDS factors. The Risk Adjustment Panel was explicitly clear on this.

For example, they used central line infections would not be adjusted. There needs to be both a conceptual basis, so that is a logical rationale or theory, as well as empirical evidence to do these adjustments. And the recommendations applied to any level of analysis, including plan, facility, and individual clinicians.

So, during the trial period we had the standing committees that are evaluating measures that have been submitted for endorsement, in charge of really looking over those measures to determine as part of the endorsement process whether the adjustment for SDS factors was appropriate. We asked the committees to consider both the conceptual and empirical basis for adjustment, utilizing standard guidelines for selecting risk factors.

If SDS adjustment is determined to be appropriate for a given measure, we do endorse one measure with specifications to calculate the SDS adjusted measure as well as stratification of a non-SDS adjusted measure. As recommended by the panel, these specifications for stratification should always accompany an SDS adjusted measure.

We want to ensure there's transparency. One of the main concerns when we implemented these recommendations and started the trial period was doing these adjustments could worsen disparities, mask them, adjust them away if you will. So --

CO-CHAIR MONSON: Erin, what's the difference between -- what do you mean by some are stratified but non-adjusted?

MS. O'ROURKE: Sure. So, the adjusted measures basically bake in the calculation for the social risk factors in its risk adjustment model. The stratification does not include those factors in the risk adjustment model but, rather, lets you break apart each group. So if you had a measure adjusted for, say, dual eligibles, in the risk adjustment model it would just basically not show you the difference. It would already pre-calculate the impact that dual eligibility would have on that person's outcome or their risk of, say, being readmitted.

The stratification would let you break it down by subgroup so you could see what was the rate for people who were dually eligible versus non-dually eligible.

Hopefully, Elisa, did I get that? It's one of those things as soon as you try to put it into plainer English you worry about losing meaning. So, hopefully that helps a little.

MEMBER CUELLAR: I think they came up with a phone call where they say they're moving to the term subgroup reporting. And you can do that on an adjusted or unadjusted basis.

CO-CHAIR MONSON: Right.

MS. O'ROURKE: That's a great term. We should probably move to that. It makes more sense.

So to really address this concern about worsening disparities, NQF brought together a Disparities Standing Committee. They have a number of items that they are charged to do.

The first is to develop a roadmap for how measurement and the policy levers associated with it can be used to actively eliminate disparities. So, to kind of piggyback on what Karen and her team were presenting, one of the main things they're working on now is to really think about what that plan for equity measurement could look like. What topics would you want to measure to really promote equity and start to reduce disparities.

And then the next step when they come back together in June will be to think about how we can push to get those measures into use to make equity a key focus in things like public reporting programs, value based purchasing, to try to capitalize on some of the shifts around payment that are happening.

We also asked the Disparities Committee to help us oversee the trial period and to provide guidance. We've been giving them periodic updates over the past two years. They are also asked to provide a crosscutting emphasis on disparities across all of NQF's work. So, as we move to the conclusion of the 2-year trial period we'll update the committee during their June meeting and get their guidance on a potential path forward here.

They actually were just meeting Monday and Tuesday. So we presented them our evaluation plan, what kind of data we could put together to support their recommendations, and see if there was anything else they thought we should look at as we start to make potential recommendations around whether we should make this a permanent change in policy, if we should extend the trial period, or we should put that ban back in place and, you know, that looking at social risk factors and risk adjustment models is the wrong way to go.

So we have asked the standing committees to consider a few key questions when they're looking at SDS adjusted measures. First, we asked them to look if there's a conceptual relationship between the factor being considered and the focus of the measure.

Was that risk factor present at the start of care, is there variation in the prevalence of that factor across the measured entities?

Do the empirical analyses, that is the ones provided by the measure developer, show that the social risk factor has a significant and unique effect on the outcome in question?

And is the information available and generally accessible for the measured patient population?

So, I did want to update you all on some of our findings to date. This is still really just at the start of gathering our data and evaluating what's happened over the past two years. So there will be more information to come, but a little sneak preview, if you will, of what we've been finding.

So, as I was saying, since April 2015 we have asked all of our standing committees to consider the potential role of SDS factors in the evaluation of all submitted measures, with a particular focus on outcome measures.

We also had the Readmission and the Cost and Resource Use Standing Committee go back and take a look at some measures that were endorsed with the condition that the developers perform some additional analyses to determine if there was a need for the inclusion of SDS factors in their models. Those measures were endorsed immediately prior to the start of the trial period, and the potential need for SDS adjustment had been a big theme in their endorsement reviews.

So, ultimately, the board of directors put this condition on their endorsements. So we worked with the developers to bring them back in and do some, some additional work around there to determine the impact of these factors.

So, probably not a shock to anyone given your conversation this morning and what Karen presented, we've had a significant number of measures come forward with a strong conceptual basis for SDS adjustment. However, when you look at the empirical analyses to support whether or not you put that factor in your risk adjustment model, frequently it just hasn't been there. It doesn't change the performance of the risk adjustment model. It's a very, very tiny effect, so developers have chosen to leave it out, given some of the politically charged nature around this topic.

To date, we've actually had a relatively small number of measures that have been endorsed with risk adjustment for SDS factors. Some examples are we have a patient reported outcome for pediatric experience of care that is survey-based and takes into effect the caregiver's education level to the person completing the survey.

We've also had some measures in the nursing home setting. One around hospital readmission and one about discharge to the community that look at things like payer mix and marital status, so as a proxy for caregiver availability.

And, whoops, I had the same bullet twice, apologies there.

So, again, really just to tack onto the conversation this group's already been having, we've really found there's very limited availability of patient-level data. We've had developers do some extensive work to get to 9-digit ZIP Code. But it's not easy to do and not really readily accessible.

Risk models using currently available SDS adjusters are not demonstrating an association for measures that have a clear conceptual basis.

MEMBER CUELLAR: Can you clarify?

What do you mean by the first point? The data are available as long as we have the ZIP Code. Is it the fact that the ZIP Code, 9-digit ZIP Code isn't available? Because the other thing's been constructed and you can just download.

MS. O'ROURKE: Sure. So basically getting to that 9-digit ZIP Code has been challenging. And when they did the 5-digit analyses it was just not granular enough.

MEMBER CUELLAR: I'm not understanding that. The 9-digit data are available. That census data has been constructed. So is it that they don't have the patient's 9-digit?

MS. O'ROURKE: So, the work to match it, yes, to get to the patient's -- to match the patient to their 9-digit ZIP Code through the claims data.

MEMBER CUELLAR: Well, they don't have an address? Okay, the problem is that many times they don't have an address?

MS. O'ROURKE: Yeah.

MEMBER CUELLAR: So maybe that's what

we're talking about.

MS. O'ROURKE: Yes. Some, yes, that's a better way to --

MEMBER CUELLAR: It sounds like we don't have the 9-digit ZIP Code census block data. But that's the piece we do have?

MS. O'ROURKE: Yes. So that is available. And we've had developers use things like the AHRQ SES index to get to that and then try to match that with what's in the claims data. But it's proven to be quite a lot of work to get that match done. And --

MEMBER CUELLAR: Because they don't, and the claims don't have the 9-digit ZIP Code?

MS. O'ROURKE: I believe that is the issue, yes.

MEMBER CUELLAR: It would have to be. Because if they're arguing we don't have census data down to the 9 digits, I can download that for them in about 20 minutes. Right?

MEMBER POTTER: Not everybody has the ability to process on 50 million records. That's really the problem.

MEMBER CUELLAR: Right, but I don't think that's -- the bullet's not capturing the issue. Either it's that you don't have the address, which means you don't have a 9-digit ZIP Code, or it's that you don't have the computational power to merge on, you know, a few more variables.

MEMBER POTTER: But I think the issue of it not being easily accessible is a legitimate issue. I mean, if we were merging at the county level as opposed to the 9-level ZIP Code, there's a whole database that HRSA puts out called the Area Resource File which has all kinds of stuff at the county level. And it's readily available.

This isn't readily available, you know.

MEMBER CUELLAR: Well, but it could easily be, since it has been constructed, it could easily be made available. And if that's the -- if it's the computational power of, gee, now I've got to merge it on, that's no different from merging on the Area Resource File conceptually. If your programs can do one, they can do the other, if the computers can. I mean, I'm not saying they all can.

But then it's a can we make this 9-digit ZIP Code file with the census information more accessible? That we could do.

MEMBER LYTLE: Without having to pay for it.

MEMBER CUELLAR: Right. Right.

MS. O'ROURKE: So we've also heard some other concerns about the factors that developers have selected and analyzed. Some of the proxies that they've been using to get to a person's actual socioeconomic status and their social risk have not really been adequate. They're just a little too blunt to show a person's actual -- I think Charlie put it very well -- the data versus the person issue.

We also had a lot of push back from some of our stakeholders about when developers have included race as a potential variable. We have not adjusted any measures with race in their final risk adjustment models. But it's been a pushback from stakeholders that developers have even looked at that.

We took it to our Disparities Committee for some more guidance there, and they came down more along only really if there's a genetic basis, emphasizing it should not be a proxy for socioeconomic status.

We've also heard from some of our stakeholders a call for a more prescriptive approach to how developers are testing these variables. For some background for those of you that haven't been involved in our endorsement committees, NQF does not tell developers what methods to use to test their risk adjustment models. We also don't give a standard set of variables that they should look at. It's up to them to make the decisions about those things and for the standing committees to determine if they agree or disagree.

We've heard some, some calls that NQF

should be a little more prescriptive in this space, but that's how we approach both clinical and social variables, so it's something to think about.

I did want to bring forward a slide that has some of the implication for your family of measures. Gave an update on this last year, so just to close that loop.

A number of measures in the family were reviewed during the trial period, mostly all dealing with readmissions. All maintained endorsement without social risk factors included in their models.

Let's skip that slide because I think we're all familiar with it after Karen's presentation.

But just to keep you informed on some of the next steps here. The Disparities Committee met earlier this week, provided some feedback to us on the evaluation plan for the trial period. We'll bring them back together in June to share the results of the evaluation and to get their input on what could be a path forward here.

Similarly, we'll take it to our Consensus Standards Approval Committee in July for any thoughts that they might have. And then in July the board is tasked with determining what should be our path forward here.

CO-CHAIR MONSON: Erin, can I go back and ask a question?

MS. O'ROURKE: Of course.

CO-CHAIR MONSON: So, on those five, or the ones that you looked at when you said that there was no -- so I'm just reflecting on the prior conversation where dual status alone seemed to have a major impact. Is there no adjustment because you're using dual as a subgroup, therefore within dual? Or is it there's no adjustment for dual/non-dual status?

MS. O'ROURKE: No adjustment for dual/non-dual status.

CO-CHAIR MONSON: So why do you think that is that the ASPE work shows something so different than what you guys have been doing? There's something that doesn't fit is what I mean.

MS. O'ROURKE: I think you hit on the million dollar question. We've heard a lot of different potential reasons. Some developers used whether it improved the C-statistic of the risk adjustment model, so the, really the stat you look at to see how well your risk adjustment model is predicting outcomes as their metric for whether they'd include or not include. And anything that didn't improve that, they didn't improve.

We've heard some concerns that some of the developers looked at all of the clinical factors first and then baked in things like dual eligibility as a risk adjuster. So that by the time you got to that, most of the risk was accounted for by the clinical factors and you just had a very small effect seen by dual status.

So, I think it's something we've gotten pushback on from particularly the provider community. And as Karen was saying, there's a lot of evidence that there's something there, and it's not showing up when developers do these calculations.

CO-CHAIR MONSON: I think Alison wanted to go first and then D.E.B.

MEMBER CUELLAR: Is it also possible, given what you've said, that NQF doesn't impose any methodology, that they didn't even test dual as one of their factors?

MS. O'ROURKE: Yes. So not every -- you were not required to test dual.

MEMBER CUELLAR: What they would have said is risk adjustment didn't matter, but they may not have looked at dual versus non-dual.

MEMBER POTTER: There's also data systems that don't adequately capture secondary payers, which in this case is the duals. And so when you're talking about clinicians and hospitals, the primary payer is Medicare, and then the hospital or the doctor would have to have another variable that said they were also covered by Medicaid and somehow captured that in adjustment.

So, I think some of it is just accessing the information at the clinician or the provider level, which is different than what CMS can do when it goes and looks at the enrollment file and it has just one little variable it can attach to a person that says they're dual.

CO-CHAIR MONSON: Tom.

MEMBER LUTZOW: I'm sure you have, you know, a good statistical consultation. We certainly have found that evaluating one factor at a time may not be the right way to go because these factors play -- they have interrelationships, interactions behind the scenes. So, hopefully, as you look at this, these things are not got from isolation but looked at in groups because as groups they may have an impact. But I'm sure you're looking at that.

MS. O'ROURKE: Yes, it's a good point, how you look at things can have a significant impact. And that's certainly something we've heard from some of the stakeholders that are telling us these measures should be adjusted and we need to push harder here. So that's a good input.

CO-CHAIR MONSON: Rich.

MEMBER BRINGEWATT: Yeah. I think this is in part related to what often happens when there's a self-evaluation done where the measures used are defined by the person doing their own evaluation, you know, and where there aren't outside standards.

And I think part of what National Quality Forum is about as it relates to provider standards is that there is some standard relative to whether the measure is adequately tested, and what are the factors that were involved, and what is the population mix that was involved? And it's, you know, the National Quality Forum has a very rigorous set of standards for whether it's a good measure or not. And I think that and all due respect to the people who have done these, done the measurement, I'm not trying to cast aspersions on anyone. But I think it's just good scientific process that there be standards set for testing of whether social risk factors are adequately addressed in any of the measures used.

You know, we, that kind of standard is applied to every other performance measurement for providers, for plans. I think it should be also used for the testing of the adequacy of whether the measure includes or excludes accounting for social risk factors.

CO-CHAIR MONSON: Alison.

MEMBER CUELLAR: I don't know that the statisticians could ever tell you whether you should adjust for a risk factor because the implications are profound. Right? So, if you serve a lot of duals, according to the ASPE result, and your, let's say your quality was a 4, if you serve a lot of duals they're going to upweight it and it's going to look like a 4.2. Right? Just sort of roughly speaking.

Whether you want to do that or not is

a policy decision; right? You're going to get, right, it has to do with penalties one way or the other, even playing fields, which playing fields do you want to even out, which ones do you not want to. Neither the Academy nor ASPE is telling CMS whether or not they should do it, they're just saying statistically would it even matter.

So, but none of them disagree with the idea that one can report it separately by subgroup. But that's innocuous. That's not moving money from A to B or C to D in ways that you may or may not -- they're tradeoffs. I mean, everyone is very clear that there are tradeoffs. There are tradeoffs. Once duals in an ASPE report is statistically meaningful, then they simulate would it have moved money to use? And the answer is yes. Is that desirable is another question.

So I don't think NQF alone is ever going to be able to tell you you must do this adjustment because it's going to depend on what your goal is. It could tell you whether statistically it mattered. And then NQF could, I would think, say subgroup reporting makes sense if statistically the duals group looks different. And that's not going to move money from A to B, it's just going to shine light on subgroup differences that appear to matter in the analyses.

So, I'm not exactly sure why the subgroup's charge is to figure out whether we should do the quality adjustment as opposed to just give us some statistical information. And if the statistical information is so heterogeneous, it's very difficult to process. It's like saying we have no definition of validity, reliability, sensitivity. Just throw any number you want at us and we'll review it and assess. I mean I don't -- it seems a little odd.

MS. O'ROURKE: Great. So, I really just wanted to get some input from this committee, continue the conversation that you've already been having previously and just now. As we start to bring this to the Disparities

Committee is there any input or guidance you'd like us to think about, particularly since dual eligible status has been one of the most common factors that our developers have tested? Any thoughts on that? Arguments for or against? Anything we should be cautious about?

MEMBER POTTER: I just have a question really. Is there someplace where there's a list of the measures? I mean, if we go to the QPS is there a way to search and have the measures that have this adjustment come up?

MS. O'ROURKE: It will be coming in the future weeks ahead. We're actually wrapping up our data collection process now, asking staff to fill out all this data so that we can go through things like QPS and perhaps create one of our -- I forgot the term -- a portfolio of measures that has that information for you.

So, stay tuned. It will be coming.

We're also doing some work around trying to get out what variables were looked at for each measure. You know, we want to be very transparent here about what happened and get the information out to you also. It's coming.

MEMBER CUELLAR: It would be helpful. Even just to know whether or not they looked at dual or not and with what other variables, and then what methods they used.

MS. O'ROURKE: Sure. Yes.

CO-CHAIR MONSON: Pam.

MEMBER PARKER: Have they looked at dual at all? And you said that might be separately reported in some of them, or not?

MS. O'ROURKE: They should. If they included it, they would have the instructions to stratify so you could report that.

MEMBER PARKER: And does that mean that the measure is used comparing the dual group to another dual group? Or does it mean it's just reported out like that?

MS. O'ROURKE: So, it would be how it's endorsed. Unfortunately, we don't always control how a measure is used. Someone could choose to use what would be the non-endorsed version that doesn't include a certain adjuster or how they reported out. So, it's a recommendation for how NQF would like to see the measure used and what we've endorsed as the best practice. But the decision about which version of a measure to use really rests with the payer or the purchaser group that's doing the evaluation.

MEMBER PARKER: Well, I can see where it might depend on the specific measure. But it would, I would think it would be good to encourage them to try to say whether or not, you know, it would be appropriate to compare duals against duals in that measure rather than duals against everybody else. You know, if you're going to stratify it and you're going to report it out in that way, then that gives you that opportunity. And so it would be good if they, you know, if they utilized it in some way and suggested that, so.

CO-CHAIR MONSON: Tom.

MEMBER LUTZOW: Yeah. I'm hopeful

that the guidance would maybe be adopted in terms of testing for SES or social determinant impact. I think we saw some measures yesterday, the HIV measure, where the developer assured us that there was no social development impact.

Rather than an assurance, it would be nice to hear that we tested this against NCQA's standards and it came up negative. Okay, now I, now I believe that. But, hopefully that is a product that you come out with.

CO-CHAIR MONSON: Rich.

MEMBER BRINGEWATT: Yeah. I, well, a couple of things. One, in looking at the options that are under consideration here, I think they're reasonable options to consider because it provides a good basis for discussion.

I would be extremely disappointed if the decision was made to rescind, you know, no longer go forth in doing this. This is a, you know, to be fair to measure developers, this is even though research on the effects of social factors in health and health outcomes is well established, you know, that has been demonstrated in multiple ways for decades. Dealing with risk adjustment or sorting out how to account for that performance measurement is a relatively new science.

And so I think at some level we have to be cautious and careful about acting too quickly. At the same time, I think we have to be careful that we do in fact act because there's clear evidence that there is an influence here. And even in relation to all-cause hospital readmissions, there are a couple of very, you know, scientifically -- scientific studies that show exactly the opposite, you know, that there is, that social factors do influence the reporting of all-cause readmissions.

And so, that raises questions about methodology that I think are appropriate to ask for an organization like the National Quality Forum. Which drives me towards, you know, a suggestion that there be more work done in this area, that now is not the time to stop. But it seems to me like, you know, there have been minimum standards for what is necessary in looking at other measures. Seems to me like there should be minimum standards for the testing of social factors in performance measurement.

And there's differences of opinion as to what that should be. And that's appropriate. But, you know, I think it would be remiss if there wasn't some leadership provided, even if it's a matter of guidelines and not a requirement. You know, at least some leadership provided on the part of the National Quality Forum for addressing this issue in a scientifically rigorous way. It is the quality standard, so define the quality standard for doing this.

MEMBER ROSS: Hi, Erin. I don't know if you were here this morning. I mentioned but could not remember the four major subpopulations in the duals population that this committee analyzed and studied in 2013 and 2014. But I would just remind you and remind all of us, those are official reports from the National Quality Forum to CMS to go back and examine what we said about those four.

We spent a whole meeting defining the four and then we studied it for another year-and-a-half. And I can't remember precisely, I could guess but I can't remember precisely what the four were.

I apologize, I didn't listen to the Disparities Committee meeting Monday and Tuesday. But previous meetings, the sole, single disability expert on the committee was disappointed. The National Academy of Medicine concluded that disability was a product of but is not a factor, a risk factor. And so the representative previous to this week -- because I don't know what happened this week -- was disappointed that this Disparities Committee just accepted the National Academy of Medicine recommendation that we shouldn't really focus on disability; it's a product and it's important but it's not a risk factor. And the disability field, obviously, would like that kind of thinking further discussed.

MS. O'ROURKE: I think those are both great points. I think to your first point, that's a great suggestion to go back to those reports and see what could be done about maybe breaking duals down. If that, if that factor is still too not getting granular enough, what could be done to break that down.

So I can pull those and see what this committee previously said.

For your point about Dr. Iezzoni's strong feelings about the NAM report, that was a prominent feature of discussion at the meeting earlier this week. So we I think will have some language in our reports perhaps challenging that and suggesting potential different directions.

MEMBER ROSS: Thank you.

CO-CHAIR MONSON: Other comments? Anything else you need from us, Erin?

MS. O'ROURKE: No, this was great. Actually I was in the back listening for most of the morning. And this has been really helpful input. It was remarkable how well the conversation tracked between this group and our Disparities Group. So that gives me some encouragement that we have some validity there and we're on the right track.

So, thank you for all of your time and input. And we'll be bringing this to our Disparities Committee in June as they start to think about what's our path forward here. So thank you so much for your time this morning.

CO-CHAIR MONSON: Thank you, Erin.

I think we're going to go to public comment.

MS. BUCHANAN: Yes, hi. Shawn, would you mind opening up the lines so we can hear from any members of the public. Additionally, if you are not connected via telephone and you would like to chat a question in your chat box, NQF staff can read it aloud.

OPERATOR: And at this time if you would like to make a public comment, please press star then the number one on your telephone keypad.

And we have no public comments at this time.

MS. BUCHANAN: Thank you.

CO-CHAIR MONSON: All right. So we are actually like a half hour ahead. But, unfortunately, our next speaker is external so we can't necessarily get them to move up faster. So the bonus to that is there's a long lunch then.

So, we need to be back here, we're going to start again at 1:15, 1:15 Eastern. Enjoy lunch.

(Whereupon, the above-entitled matter went off the record at 12:16 p.m. and resumed at 1:15 p.m.)

CO-CHAIR MONSON: All right. We are reconvening, finishing lunch.

DR. ABERY: Okay.

CO-CHAIR MONSON: All right, guys. We've all been guilty. So, we're sitting. We're sitting. We're not chatting. We're listening.

Rachel is taking over -- or Kate is.

MS. BUCHANAN: Hi. Thank you very much. And, Brian, I just want to make sure that we can hear you.

DR. ABERY: Yes. This is Brian here and Renata Ticha.

MS. BUCHANAN: Great. So, we just want to briefly introduce Dr. Brian Abery, who is co-director of the Rehabilitation Research and Training Center on HCBS Outcome Measurement at the University of Minnesota's Institute on Community Integration.

And he is also the co-director of the Institute's Educational Assessment and Intervention Program.

And we'd also like to introduce his colleague, Dr. Renata Ticha, who is a research associate and a principal investigator at the University of Minnesota's Institute on Community Integration.

And if you all wouldn't mind just saying "next," we'll move the slides along for you. And with that, we'll take it off.

DR. ABERY: Okay. Well, thank you very much. We appreciate the opportunity to share the work we're doing with our Research and Training Center on Home and Community Based Services Outcome Measurement with you.

And we will try to stay within our time limit today. I know I sent a lot of slides and we're going to move through these fairly quickly, but we wanted to give you some examples of some of the learning that has occurred since we began our project approximately a year and a half ago.

Can we go to the next slide, please. Although some of you may know a little bit about our Research and Training Center, I just wanted to introduce you to our primary partners before we got into content.

In addition to the University of Minnesota and our Institute on Community Integration, we are working with Mark Salzer, Gretchen Snethen and Beth Pfeiffer at Temple

University; with Steve Kaye from the University of California-San Francisco; with John Corrigan from The Ohio State University; and Joe Caldwell from the National Council on Aging.

We are also working with a number of the organizations that are involved in administering HCBS outcome measurement programs. And of course we're funded by NIDILRR.

Next slide. So, I'd like to begin by just kind of letting people know that we are a five-year research and training center.

And our goal was to undertake a program of research that would really help us collect and analyze the data necessary to be able to report to a wide variety of end users, specific measures related to HCBS outcomes and quality that were both psychometrically sound and could be used with multiple populations ranging from individuals with intellectual and developmental disabilities, physical and psychiatric disabilities, traumatic brain injury and age-related settings.

And we really wanted to focus on the development of measures that could be used in specific settings and contexts along with relevant risk adjusters.

In addition to the development work we are doing here, we're also providing training and technical assistance to a variety of stakeholders on outcome measurement in the HCBS field both in Minnesota and nationally.

Next. Next slide, please. What I really want to kind of focus on before we get into the content, is the fact that, you know, as we have thought about our research and were developing our center, we really wanted to start at Base 1 with the National Quality Forum's framework and really take it out to stakeholder groups, a variety of stakeholder groups nationally and figure out, first, what they thought was most important to measure.

We have in the NQF framework, what experts in the field thought was most important to measure. We really wanted to get out into the community as our first step and find out what those other stakeholders are thinking.

As a second step, we wanted to identify gaps between the current measures which are out there and available for use in both the NQF framework and federal and state policy operational drivers.

And then, to identify which measures are currently psychometrically sound and robust enough to be used across multiple populations.

And then, to really start a three-and-a-half-year process that includes development - field-testing to support the refinement and development of new measures to get at those gap areas.

Next slide, please. So, our goal is not to create a master instrument. I mean, we're not trying to replicate what, you know, other groups who have national projects are undertaking with respect to HCBS outcome measurement.

It was really to try to take measurement in this area, kind of that next step, so we could report to end users, you know, what measures are psychometrically appropriate to use with specific populations and contexts and settings and with the eventual goal of NQF endorsement in those areas that our stakeholders indicated were most important and which we identify gaps within.

Next slide. So, we have kind of conceptualized our work over the five-year period as consisting of six different studies.

Study 1 I've already alluded to; soliciting broad stakeholder input on the NQF measurement framework.

Study 2, which is also well underway, is a gap analysis and we're taking a look at the NQF measurement framework and current instruments which are being used to measure HCBS outcomes.

Study 3 focuses on the identification of high-quality/high-fidelity implementation practices in HCBS measurement.

In Study 4, we'll be working very closely with our colleagues at Temple in the refinement and development of new measures.

Study 5 is a large-scale study focused on kind of a national study to ascertaining the reliability, validity and sensitivity to change of the measures we have refined and newly developed.

And then, finally, Study 6, which is a little bit out of order there, which we're already working on, is to identify and eventually test in Studies 4 and 5, relevant risk adjusters for home and community service-based outcomes.

Next slide, please. So, I'm sure all of you are familiar with the National Quality Forum framework. So, we can just kind of put this slide up for just a minute or so with its 11 domains and two to seven subdomains within each domain.

This was kind of our beginning point of our research in Study 1. We can go through the next two slides quite quickly just to give the group -- you can see we have the slides that focus on human and legal rights, service delivery and effectiveness, workforce, system performance and accountability. And the more individualized outcomes, choice and control, community inclusion, equity, holistic health and functioning.

So, if we go to the next slide, we can talk about our first study. And that really was designed to gain the input of a critical set of representative stakeholders from around the country.

Next slide. What we are using for this process is a group of stakeholders who include individuals with disabilities across that disability group that we are working with, family members of persons with disabilities when that is appropriate, providers, and program administrators both at a state and national level.

And, again, our disability populations of focus were individuals with intellectual and developmental disabilities, physical disabilities, traumatic brain injury, psychiatric disabilities and mental health challenges, and age-related disabilities.

Next slide.

CO-CHAIR MONSON: Brian, can I just interrupt for a second? This is Michael Monson from Centene.

So, the one group I didn't see on there, and I say this as a self-interested party, was health plans.

Are you planning on incorporating them as well?

DR. ABERY: I didn't catch that group.

CO-CHAIR MONSON: Health plans, managed care organizations.

DR. ABERY: Oh. Yes, they were actually part of our program administrator and provider groups.

CO-CHAIR MONSON: Thank you.

DR. ABERY: So, we're using a process that we have developed here at the University's Institute called participatory planning and decision-making. And I'm just going to quickly kind of go through the process so you have a feeling for what we did as part of this group.

In each phase of the process, stakeholders are distributing ideas to the framework --- to the NQF framework under discussion.

They have the opportunity to add new domains and broad domains, if they want, subdomains. They can suggest the removal of domains or subdomains they view as unimportant.

They then provide importance weightings for each domain and subdomain. They discuss their thinking while undertaking these importance weightings. And then they do a second round of importance weightings.

Following the weighting of both domains and subdomains, we then are able to determine proportional importance weights that are assigned to each subdomain.

So, it's a very interactive process that takes about two to two-and-a-half hours to implement.

Next slide, please. And we basically did our implementation with 54 groups over the course of the first year-and-a-half of the project, including over 280 participants.

We have three more groups left to collect. But, again, those groups cut across nationally all of those --- all of those stakeholders that we previously discussed.

You can see from the next slide, which is just a map of where our participants resided, number of groups conducted on the east coast, but we ended up trying to get groups in both rural and urban areas across the U.S.

The next slide shows our disability population. I'm not sure whether that's going to come up large enough for you to really see. But as you can see, we cut across all of those groups that we are responsible for as part of our Research and Training Center.

So, the measurement framework question that we really wanted to ask in this study, was which NQF domains and subdomains were viewed by our various stakeholder groups as most important to measure, what if there were differences between our stakeholder groups and disability populations and how they prioritize those domains, and really to look at the extent to which the stakeholders supported the current NQF framework as including everything that they thought was a critical importance to measure and looking at HCBS outcomes.

Next. So, this next slide just gives you an indication of our importance weightings across the various domains of the NQF.

Those numbers on the right-hand side include both the importance weightings, which are weighted on a zero-to-100-point scale for most groups and our standard error.

And as you can see, the individuals who were part of the NQF group that developed the framework, you know, basically hit the nail on the head, so to speak, in that all of the basic domains that were identified by the group of experts were considered by our stakeholders across the country, across these disability and stakeholder groups, as to the importance.

When we looked at the subdomains, there were some additional differences. And we can talk about those in just a minute.

Next. We took a look at both the domain and subdomain level using multi-analysis variants, to really take a look at whether domains were evaluated in a similar manner by disability population and stakeholder type using both a full factorial design and some post-hoc comparisons to really see whether we had some significant effects at the domain and subdomain level.

Next slide. So, we did find that there were, in general --- there was, in general, a high degree of agreement among the groups, but there were some areas where stakeholder groups differed with respect to how important they thought these outcome areas were to measure.

And that's important to understand that they were basically working off the NQF framework and the operational definitions that each of those areas provided and were really providing weightings of measurement importance rather than how important are these domains to them personally.

The areas where we found some significant effects were in choice and control, consumer leadership in system development, human and legal rights, community inclusion, and service delivery and effectiveness.

Next slide. I'm going to just go through the next several slides pretty quickly to just give you kind of an overview of some of the things that we found. And many of these differences are things that, you know, we really expected to find.

So, for example, you know, we found that, you know, across the groups, the group of family members -- or the groups of family members tended to weight personal choice and control significantly lower in importance than both providers and individuals with disabilities.

Next slide. You know what? Go ahead. Next. We'll go one more. At the subdomain level, we found some interesting differences in that persons with age-related disabilities rated personal choice as significantly more important to measure than the other disability groups.

And, again, these are based upon Z-scores. So, we're taking a look at kind of the average importance weightings that people assigned across the domains and subdomains, and then the extent to which their weightings and specific areas were significantly above and below that.

Next slide. Still staying within the domain of choice and control, you can see that when we discuss self-direction of persons with physical disabilities, so within that group, there was an interaction between the stakeholder group and the disability types. Persons with physical disabilities rated self-direction as relatively important.

Next slide. Providers for the IDD

group rated self-direction as below average importance to measure. And, again, this is directing your own -- excuse me -- directing your own service planning and service delivery.

Families of individuals with disabilities, on the other hand, rated self-direction as of average importance.

Next slide. In human and legal rights, you can see, as expected, you know, all groups rated them as important. But individuals with disabilities rated that area as significantly more important than provider members and family.

Next. The highlighted areas are just the areas that we're going to focus on. So, if we can go to the next slide, okay, you can see that, again, there were some differences in providers rating the optimization of legal and human rights as significantly less important than family members.

Next. Persons with IDD rated community inclusion as significantly more important to measure than all of the other groups.

Next. Next slide. And the subdomain of meaningful activity was rated significantly more important by persons with physical disabilities and intellectual and developmental disabilities compared to the aging population.

Again, just to remind the group that we're talking about how important people felt these were to include in measurement systems.

Next slide. So, in addition to collecting quantitative data, we also got a lot of input from individuals which we have partially analyzed, we certainly are not finished with that yet, that took a look at whether there were new domains or subdomains that the groups thought needed to be added to the NQF framework.

And we found that, you know, in the areas of community inclusion, choice and control, system performance and accountability, and holistic health and functioning, we had a number of groups -- and those numbers reflect the number of groups that suggested additional domains or subdomains among all of those groups that we -- that we had derived.

Next slide. Okay. So, again, I'm just giving you some quick examples. Recommendations for the domain of community inclusion included, you know, the need to ensure that measurement developers are focusing to a greater extent on diversity and cultural sensitivity, community outreach and education, and individuals with disabilities feeling welcomed and valued.

This is an area that we spent a lot of time discussing with our group who had very strong feelings about how community inclusion is currently measured in the most widely used measurement frameworks where it tends to be more focused on how often individuals get out into the community rather than their experiences within the community.

What we heard again and again from groups is, we really should be doing a better job focusing people being of the community than rather just being physically in it.

Next slide. Recommendations for choice and control indicated the need for individuals with disabilities to be more effectively supported and empowered and to have more choices available to them.

The group felt that the current framework really didn't place enough emphasis in either of these areas.

Next. This area was an area that we, when we originally focused on the NQF framework ourselves, thought was missing. And we were interested in seeing across the stakeholder groups, what groups felt that employment needed to be more attended to than it currently is.

Currently, as you know, in the domain of community inclusion, there is a subdomain called Meaningful Community Activity. And within that is employment.

We heard from a number of our groups, specifically persons with disabilities, family members and providers, of the need for the framework to include employment as a separate domain, because they really had concerns that it would not be effectively addressed by measure developers if it did not have inclusion at that level.

Next slide. So, that just gives you a quick overview of some of the things that we learned in our Study No. 1.

I think, you know, the major take-aways is that the work group that developed the HCBS NQF measurement framework did a good job in hitting on most of those areas.

We have a lot of information from our various stakeholder groups about how each of those subdomains might be further improved and we are in the process of putting together a technical report that we will share with everyone that will lay out in much more specific form, you know, each of those recommendations.

So, that study was a study that really needed to take place, we felt, before the development of our measures.

As in Study No. 2, which my colleague, Renata Ticha, will share with you, which is the gap analysis, taking a look at the NQF framework and its existing domains and subdomains, and existing instruments and measures that are being used to take a look at HCBS outcomes.

Renata.

DR. TICHA: Yes. So, Study 2 is basically an effort that started about a year ago. And we have a team of us, about five of us, working on that study.

And we are reviewing and cataloging all the instruments that we can find across the disability areas that already exist and mapping their areas or subscales and their items onto the NQF framework as our first step.

Next slide. So, some of the main questions for the Study 2, is the extent to which the measures are mapping onto the NQF framework and looking at the different characteristics of those instruments, which is -- and I'm looking at the following slide right now.

So, some of the characteristics include response options, respondent type, type of data, whether the items within the instrument are person-centered, and if the instruments have psychometric qualities, including reliability and validity.

So, we can look at the next slide. Some of the instruments that you can see there, these are just, really, examples, because we are close to coding a hundred instruments at this point, but you can see some of the really big ones that include National Core Indicators and the different types of surveys under that program, as well as some of the other ones that you can see that are big.

One of the one's that's a little different that you can see, number 19 is the PEONIES assessment developed in Wisconsin.

So, as you can see, we have a variety and different types of assessments that we have coded so far.

We can look at the next slide. How has -- the method that we have used apart from just cataloging the 95, and now it's about a hundred, instruments, is really code every single item within each of the instruments that has mapped onto the NQF framework.

So, we are in the 5,000s of coding the different items across the different characteristics that we have just gone over. And a lot of them have been assigned to the different codes that we are coding. Some of them that have not been assigned are just demographic questions or questions that don't directly map onto the NQF framework.

Next slide. This slide, in detail, lists the different variables that we have coded each of the items within each of the instruments.

So, apart from the NQF domains is subdomains. You can see I was mentioning the person-centeredness, the target population, the purpose, psychometrics and also coverage area. So, we are looking at very primarily those instruments are being used.

Next slide. So far, we have learned that of course there are some items in some instruments that are, more or less, covering the NQF framework. And I will show you some numbers in the following tables.

We have learned that a lot of the subdomains overlap. So, we end up double or sometimes triple coding some of the questions within instruments.

We are also learning that there are less questions or items that cover the systems-level performance and accountability versus the individual -- more individual-level domains and subdomains.

We have fewer items covering the caregiver and caregiver support subdomains. And also, not many items are covering consumer leadership and system development. So, those are some of the broader areas of coverage by the NQF framework.

If you look at the next slide, this is

information -- we are currently developing a database of these measures as they map onto the NQF framework.

And the purpose of the database really will be for people to be able to see what measures that already exist, but also the ones that will be developing and refining, how do they map onto the NQF framework as they exist, but also how it's being refined by the results of Study 1 that Brian has gone over.

I'll show you on the next slide an example of a dashboard that is kind of a precursor of the database that we are working on.

So, you can see in the table on the left-hand side, the coded items by domain. And so, you can see how many codes are for community inclusion, for example, for choice and control in relation to some of the ones that I highlighted that have not been covered as well that goes to equity, consumer leadership, and system performance and accountability.

And on the pie chart on the right-hand

side, you can see the percentage of coverage rather than just the raw numbers that you have in the table.

Moving on to the next slide, the next slide that's titled Instrument Heat Map, conveys similar type of information, but in a slightly different format.

So, you can see on the x-axis, you have the names of the instruments, the National Core Indicators, PEONIES and so on. And then on the left-hand side on the y-axis, you have the domains. And then you can see kind of the level of coverage of the NQF domains within each of the instruments. So, that's also available on our dashboard.

DR. ABERY: And then one thing which we aren't able to show you just because of size, is the fact that one can also look at that across each of the subdomains.

So, in some areas where there appear to be a lot of -- some domains where there appear to be a lot of questions, you know, those questions really aren't equally distributed across the subdomains. So, we may have many measures that focus on only one subdomain within a domain rather than adequately covering all of the subdomains.

DR. TICHA: Yes. So, on the next slide, that's really an interesting slide, because it does combine the data from Study 2 that Brian was describing, the groups, the PPDM groups, and also from the gap analysis.

So, you can see the coverage of items of the different domains, but also you can see how the groups across the country have rated them in importance and if there are any similarities and differences.

MEMBER POTTER: Hi. Could you give us a quick walk-through of how to interpret the colors on this, please.

DR. TICHA: Yes, absolutely. So, the darker the green color, the better coverage of items of the different domains.

So, for example, if you look at the

domain of community inclusion, we have not -- we have coded 972 items that in some way cover the domain of community inclusion.

It doesn't speak to the quality of the questions, but this is just a quantitative coverage.

In the same way in the same line, there is about 92 percent of the way that people saw this domain as important. So, it's not as important as, for example, choice and control, but choice and control only has about 620 items that cover -- which is still very large -- of that particular domain.

Does that help?

MEMBER POTTER: Yes. Thank you.

DR. TICHA: Okay.

CO-CHAIR MONSON: So, I had just a question around the instruments. So, you said there was like 94 instruments, but you've listed a bunch of them, all very good ones.

I didn't see any on there that were state-based assessments like the 701 B from

Florida or the NRI, which is utilized.

So, are those included as well and they just didn't make it onto these charts?

DR. TICHA: Yes. So, that's a really good point. We have included a number of the state-level assessments. And we would be very happy to share with you the, you know, the whole dashboard with all of the assessments.

But as you say, only the ones that have really large coverage are -- we coded -- some of the first ones were included on that particular slide.

MEMBER ROSS: Hi. I have a question. This is Clarke Ross with the Consortium for Citizens with Disabilities. Two questions. These are instruments currently coded.

When I joined this National Quality Forum group in 2012, I asked for an analysis of PEONIES and was told that only two counties in Wisconsin currently use it and that it didn't meet a meaningful threshold of implementation even though the domains were wonderful.

So, this is my first question of two: is -- do you have a threshold on use, and has PEONIES increased its utilization of implementation?

DR. ABERY: Okay. We did not have a specific threshold of use. One of the things that we wanted to do is to include instruments which were new.

We were also encouraged by our NIDILRR project officer and the individuals who we've met within ACL, to include instruments that had been developed in this area specifically for, for example, federally-funded research projects which while they might not be widely used, in some cases are the instruments who have the best psychometric data available on them to attest their reliability and validity.

We did include PEONIES within our analysis and have met with its developer, Sara Karon, on a number of occasions.

What has happened in Wisconsin, is that PEONIES has morphed into a similar type of instrument which is referred to as -- Renata, do you remember?

DR. TICHA: IRIS.

DR. ABERY: IRIS. You know, so that it is still being used in a sense; however, it is not being used as far as we have been able to track down outside of the state of Wisconsin. Although, it does include, we think, some very unique ways of taking a look at HCBS outcomes which are significantly more person-centered than some of the other instruments that -- many of the other instruments that we've reviewed.

MEMBER ROSS: So, I agree on the value and importance, but -- that's why I suggested it years ago.

But when you think about health plans and payers, something that's academically developed that's only been used in a couple counties and is no longer currently used is not really relevant.

It's a helpful academic conceptual thing, but it's -- the implementation experience is important.

And then my second question, I don't see the council and quality and leadership personal outcome measures on the currently-coded instruments. Do you plan to analyze that?

DR. TICHA: Actually, that's -- that has already been coded and analyzed. That's one of our main instruments that we have dealt with.

MEMBER ROSS: Thank you.

DR. TICHA: Yes. I just wanted to also add to Brian's response to the PEONIES.

So, we -- as we have completed -- or are completing Study 1 and are learning from the quantitative, but also the qualitative information what the stakeholders are seeing as critical areas, one of the things that has come up, for example, under community inclusion, is a social inclusion that wasn't included as a specific subdomain of the particular domain.

And instruments like PEONIES really have the capacity to address some of the more softer subdomains that are not as easily measured with some of the hard-core quantitative instruments. So, that's why we are still keen to include that instrument in our work as we move forward through the Center.

DR. ABERY: Okay. If there are no other questions, our next slide, please -- oh, and I should say before --

MEMBER ROSS: Well, this is Clarke again. I just -- PEONIES is a wonderful instrument. That's why I advocated it years ago here.

But if it's not used as an advocate, it's hard for me to convince the health plan folks around the table, and others -- I mean, it's a nice resource. It's well thought out. It has great domains.

You said it had good definitional properties, but -- that's my frustration is -- and I work closely with Joe Caldwell and I'll talk to him about this, but how much attention do we give to something that's not implemented almost anywhere.

DR. TICHA: Right. And, you know, so maybe we should go back to the beginning of our presentation a little bit.

So, the purpose of our center is not to promote, you know, some instruments over others or say -- sort of recommend an instrument over others, but it's really to pull items or measure concepts and measures eventually when we get there, that will be based on these instruments.

So, for example, if we are suggesting to measure social inclusion or we are suggesting to measure choice and control or decision-making, we are going to be recommending measures that will be sets of items that we are basing on these existing instruments and then filling gaps and refining them to recommend those.

So, I hope people understand what our mission is that it's not just recommend an instrument, it's recommending measures at the end of our cycle.

CO-CHAIR MONSON: So, this is Michael

Monson from Centene again. I think I understood all of that.

I think it's really important, though, that as you do that, you do that in the context that, first of all, some are proprietary instruments, some are not.

So, if you're going to take specific questions, for instance, from instrument A and say that that's the right question to get to, let's say, a choice and control measure, but instrument A is a proprietary measure and then you're taking question 3 from instrument B that's also a proprietary measure that gets you to another measure, we could end up with kind of a mash-up of different questions from different instruments and it will be extraordinarily difficult to implement.

So, I guess I would encourage you to think about how you use -- because where you end up will direct us back to what we end up doing in the real world.

And I think we look -- I know that

from a health plan perspective, we would be thrilled to have one standard assessment across all states for managed long-term services and supports, assuming it's a good assessment, you know, and assuming that we can always add additional questions to it.

But what would be really difficult would be that we end up with states saying, well, we want to measure this. And, therefore, you need to go buy this question from NRI and this question from PEONIES and this question from over here and having resistance from them because they won't parse it out that way.

So, I just -- I don't think you can -- I just would really encourage you to think about the end point as much as the journey you're on right now, because it needs to be something that's actionable for all of us so that we can actually get to where we all want to get to with some standard measures that we can all use to improve the system.

DR. TICHA: Yes. And thank you for

the question. It's a real important one. We are kind of in the thick of figuring that out right now, the proprietary material versus the one that is not. Also, versus -- we are looking at the questions that are covering certain measure concepts right now. And some of them are great, but some of them are not so good.

And so, we are really refining and developing additional questions for each of the -- each of the domains and subdomains -- well, not each, but the ones we are focusing on.

And the other piece of this is that we really have to pay attention to if -- when we were showing you some of the dashboard diagrams and some of the heat maps, if the domain or subdomain really doesn't have sufficient coverage and it doesn't have questions that a gap -- like Brian was giving the example of community inclusion domain didn't really capture what people are telling us in our groups, which is the engagement -- the active engagement of the person with a disability in the community.

So, we'll have to come up with questions that are not really included in any other instrument that, you know, be it -- I don't want to name specific instruments, but be it some of the ones that are the big program of assessment for the disability populations.

And so, we are really kind of in the thick of thinking this through and we certainly take your point into consideration very seriously.

CO-CHAIR MONSON: So, the only other thing I would add to that would be -- and I appreciate you taking it into consideration.

The only thing I'd add is as you think about adding questions, remember that when these assessments are conducted, they're already 90 to 120 minutes, sometimes longer, which is a burden on all participants.

And so, I know one of the NQF principles is parsimony. So, just bear that in mind, I guess I would ask you, about what's really important, right.

And, I don't know, maybe at the end of this you are going to have a ranked order of questions and some new standard assessment that you're proposing, which would be great assuming it's one that is not proprietary.

But just bear in mind the burden level that comes from a family caregiver or a participant having to participate in that. And it's a long -- it's a long session.

And then, also, obviously the cost that comes with that from either the fee-for-service side or the managed care side.

DR. ABERY: Those are all excellent points. One of the things that we probably should have indicated more clearly is we are working quite closely with the major measurement development programs, several of them, as part of this process.

We see ourselves, you know, not developing these -- any additional measures or refining measures in isolation, but rather working with those groups to help improve those measures.

I mean, we do not intend to produce the instrument, that was not the intent of our Center, but to improve measurement quality among all those groups that are, you know, currently using instruments which are widely used across the country.

MEMBER POTTER: Hi. This is D.E.B. Potter from ASPE. I guess I was a little confused by your last statement of not developing the instrument.

Because when I looked at one of your first slides, you say eventual objective is NQF endorsement. And the only way you can get NQF endorsement is, quote, to have an instrument.

DR. ABERY: Well, actually --

MEMBER POTTER: So, maybe you could expand that a little bit.

DR. ABERY: We will not be -- we will not be asking for NQF endorsement of an instrument.

We will be asking for NQF endorsement

of measures that correspond to the NQF subdomains in most cases, as opposed to attempting to develop an instrument such as the NCI, which covers multiple domains and multiple subdomains.

Our goal is to improve measures. And once those -- that measurement testing is done, you know, we do not see the measure that we developed there as proprietary. We are more than willing to share them with all of the measurement organizations that are currently being used by states.

So, I think we're basically differentiating between an instrument and measures that together would comprise an instrument.

MEMBER POTTER: I guess I would urge you to read the updated NQF requirements for person-reported outcome measures and then consult with the NQF staff on how that would happen given the current endorsement process for person-reported outcomes, which is what you're talking about.

MS. MUKHERJEE: And, D.E.B. and Brian, I can give a quick sort of update on that. So, NQF does not endorse a tool or a survey.

What they are doing is endorsing patient-reported outcome measures based on elements within the survey and within the tool.

So, in the past, when surveys first came out, NQF did endorse a survey as a measure, but right now because the field has evolved and moved on, if it's a survey or a tool, we do not endorse it as a measure.

What we will do is ask for the development of patient-reported outcome measures, which is taking elements of the tool or the survey and creating performance measures.

DR. TICHA: Yes. Thank you. And that's exactly what we are doing. And we do have a number of the members of the NQF committee on RRTC. So, we consult them on a monthly basis. So, but -- yeah, thank you for that point.

MEMBER ROSS: This is Clarke Ross, if I could ask one other question. You say you've been closely consulting with some of the measurement organizations, I assume National Quality Forum and Council on Quality and Leadership on personal outcome measures.

They've been reluctant to submit anything to the National Quality Forum, because their philosophy is you cannot pull out one measure from the mosaic of domains that they capture.

So, as an advocate, it would be helpful to pull out one measure and have it endorsed, but it runs counter to the philosophy of both NCI and CQL.

And so, I'm just trying to get my handle on the utility of your federally-funded project with all the experts saying, "Yeah, we'd like these four measures of 22 endorsed because of their importance and validity and all that stuff."

And so, I'm -- it's more a question -- I'm just trying to sort this out in my own mind on the utility of recommending unique measures that the developer themselves say need to be viewed in the total context of what they're trying to do.

DR. TICHA: Right. And, you know, I think we really appreciate that question and that's a question we have been thinking about from the time we wrote the proposal.

And I think, you know, Brian already indicated that we work very, very closely with the developers of the major assessment programs.

And, really, we see our charge as looking at not necessarily the domains/subdomains in isolation. But if we look at the psychometric properties of some of these measure concepts, they essentially have to be looked at in isolation to the point that they constitute a unique and differentiated concept or construct.

So, we have to confidently -- we have to be able to measure, for example, you know, social inclusion or choice and decision-making as its own entity regardless of how it relates to the others, for example, using factor analysis or using other statistical techniques.

So, I think they are not mutually exclusive and we just -- our charge, really, is to find the best possible questions to measure a measure concept or a construct across different instruments not to say that one is better than the other, it's just how do we conceptualize a measure concept in the best way possible.

CO-CHAIR MONSON: Yes. Go ahead.

MEMBER POTTER: Hi. This is D.E.B. Potter again. Additional clarification.

So, in the end, you might say -- this is just an example -- questions 3, 4 and 5 from survey A are the best way to measure community inclusion, but will you also say these other ways are valid?

I'm sort of back to the operational if we're going to measure community inclusion from survey A and we're going to measure employment from survey B and we're going to measure X from survey C, no one organization can administer all of those and how do you get it down to something that's operational without overburden, or am I not understanding your goal?

DR. TICHA: So, I think -- I think the way you're asking the question, you are understanding the goal well and I think there are two pieces to this question.

One is when we conceptualize a measure, it's essentially a theoretical concept or a theoretical construct such as social inclusion, choice and decision-making, and the questions that get asked are really secondary, right?

They are really -- the goal of asking the best possible questions is to be able to saturate a construct to ask the question in such a way that they are valid, that they really get at that subdomain or that measure concept.

And we can only determine that after we have conducted really sound, statistical analysis that include internal consistency, factor analyses and other analyses that differentiate a measure concept and --

MEMBER POTTER: I buy all of that, but where is question order effect in here?

DR. TICHA: Where is what again? Sorry.

MEMBER POTTER: Question order effect, i.e., in this survey, the questions are question 10, 11, 12, and in this survey, they're questions 24, 25 and 26.

DR. TICHA: The --

MEMBER POTTER: Question order effect is a well-known survey research construct that affects the validity and reliability of whatever you're trying to measure.

DR. TICHA: Yes. You know, at this point, we are just working with our data from Study 1, really, and then the information from Study 2.

Our next step is to -- we are about 10 days or two weeks away from identifying our measure concepts that we'll be piloting.

Once we have the measure concept identified, we are going to look at the, you know, 6,000 questions we have coded for across the hundred instruments. And we are going to map those onto those measure concepts.

And then once we have those, we are going to contact the developers of the instruments that these questions are potentially coming from, or the questions we are going to develop if they are missing. And then that's going to be our next step in taking it into the piloting stage.

So, we already are in very close contact with the programs that might have some other really good questions.

CO-CHAIR MONSON: So, you guys -- you all on the phone can't see us, I know, but you're seeing a lot of kind of furrowed brows not because -- everyone is really interested in this and very concerned that you get it right, right, because this is a really important endeavor. So, I think there are people in this room who probably have a lot of value to add.

And one thing I encourage you is to

maybe have another forum, because this is -- we're going to run out of time soon where, you know, people -- the experts in this room potentially who want to participate can provide you some more insight.

DR. TICHA: And we really value the feedback because, you know, we have a certain approach to this, the way we have written the grant.

We do have an advisory board that some of you -- actually, some of the people who ask questions are on our advisory board, but we would love to get as much feedback as possible especially at this critical stage when we are selecting our measure concepts for piloting.

DR. ABERY: Keep in mind that those first two studies really were to orient us to what is viewed by stakeholders as most important to measure, and what are the things that are not being measured right now.

There is a lot of measurement taking place which has little psychometrics to back it up at this point, you know. There are constructs that we say we're measuring and they would appear to the typical person to be measures where the necessary research and analysis has not been done to demonstrate that they are valid and that they are reliable.

And some of those are included in some of the most frequently used measures both nationally and within our own state of Minnesota.

CO-CHAIR MONSON: So, the one last thing I would say on this, the conversation heretofore has been really focused on patient-reported outcomes, which is hugely important, but those are not the only outcomes.

To hit all those domains, there are other measures and instruments that you'll need beyond patient-reported outcomes; claims-based data, observational data.

And so, I guess I would just encourage you all to think about the measure developers that you're working with are great measure developers for patient-reported outcomes, at least the ones you've mentioned like NCI, right?

They're fantastic, but they may not be the ones that are helpful in figuring out how do we best, you know, figure out the transitions for a nursing facility, net transition number, or how do we think about functional improvement for individuals who are receiving these services.

So, I would just leave that with you as well.

DR. TICHA: Yes. Yes. Thank you. And we would encourage all of you to, please, get in touch with us with any comments or suggestions.

We are also open to doing another forum like this. We have compiled a -- more of an interactive framework for feedback for the stage that we are at.

So, we would be -- our colleagues who work with us as project coordinators would be happy to schedule something like this to go into more depth. So, thank you again.

CO-CHAIR MONSON: Great. Was there

more that you guys -- we've kind of hijacked the conversation, but was there more material that you wanted to cover?

DR. ABERY: Well, there was quite a bit more material, but we don't realize how much time you have.

So, if you could kind of direct us as to how much additional time we might have, we can quickly go through kind of the meat of the RRTC's proposed work over the next four years or so.

CO-CHAIR MONSON: We've probably got another 20-25 minutes that we can keep going.

DR. ABERY: Okay. So, we can probably give you a pretty quick summary of what we are planning to do in Studies 3, 4, 5 and 6.

Study 3 is just starting. Study 4 and 5 will be starting with -- Study 4 in the next few months.

So, Study 3 basically is focusing -- if we can have movement to the next slide -- on identifying existing outcome measurement programs in which NQF-related HCBS outcome measures are being implemented and then to conduct a series of case studies to take a look at the quality of those measurement approaches and programs.

So, essentially, the idea behind this study is to take a look at the fidelity of administration and implementation of these measurement programs via the major players/organizations that are doing HCBS outcome measurement across the states.

Next slide. And essentially what we're going to be hoping to identify are the components that need to be in place to ensure that there's a high degree of administration fidelity, we're going to be identifying the strengths and challenges of the various approaches that measurement programs are currently taking, similarities and differences, and really looking at the factors that either facilitate or serve as barriers to effective implementation of the programs.

So, this is a qualitative study. Our only real qualitative study of the Research and

Training Center is being PI'd by Dr. Amy Hewitt from our center. And that is a study which is starting within the next few weeks with some work with the people from CQL.

So, we'll be looking at instrumentation, sample and recruitment, you know, across a number of different sites.

Study 4, and if we can move ahead with a couple of the slides and get to Study 4, is where after we select, based upon the gap analysis, those areas in which we feel that there is the greatest need for further measure development.

We will be working with several technical expert panels and with our colleagues from Temple University, UCSF, and The Ohio State University to either refine existing measures or to develop new measures and then to go through a development and testing process using the CMS criteria as we go through the process.

And I think there are a couple of slides which just look at the measure life cycle which had been incorporated into the process that we will be using.

So, if we can move ahead to Slide 56, which focuses on measure evaluation criteria, we will be looking at all of these in selecting the measures that we will be working to either newly develop or refine focusing on importance, feasibility, usability, harmonization, and of course scientific acceptability.

So, if you look at the next slide, that really is the focus of kind of the pilot study which is done in Study 4 which will be looking at inter-rater, test-retest and internal consistency, reliability and a number of aspects of validity.

Next slide, please. As part of this iterative process, we will be, again, prioritizing measures to develop based upon the combination of stakeholder input from Study 1.

Those areas that currently are not being measured or not being measured well where there isn't adequate saturation of the concepts will be our focus.

We will then be bringing together both among our research team members and technical expert panels, a group of individuals to work on this.

We are addressing the need for proxy reports and minimizing the burden to the individuals who we are requesting information from.

And then, again, we'll be looking at going through a process where we have kind of expert content review, cognitive testing and doing a relatively small, n=100 pilot study.

One of the things that we are really keen to do is the cognitive testing to ensure that since we are trying to develop measures that cut across groups, that the different groups are interpreting the questions that we are asking in order to get information, collect data so that we can have measures, or actually interpreting the questions in the same manner and in the manner that we intend.

That piece has really been absent from a lot of the work that has been done in this area.

Again, the next slide, if we move on to it, just gives you kind of an indication of what we'll be doing in the pilot study, identifying issues or concerns with administration and scoring, determining the acceptability of measures and the acceptability of measures to the recipients, obtain feedback on the response formats and wordings, and then determining the extent to which there's variability within -- between items and within measures.

Renata, you want to talk quickly about Study 5, which is our large-scale study?

DR. TICHA: Yeah. So, Study 5 is just the logical progression in this whole process toward NQF endorsement.

After in Study 4, we have piloted some of the really promising measure concepts based on the NQF framework. We will then work with those measure concepts a little bit more to refine them, to add some, perhaps.

And then in Study 5, it's our large-scale study, hopefully a national study, that we will then test those concepts on a thousand participants across all of the different visibility groups.

And so, we are currently in the process of recruiting sites. So, if there are any interested sites across the country who would like to work with us in the testing study that will start in about a year, we would very much be interested in that and will be looking at those psychometric properties that are listed on the slide; reliability, validity, item discrimination, sensitivity to change, but we will also cover some of the more basic criteria for feasibility, usability, importance and all of the CMS criteria that they have.

And we already will of course be working with some of the big programs like NCI, CQL, but we would love to work with any sites that are interested for that big study.

The data collection from Study 5 will be longitudinal to look at some of the sensitivity to change of some of those concepts and would like, as I said, would like to collaborate. And we do have some funds built into the grants for the collaboration.

And just briefly and in the rest of the time that we have, Study 6 is already underway. It is a study that has been reviewing different risk adjuster variables from studies across the different populations, identifying the most promising risk adjusters across the NQF outcomes.

And we will be including some of the risk adjusters that we have identified over the last six months or so into the pilot study, but also into Study 5 to test alongside the measures of the outcome so we can look at potentially also at some of the relationships.

And, again, if you have any ideas or any input into risk adjusters, we would be happy to share more detail about the study. We are currently working on a manuscript of the Study 6 results. So, that's where we are with that.

And let me just -- I'm sorry, I'm skipping a little bit here, but on one of the slide that's titled "Coding Progress," you can see that as of about ten days ago, we have been able to code almost a thousand different variables from about 60 studies. And you can see the breakdown by systems and by individual levels.

And of those, we have identified about 42 promising risk adjusters that will be then narrowing down for the testing phases.

And on the following slide, there are some examples of the categories of risk adjusters that cover functional disability, chronic conditions and risky behavior. So, those are the ones that are appearing more frequently.

MEMBER ROSS: Excuse me. This is Clarke Ross again. So, a lot of folks in the health policy arena and the National Quality

Forum have been trying to get away from a diagnosis as a risk category, because, you know, there are hundreds of chronic illnesses that are debilitating -- potentially debilitating.

So, how are you building diabetes, cancer and epilepsy and all the forms of cancer as a risk adjuster?

DR. TICHA: Yeah. So, maybe I should just give a little bit of a context. So, this initial stage of our Study 6 on risk adjusters has been very broad.

We have cast a very, very broad net to really provide interested audiences with what has been done so far in the field of risk adjusters with these populations.

So, some of the ones we have got on the slide here that you are referring to, are just to say that they have been studied and included in the past, but it's not necessarily something they'll be considering to include in relation to the NQF framework.

CO-CHAIR MONSON: So, this is Michael

Monson. Again, I would just say, also, on the functional disability, that is going to be very dependent on having a standard set of observational questions.

DR. TICHA: Uh-huh.

CO-CHAIR MONSON: I would encourage you to look at the work that's happened in New York and Wisconsin around functional-based risk assessment for rate setting on the managed LTSS side, because they've actually done a lot of good work there.

DR. TICHA: Uh-huh.

CO-CHAIR MONSON: But it is going to -- that -- if you're going to use that -- and it should be used as a risk adjuster, but it will require us to have standard elements across every assessment to be able to do that.

DR. TICHA: Yeah. You're absolutely correct. And if you have a specific reference that we should look at, please share it with us.

CO-CHAIR MONSON: The -- we will.

DR. TICHA: Okay. Thank you.

Yeah. So, again, so these -- the ones that we put on the slide are just some of the ones we have seen as reoccurring across the studies that we have reviewed. And that goes back to the current take-aways.

We did find that most of the risk adjusters across the studies were not well-matched to the HCBS outcomes. And there's a strong tendency to focus on the personal characteristics of the recipients. And there are many newer risk adjusters that have been studied well that are at a systems and the organizational level.

And so, we'll be making a big effort to get input from our -- from stakeholders from our advisory board and our leadership team on which variables to actually include in our pilot in our big study for risk adjustment.

So, our next steps will be to prioritize risk adjusters in a similar way we are doing with the measures of the outcomes based on CMS criteria; importance, feasibility and usability and looking at the magnitude of effects in the previous studies. And then linking the risk adjusters to the outcomes that we actually will be testing and piloting so that we don't have risk adjusters that don't relate well to the outcomes. And we are kind of parsimonious about examining all those variables.

But, you know, as has become clear through this webinar, but from your comments it's critical that we engage as many stakeholders as possible, and we very much appreciate your input into this work so that we make it as applicable and relevant to all the groups that we need to work with.

DR. ABERY: So, if there is anybody out there in the audience who would be interested in participating in the technical assistance panels that we will be putting together over the next, well, several months and years, you know, you can contact either Renata or me or Amy Hewitt, who serves as co-director, or we can work through our colleagues at NIDILRR, Amanda

Reichard is our project officer, to make those connections, because we do want as much community input as Renata has suggested, as possible.

CO-CHAIR MONSON: Great. And NQF staff will send out their email addresses to everybody so that we can be in touch with them.

Yes, D.E.B.

MEMBER POTTER: Hi. This is D.E.B. again. Thank you so much for this presentation. This is the third time I've heard variations on this presentation, and every time it's richer and I have a better understanding of all of the moving parts.

One thing I still don't have a good handle on is what is the accountable entity? Who are you holding accountable for the performance of the quality metric?

So, there's measures that hold health plans accountable, measures that hold hospitals accountable, measures that hold programs or states accountable.

If you could speak a little bit to the

accountable entity, because as you pointed out in one of your slides, one of the criterias of importance is a performance gap. And the performance gap is between the accountable entities.

So, if all the accountable entities measure at 99 percent or 70 percent, it's not differentiating and helping you to improve.

So, if you could speak to that, that would be helpful.

DR. TICHA: Yes. Thank you very much for the question. I think there are a couple pieces to that.

So, when we were charged with the Center, one of the initial -- one of the major, initial pieces was a focus on developing measures that were psychometrically sound and important.

And so, we -- our first -- we have not even reached our second year of the Center. We have been really focused on developing measure concepts that will stand the test of psychometric rigor.

Then after we have established some of that, then I think we are going to more closely look at the differing accountability frameworks.

You alluded to the high percentages of agreement between the different NQF domain and the ratings. I think that is -- that is important to consider.

However, we -- in addition to the quantitative data that you saw up there, we have a very, very rich data set of qualitative data that provides really good information on differentiation among those domains and subdomains, and also that introduce additional subdomains and domains to the framework.

And so, I know I'm not quite answering your question, but I think we are sort of in early stages of the Center and that we'll provide initial information within the next, I would say, three to six months.

DR. ABERY: Also, to build off what Renata was saying, you know, in the FAL that was issued by NIDILRR, it was made clear that the measurement work that the Center was to do, you know, was to kind of cut across a variety of different entities that are responsible for the quality of home and community-based services ranging from provider organizations to states.

And so, we're looking at our work in that way and currently we're doing a fairly large amount of work with the state of Minnesota in helping them to improve the quality of their measurement in an area working with three different regions of the state and have come up with over in about a year period of time, what we believe is an approach that will get at, you know, both the standard -- based upon the standard questions, survey questions asked of individuals, but also will allow the state to dive deeper into the extent to which the supports people are being provided with are truly person-centered and the degree to which they are actually able to achieve some of the life outcomes that they desire.

So, on responsible entities, it is

going to be -- or our goal is to cut across the different entities. And, again, we are attempting to do this in a way which is consistent with the NQF framework, but also is going to build upon it.

CO-CHAIR MONSON: Well, thank you all very much for sharing that with us today. As you can see, this is a group that's very interested in this topic and is -- I know many individuals in this room will be very eager to continue to be engaged with you outside of the role here on the MAP Committee.

It's exciting work that you're doing, it's important work. And so, we thank you for what you're doing and for taking the time to speak with us today.

DR. ABERY: Thank you very much.

DR. TICHA: Thank you.

CO-CHAIR MONSON: All right. And so, we're going to take, I'm going to say, a ten-minute break, because we're running a little behind schedule. So, back here at 2:35 Eastern.

(Whereupon, the above-entitled matter went off the record at 2:25 p.m. and resumed at 2:35 p.m.)

MS. BUCHANAN: Hi, all. Thank you very much and welcome back. So, we wanted to make a quick announcement.

We originally had on the schedule, voting for inclusion both in the family and the starter set on the CAHPS measure. What we are going to do is we are going to send that via SurveyMonkey tomorrow so people can vote on it, because we don't want to lose quorum and we do want to hear this presentation first. So, just an FYI to keep a lookout for us asking for a vote.

And other than that, that we'll have just the --

MEMBER POTTER: How will we find out the results?

MS. BUCHANAN: Yes. So, they will be emailed through communication. And then, so, if we are going to take away the voting for the measure, we also will be talking, though, about the strategic direction for the duals population. And then we'll open it up for member and public comment. And then we'll adjourn for the day.

But prior to that, we are very lucky to have Dr. Kerry Lida, who is the team lead for the testing experience and functional tools demonstration within CMS, and the TEFT demonstration is within the Disabled and Elderly Health Programs Group, Division of Community Systems Transformation.

She is also joined by Elizabeth Frentzel, who is a principal research scientist at the American Institutes for Research, with almost 20 years of experience in developing CAHPS surveys, including the HCBS CAHPS, Medicaid CAHPS, home health CAHPS, nursing home family CAHPS and cancer CAHPS pilot version.

We also have Coretta Mallery, who is the principal research scientist at AIR and have over five years of experience doing CAHPS survey data analysis. She was the analytic lead for the

HCBS CAHPS survey and related measures.

And then last, but not least, we have Susan Raetzman, who is the director at Truven Health Analytics and the current TEFT TA lead for the Experience of Care component under which the HCBS CAHPS survey and related measures were developed.

And with that, I will turn it over to you all. Just let us know when you want us to move the slides. Thank you.

DR. LIDA: Wonderful. Thank you for inviting us to talk with you today about the CAHPS Home and Community-Based Services Survey and related measures that were endorsed by NQF.

My name is Kerry Lida and I am delighted to be sharing this area of work with you and very happy that we have expert colleagues on the line with us, including Susan Raetzman, Elizabeth Frentzel and Coretta Mallery.

If you could advance to the next slide, please. Today we'd like to cover the following topics; development of the survey, including the need for a CAHPS survey for home and community-based services, HCBS; key features of the HCBS CAHPS survey; National Quality Forum-endorsed measures derived from the survey; state use of the survey and some of the resources that are available for use.

Next slide, please. And briefly in the next few slides, we'll be giving an overview -- a very brief overview of the development of the survey. And then we will walk through some examples which we believe will be of interest to this meeting today.

Next slide, please. Before we do that, let's talk about why a CAHPS survey focusing on the HCBS setting is important.

Historically, Medicaid systems for people who needed long-term services and support, LTSS, had an institutional bias in that services overwhelmingly were provided in institutional settings such as nursing facilities, long-term care hospitals and intermediate care facilities for persons with intellectual and developmental disabilities.

However, the percentage of Medicaid LTSS expenditures for HCBS has grown since the early 1980s.

By the late 1980s -- 1990s, HCBS accounted for more than 25 percent of the 70 billion spent on Medicaid long-term services and supports.

And by fiscal year 2014 for the first time, over half of Medicaid's LTSS dollars, 53 percent, was spent on community-based supports.

With such a large portion of LTSS provided in the community, it is imperative that there be mechanisms for ensuring the quality of care delivered in those settings.

Next slide, please. The entire development and testing process beginning in 2010, was funded by CMS under two projects. the most recent is the demonstration grant for testing, experience and functional assessment tools, TEFT, in community-based long-term services and supports. Prior to that, it was the

National Quality Enterprise, the NQE grant.

There are four test components. And each component has different activities and time lines which you will see in front of you.

In the upper left quadrant is the Experience of Care Survey component under which test grantees and researchers from Truven Health and the American Institutes for Research test and survey, which contributed substantially to submissions to CAHPS and NQF.

The other three quadrants we will not be discussing, but they are available on medicaid.gov if you are interested in further -- and those include the functional assessment standardized items, the eLTSS plan standard in collaboration with ONC and the Personal Health Record.

Next slide, please. Test grantees participated in the pilot and field tests that were conducted to test for reliability and the validity of the instrument and measures derived from it.

This table shows that beneficiaries from up to four HCBS programs in each of the ten test states were surveyed in the pilot and field tests.

Individuals and Medicaid HCBS programs in Louisiana and Tennessee participated in the pilot test between October 2013 and October -- April 2014.

Field testing took place in Louisiana, New Hampshire, Connecticut, Maryland, Minnesota, Colorado, Arizona, Kentucky and Georgia between July 2014 and February 2015.

Next slide, please. As noted earlier, the HCBS CAHPS survey felt a critical need in LTSS quality assurance because it focuses on Medicaid, HCBS beneficiary-experienced outcomes and quality of life as the result of receiving services and supports.

What distinguishes it from other HCBS surveys is that it was designed to be completed by the broad range of beneficiaries through the Medicaid HCBS programs, including individuals who are frail/elderly, individuals with a physical disability, individuals with an intellectual or developmental disability, individuals with brain injury and individuals with serious mental illness.

Next slide, please. This is a very brief summary of approximately six years of work and we have a team who will be available to answer questions you have in any of these areas.

The development team follows the CAHPS survey development process, a rigorous and beneficiary-involved process. It is also well-known and highly regarded by developers of experience of care surveys.

This diagram outlines the three major phases of the development process. The process of developing and testing the survey instruments began at the formative research stage to identify key domains and constructs. Then the team drafted the survey.

In Phase 2, survey development continued with multiple iterations of cognitive testing of the draft survey instrument in both English and Spanish.

Once the items were refined, the survey was pilot and field tested on over 3200 individuals in 10 states and 26 HCBS programs using both fee-for-service and managed long-term services and supports.

In Phase 3, pilot and field test data were analyzed to determine which items worked or did not work allowing the survey to be finalized.

The survey instrument receives the CAHPS trademark in June 2016. NQF endorsement of the related measures occurred in October 2016. Survey and measure maintenance activities continue from those points in time.

And it is -- as noted in the beginning of this slide, we mentioned the high level of integration and input that we sought from individuals who are participating in this. And it's very unique and it's a very critical survey in this regard and we're happy to speak more about this.

Let me hand the presentation now off to Elizabeth Frentzel from AIR, who will continue. Thank you.

DR. ROILAND: Operator, is Elizabeth Frentzel on the line? Operator?

THE OPERATOR: I do not see her connect-in.

DR. ROILAND: All right. Apologies, Kerry. We seem to have lost Elizabeth. Just one moment, please.

DR. LIDA: Okay.

(Pause.)

DR. ROILAND: Elizabeth, if you are logged into the webinar platform - we did have you noted as calling in earlier, but you may have dropped off accidentally. If you could call back in, that would be great.

(No response.)

DR. ROILAND: Kerry, is there another section of the presentation we could -- I know this is probably an important part of the presentation.

Do you just want us to wait for Elizabeth to try to get her back on?

(No response.)

DR. ROILAND: Kerry?

(No response.)

DR. ROILAND: Operator, are you still there?

(Laughter.)

CO-CHAIR MONSON: I'm waiting for someone to say --

DR. ROILAND: Hello? All right. Just stand by, please. We're having some technical difficulties.

(Pause.)

DR. ROILAND: All right. Operator, can you hear me now?

(No response.)

DR. ROILAND: Operator, we can't hear you, if you can hear me, just so you know.

(Pause.)

DR. ROILAND: Hi, all. I don't know if you can hear us, but we are having technical difficulties. So, just, if you give us one minute, we will hopefully get this remedied.

MS. JUNG: This is just the time we would have waited for the, you know, clickers to work. So, we're just moving it around.

DR. ROILAND: Kerry, this is Rachel with NQF, we can hear you now. I apologize, I don't know what happened there for a moment.

Just give us one second while we get reset here.

MS. FRENTZEL: Great, thank you.

DR. ROILAND: All right, Elizabeth, are you on now?

MS. FRENTZEL: Yes, I am.

DR. ROILAND: All right, thank you. I really apologize about that, but the floor is yours now. And, you can just let us know when you want to advance the slides by just letting us know next slide.

MS. FRENTZEL: Thank you.

So, next slide?

So, the survey is intended to, in

reports about a particular HCBS program's performance, these are the beneficiary-reported experiences.

The unit of analysis is either the HCBS program or the accountable entity.

An accountable entity is the operating entity responsible for managing and overseeing a specific HCBS program within a given state. For example, a managed care organization.

The HCBS CAHPS survey was developed so that the comparisons about the quality of services and support can be made across programs or between managed care organizations or other subgroups.

Next slide, please?

The development process identified the home and community based services and supports and providers that would be appropriate for beneficiary input across the disability and HCBS program spectrum.

The services and providers listed on this slide are those that are common across

Medicaid HCBS programs with one exception. Although employment assistance services are not offered across all programs, the TEP, or Technical Expert Panel, encouraged the inclusion of items on these services because they are so vitally important for full community participation, especially for working aged-adults served in HCBS programs.

Next slide, please?

As a result of the pilot and field test analyses, as well as feedback from the TEP -- I'm sorry, I think this was -- is this the slide?

So, as a result of the pilot and field test analyses as well as feedback from the TEP and CAHPS consortium, the finalized survey includes the items and measures depicted on this slide.

First is a set of three cognitive screening items that help identify individuals who may or may not be able to provide reliable information.

These individuals may assent to having a proxy respond to the survey.

Next, survey identification items and screening items ensure that the beneficiary answers only questions about the services that they receive.

Because not every beneficiary answers all questions, the average survey administration time is 30 minutes.

There are 34 items that make up composite measures that will be described later in this presentation.

Stakeholders deemed some single item measures as important to retain even though they did not fit into a composite. These include a series of questions that assess a person's unmet needs and physical safety.

Six global rating and recommendation items provide information on the person's overall experience with the three main types of staff.

There are 15 items that collect demographic information, some of which are used for case mix adjustment.

And, finally, a 21 item supplementary employment module is an option for programs that provide employment services.

Next slide, please?

Now we provide more detailed information about the survey and beneficiaries with cognitive challenges.

The survey was designed including question wording and response sets to be accessible to as many HCBS beneficiaries as possible.

However, it's also important that those using the results of the survey have confidence in those results.

And, for this reason, the survey starts with a three cognitive screening questions shown on this slide. If all three questions are answered in a meaningful way, the interviewer continues to administer the remainder of the survey.

If the three questions are not

answered appropriately, it is an indication to stop the interview and inquire about a potential proxy respondent.

Next slide, please?

Here, we provide more information about the survey's incorporation of program and provider-specific terms.

On the basis of the formative research, we knew that there were very few uniform naming conventions for providers across programs in terms that individuals used in referring to their providers.

Thus, the survey was designed so that sponsors can incorporate program-specific terms for categories of staff and provider-specific terms for individual staff.

The preferred terms can be used throughout the survey. You can see the bracketed, italicized text that alerts the person administering the survey to administer -- to insert program-specific term for these types of staff.

Next slide, please?

Now we provide information about another survey feature aimed at increasing beneficiary participation. On the basis of findings from the cognitive testing as well as an experiment conducted as part of the field test, a simplified response option was determined to be accessible for some individuals.

Using both response modes allows for more people to participate in the survey, including individuals with intellectual or developmental disabilities.

This slide shows the alternate response approach in the survey itself. The interviewer starts with a standard CAHPS response option of never, sometimes, usually or always.

If the respondent has difficulty using that question and response format to answer, the interviewer then asks the alternate version.

The interviewer does this three times, and if the respondent prefers the alternate version, the interviewer then uses only the alternate version for the rest of the survey.

Next slide, please?

The HCBS CAHPS survey asks about several categories of HCBS services, some of which respondents do not receive.

To help interviewers ask only the relevant questions, the instrument was developed with skip patterns imbedded throughout.

The survey also was developed with addition skips related to screener questions. This helped to ensure that specific experiences could be identified on which programs would be able to act.

In the examples shown on this slide, there is a set of questions focused on whether the individual goes without help in bathing or getting dressed because personal assistance staff are not there to help.

The first screener question in the series asks whether the person needs help from the personal assistance staff to bathe, shower or get dressed. If the person says no, in other words, that he or she has no need for help, the interviewer is instructed to skip the next three items and move on to an item on personal privacy.

If the interviewer -- if the beneficiary says yes to the screener question about needing help, then up to two additional follow up questions are asked to elicit information on an unmet need.

Next slide, please?

Because survey response rates are an important issue, we want to share with you the pilot and field test results.

The overall response rate was about 22 percent with the highest response rate, over 30 percent among participants of programs serving individuals who are frail elderly and individuals with a physical disability.

The lowest response rate, about 10 percent, was among participants of programs serving individuals with an intellectual or developmental disability.

In addition, in the pilot and field

tests, participants of different programs preferred different modes.

Next slide, please?

Although the HCBS CAHPS survey was intentionally designed to be as accessible as possible, survey vendors indicated that guardians could act as gatekeepers by refusing on behalf of the beneficiary or wanting to be their proxy.

Survey vendors also reported that other individuals in the person's life were willing to respond to the survey as proxy respondents.

Thus, part way through the field test, we began allowing proxy responses.

For the purposes of the pilot and the field tests, a proxy respondent refers to any help that the respondent received in completing the survey.

This ranged from answering all questions for the respondent to providing prompts, translation or help with communication technology.

This table presents numbers and percentages of proxies for each HCBS population represented in the pilot and field tests.

As you can see, across the board, proxies were used to some extent and there was substantial variation and use by population and by state.

These results are likely an underestimate of proxy respondent participation had they been allowed throughout the entire data collection period.

Next slide, please?

Going forward, proxy respondents are being allowed by CMS for administration of the HCBS CAHPS survey. This flexibility has a few implications for survey sponsors.

For example, it's up the sponsoring entity to decide whether to use proxies and which proxies to include.

There are certain qualities that make an individual more likely to be a good proxy respondent.

If a proxy respondent is being used, the IRB may require that the assent of the beneficiary as well as the consent of the proxy be obtained and documented.

If proxies are used sponsoring entities introductory script for reaching out to the beneficiaries will need to allow for talking with proxies. This script should reflect decisions that the sponsor makes about which proxies to include.

While fielding the survey, sponsors may want to monitor the percentage of surveys that are completed by the proxies.

Finally, the data analyses should adjust for the use of proxies.

Next slide, please?

Once an interview has been completed, there are a few interviewer questions that ask about proxy respondents. The questions shown on this slide are used in the HCBS CAHPS survey to distinguish HCBS (a) beneficiaries who were not helped in completing the survey from either beneficiaries who received assistance from another person in completing the survey; or, (b) beneficiaries whose survey was completed by someone responding on their behalf.

And, now, this goes over to Coretta.

DR. MALLERY: Great, thanks, Elizabeth.

Next slide, please?

So, I'm going to give you a brief overview of the NQF endorsed measures derived from the HCBS CAHPS survey.

So, next slide, please?

Okay, so the HCBS CAHPS survey items provide information about specific domains of the HCBS experience.

The HCBS CAHPS survey has 34 items that support seven scale measures and 12 items that support single item measures from the domains shown here.

So, this slide gives a nice overview of our scale measures and the different domains.

After the composites were identified

in the analyses, the developer team went back to a group of beneficiaries and talked with them about the best labels of each of the composites.

So, this slide reflects these labels for the final set of composites.

The next several slides describe each scale measure and the survey items that make up each of them as well as the single-item measures. And, each of these are NQF endorsed measures.

Okay, so, this slide shows the first composite, staff are reliable and helpful. There are six items that pertain to the personal care attendant, behavioral health staff and/or the homemaker.

These include whether staff come to work on time, whether staff work is done when they're supposed to and how long staff makes sure beneficiary has enough privacy when bathing and dressing.

Next slide, please?

Okay, so the next composite, staff listen and communicate well has 11 items.

These items focus on aspects of service and support such as how often staff show courtesy and respect and how often staff listen carefully to the beneficiary.

And, I'm just giving a brief overview because you all can -- will have access to the slides and can see each item in each measure.

Next slide, please?

Okay, so the next measures, case manager is helpful, so, this slide shows the first three items in this composite, addresses things such as whether the case manager works with beneficiary when he or she asks for help with getting changes to services and a few other constructs.

Next slide, please?

Okay, so, this slide shows that the composite, choosing the services that matter to you consist of items on whether the person's service plan includes things that are important to the beneficiary, whether the staff know what's in the person's service plan, including the things that are important to the beneficiary.

Next slide, please?

This slide shows that transportation to medical appointments composite consist of items about how often the beneficiary has transportation to medical appointments, whether the beneficiary is able to get in and out of the ride easily and how often this ride arrives on time.

Next slide, please?

So, there are three items in the personal safety and respect composite that ask the beneficiary about whether there is a person with whom they can talk if someone hurts them or does something to them that they do not like, whether any staff takes the beneficiary's money or their possessions without first asking, whether they have staff that yell, swear or curse at them.

Next slide, please?

Okay, so, this slide shows the composite, planning your time and activity. And, this is a community integration measure that asks about how often the beneficiary gets together with family members or friends who live nearby if they want to do so, how often the beneficiary does things in the community that he or she likes when he or she wants to do and similar constructs.

Next slide, please?

Okay, so, in addition to the seven composite measures, there are global ratings, measures and three recommendation measures that are supported by the HCBS CAHPS survey.

Both the global ratings and recommendations are specific to three different types of service providers shown here, so personal assistant, behavioral health staff, homemakers and case managers.

Next slide, please?

So, in our psychometric analyses, there were 13 items that were not part of a composite that the Technical Expert Panel felt that should -- that were important enough to be retained in the survey as standalone items that were then, we put forward a standalone measures.

This slide shows a subset of these that were submitted to NQF and endorsed. So, these include measures that address unmet needs and physical safety.

Okay, now, Susan Raetzman will be talking about potential uses of the survey.

MS. RAETZMAN: Thank you, Coretta.

Next slide, please?

MEMBER ZLOTNIK: Yes, can I just ask a question before you go on?

Related to the global ratings related to workforce issues, I had a question about survey item 35, global rating of personal assistants and behavior health staff.

Those two things sound very different to me, unless you are just talking specifically about kind of more like residential staff that might be in a behavioral health setting.

The behavioral staff could be everything from a psychiatrist to someone who is a group home staff person.

So, I don't -- that's a very big area. So, is there a definition of it?

DR. MALLERY: Yes, so, Elizabeth, I'm not sure if there might be --

MS. FRENTZEL: Yes, I can answer it.

It's the behavioral health staff and these are the people who are providing services in home. So, they might be, for example, queuing an individual to help them know the next behavior. It's not -- it would not include psychiatrists or anyone highly clinical.

MEMBER PARKER: But, how would the -- hi, this is Pam Parker on the Committee -- how would the person know what you're talking about?

MS. FRENTZEL: In the initial -- there's an initial set of questions on providers and they provide definitions there.

MEMBER PARKER: Well, I'd just say, within my experience, I had the same question about case manager, because there are multiple levels of case management in these programs and

I'm not sure which one the person thinks we would be talking about.

And, we've used, actually, in Minnesota, in the CAHPS at the state level for other programs, the part of the Medicaid program for the big CAHPS that's done, we've used questions on care coordination and case management like type questions and we know that there's a lot of fuzziness around that.

We've done focus groups on that. And, we know that people define case managers as everybody from their family friend to their doctor to everybody in between.

And, so, I'm just saying that, you know, just because you put in the survey, maybe not clear to the member that's looking at the survey, even if you explain it to them.

MS. FRENTZEL: Absolutely.

MEMBER PARKER: Yes.

MS. FRENTZEL: And, that's what we found in the focus groups and that's why, you know, we've -- after the cognitive screening questions, there are all these -- there's several program questions to really narrow it down because there's such variation.

Now, with case managers, it was more of a titling issue, not so much kind of the work they do. So, we have a definition for case manager for personal health staff or behavioral staff and homemaker staff and, so, that, because of the variation.

But, you're right. And, if they could have potentially two case managers --

MEMBER ZLOTNIK: Some have 23 case managers.

MS. FRENTZEL: A single person could have 23 case managers?

DR. JACKSON: This is Beth Jackson, I'm also on the -- I was on the development team.

And, because the accountable entity is typically a program, the sponsor of the survey would be able to say this is what we call a case manager in this program.

And, you know, typically, my

experience has been that a person has one case manager in an HCBS program.

Granted, that an individual could have more than one case manager across programs and be enrolled in more than one program, but I think it would be very clear to the individual what program they were being surveyed about.

And that, in fact, it's even possible to customize the survey in such a way that actually talking about, you know, your case manager and the person's name. So, I think that helps.

And, in terms of the behavioral health staff, it's really behavioral health staff coming into the home to help with ADLs and IADLs.

I don't know if that helps answer your questions or not.

MEMBER ZLOTNIK: I mean, I think it does by using the name because I think I've worked with a lot of people who really can't keep track of who belongs to what program who are coming into their house or what they're role is.

CO-CHAIR MONSON: Well, and, actually, even within a program like, in Ohio, for instance, if you're in the MMP and you're over 65, you have two case managers. You have a health plan case manager and you have a AAA case manager.

Same thing in Kansas, if you are in the LTSS program there, you've got if you're an individual with IBD, you've got a targeted case manager from the county and you've got a health plan case manager. So, you're definitely going to have two.

MEMBER ZLOTNIK: Yes, and if I could just say, where are the term case manager is often also used is at the provider level.

So, if you have a personal care provider and another kind of behavioral health provider and, you know, several other kinds of service providers, that's where you might find a home health like skilled nursing, you'll find a case manager.

And, then you've got, in addition, you

might have a home community based case manager and an MMP or a D-SNP case manager or usually called care coordinator or something else, but that's where --

But, I will say that I think your idea of trying to get the name of the person is a good one. We did do that in Minnesota in our early testing of not CAHPS survey, but another consumer survey that we were trying and that did help.

But, it was very difficult to keep track of the name because, depending on which one you're talking about, because you've got to have that bin ahead of time in the survey. And, if you're administering the survey by mail or something, you know, that can change by the time you get it printed up.

I don't even know how you do it exactly, you know, it takes manual work to do it then. So, it's difficult.

CO-CHAIR MONSON: Is it relevant to this particular topic, Beth? Because Tom has been waiting.

MEMBER POTTER: Yes, Beth, maybe you could day, this is D.E.B. Potter, a little bit about how this survey is administered using a computer-assisted application and all of that. That might be helpful.

DR. JACKSON: Yes, yes, it was designed to be face to face or by phone. So, it's designed to be CATI, computer-assisted telephone interviewing, or CAPI, computer-assisted program interviewing.

So, it's not -- given the complicated skip patterns and given the desire to tailor it to the program and the individual as much as possible, doing that with a male survey would be virtually impossible.

We were also counseled early on by members of our TEP to have this be actually an in person. They felt with the HCBS population, particularly with so much cognitive impairment as well as hearing impairments that it would be important to do this survey face to face.

CAHPS does require at least two modes

of administration and we did add the telephone interviewing, which we found, and which states are finding now, at least, well, I'll say at least one state is finding in the test program that individuals in these programs are -- in some of the programs are preferring the telephone interview as opposed to being visited.

So, it's CAPI or CATI, CATI or CAPI is highly recommended and not telephone, I'm sorry, not mail. Does that do it?

MS. FRENTZEL: Yes, it did.

CO-CHAIR MONSON: So, why don't we go to Tom? We've got about 13 minutes left for this topic. So, we'll go to Tom's question then we'll keep going.

MEMBER LUTZOW: Yes, this Tom Lutzow, iCare, Wisconsin.

Self-directed support is a big thing in Wisconsin. We're a FIDE SNP, 30 percent and it's growing. Maybe this year, it will reach 40 percent of the members of self-direct some of their home and community-based waiver services, especially personal care, homemaking and support of home care kinds of things.

Do you make note of that? It's a co-employment situation, there's a fiscal agent that is the employer of record but pretty much, the member picks their own worker.

And, even sometimes, when we might otherwise prefer they pick someone else, it's a relative, maybe even someone who is, you know, has a court record that would otherwise not make the cut. But, because it's a matter of preference, they're selected and hired.

Where is that accommodated here?

MS. RAETZMAN: So --

MS. FRENTZEL: Go ahead, Susan.

MS. RAETZMAN: This is Susan Raetzman.

The survey accommodates it by not actually making a distinction. The questions can be asked for a person who is, you know, providing services under that arrangement as well.

If the survey sponsor wants to be able to identify those surveys when they're doing the analysis, they could add that information to their analytic file so that they could look at that separately.

CO-CHAIR MONSON: Okay.

Do you have more of your presentation?

MS. RAETZMAN: Yes, we have a few more slides.

CO-CHAIR MONSON: Okay.

MS. RAETZMAN: And, I'll go fast.

So, could you go to the next slide, please?

So, this is just a brief review of some of the aspects of the survey that states are encouraged to consider in their thinking about, you know, whether the CAHPS -- HCBS CAHPS survey, you know, fits into their needs.

So, it's the person centeredness, the fact that you can use the survey with a broad range of individuals with disabilities and do that with a single instrument, you know, across these different programs.

It's designed to be as accessible as

possible to all beneficiaries through a variety of design features that we've talked about.

It aligns with CAHPS. It supports several quality metrics such as the NQF ones and there's an opportunity for users to also add their own specific questions that they're interested in or as a supplemental employment module.

And, because it is provided by CMS to the public, there's no cost to access the instrument or the supporting documents.

Next slide, please?

We have also been explaining that this CAHPS -- the HCBS CAHPS survey can be used in a variety of ways for program quality management.

First, it can help to document program successes. In addition, by fielding the survey over a short period of time, users can get a point in time performance snapshot to identify areas needing improvement.

And, the, of course, you would want to do further investigation to determine the cause of any problems identified.

Another way to use the survey is to repeat survey administrations in order to track performance over time and monitor changes.

For example, before and after implementation of a program improvement project.

And, provided there's a sufficient beneficiary sample for each program or subprogram group, the survey can make comparisons among programs serving individuals with different types of disabilities.

It can be used to convey performance information to a variety or stakeholders within the organization, beneficiaries, providers, even state legislatures.

And, because the measures align with some of CMS's quality requirements, the survey can assist states engage in compliance with regulatory requirements.

Next slide, please?

So, we do have some examples of state uses. Right now, there are seven grantees from the test states that are demonstrating use of the survey in their HCBS programs.

And, their plans range from comparing performance across programs to identifying quality improvement opportunities, to exploring whether future managed LTSS programs should use the instrument.

We have one grantee that is further studying response rates from different administration modes including an online version.

And, we have another grantee, the State of Connecticut, that is using the survey to set quality benchmarks for performance incentive payments to case management agencies.

And, if we go to the next slide, I just have a little bit of information about their initial results.

They are the furthest along, and that's why they've been able to share their results to date.

They are focusing on three waiver programs, those serving older adults, individuals with physical disabilities, that's their personal care assistance program, and individuals with brain injury.

They let the participants choose whether to take the survey over the phone or in person and they allowed both assisted surveys as well as proxies to complete the survey.

Next slide, please?

So, this slide shows preliminary response rates from Connecticut. These are in the 60 to 70 percent range at the bottom of each column for two of the populations.

And, Connecticut recently updated that the response rate for the brain injury group was about 61 percent. And they attribute their response rates, first and foremost, to allowing the assistance with the interviews and also proxy respondents.

Next slide, please?

And, then, this slide of Connecticut's preliminary participation results indicates that, when given a choice of survey mode, more than 9 or 10 are choosing telephone. I think it's more like 95 percent actually.

Among the brain injury population, just under 20 percent have been in person. So, the majority still prefer the telephone option.

And, Connecticut indicates that people who prefer an in person option are usually those with a communication issue where they need to meet in person and also the assisted interviews are sometimes preferred to do in person.

And, this slide also shows that the PCA waiver participants were most likely to complete the survey by themselves compared to older adults.

And, again, in terms of the brain injury group, about 65 percent completed it alone as opposed to having either a proxy or some help.

And, the last slide is just a resource slide in terms of -- yes, just different places that people can go for more information about the survey.

And, that's it.

MEMBER POTTER: Hi, this is D.E.B. Potter again.

Is this Truven mailbox still up to date? I've got other Truven contractors that now have different email addresses. So, I bring that to your attention.

DR. JACKSON: Yes, we are aware of that. We have not yet migrated that mailbox, but we will. We are working with CMS to make sure that it is seamless. Thanks for noticing that.

MEMBER PARKER: Hi, this is Pam, again.

I'm quite interested and excited about the -- you're doing a couple of different things in here than what we've heard the standard CAHPS approach. Now, I wish that Stacey was still here, Stacey from CMS.

But, first of all, the fact that you're allowing someone to translate it verbally. I know that the regular CAHPS methodology just has been, you know, absolutely opposed to that.

And, I think this is so important for

this population and not just having the instrument in a different language, that isn't all that useful, but having someone to be able to do it, you know, verbally.

And, of course, then the other is that you're doing it by telephone and you're allowing proxies. So, the proxies, you know, actually administering it by telephone with a person or in person, but I think telephone's the most important.

And, then, the other one in terms of the language.

In Minnesota, we were doing a project where we've been trying to -- we cover much of the same population with the regular CAHPS and we've integrated the Medicare and the Medicaid CAHPS survey instruments with some success, though, because of these methodologies that CMS insists on, we don't really -- oh, here's Stacey -- we don't get to -- Stacey --

MEMBER LYTLE: Just in time.

MEMBER PARKER: What I was excited

about here is, I was hoping that what they're testing here with proxies, the language and the in person or on the phone surveys, you know, being able to use addition people, this would be so good for so many dual eligibles across the country if that testing, you know, that they've done now could be transmitted to other CAHPS documents like for the MMPs and like Minnesota with all its, you know, immigrant seniors that are left completely out of these surveys.

MEMBER LYTLE: If only that call had lasted two minutes longer.

So, I think the short answer is it's not that easy to translate it to the other CAHPS surveys. We have been in discussion with like the Medicare CAHPS people and others about what it would mean to do some different testing and to use testing that we've seen elsewhere.

But, survey order and things that we all know about come into play then.

However, we have been in MMCo trying to figures out just where this HCBS CAHPS survey fits and how we can better use the information that it presents because of our population with the MMPs and others.

So, it's definitely on our radar and something we've been thinking about because it's information we don't have and we really want, number one.

And, number two, because it's done all the testing. So, we know that it works as it's structured. So, we're working on it. I don't know if it gets to the Minnesota issue because there's several questions that you all asked and that the plans asked that are a little different. But, we are looking into how to utilize what's available there.

MEMBER PARKER: Yes, and if I could just add that, you know, there is the issue then, too, of the juxtaposition of just how many CAHPS do these poor people have to have?

And, in this case, let's say it's, you know, well, Minnesota's the only that's tried to integrate the Medicare and Medicaid so Medicaid may use the CAHPS and Medicare may use it. And, there may be in this field at the same time and then you have this third CAHPS of this one being, you know, discussed and proposed.

And, you know, that reduces response rates terribly.

MEMBER LYTLE: And, that's our concern that --

MEMBER PARKER: Yes.

MEMBER LYTLE: -- the survey burden is just too much --

MEMBER PARKER: Right.

MEMBER LYTLE: -- when we say take these for --

MEMBER PARKER: So, there's got to be some kind of hierarchy or some plan around that.

But, then -- and then, the other last thing I was going to say was I think everybody should know and maybe you all are expert enough to know this, but I think it's pretty atrocious, I have to say, I know Stacey knows how we feel about this, that when there isn't even a real language block on the CAHPS that, you know, has in different language of what this item is.

And, when we had to switch from the Medicaid CAHPS to the Medicare, they made us use all the Medicare approaches to the CAHPS. And, so, we couldn't put a language block on saying, this is, you know, in many different languages saying that this is, you know, even if you don't fill this out, it's okay and but it is.

CO-CHAIR MONSON: Yes.

MEMBER PARKER: So, they don't know what it is. They get it and they know that they don't have to fill it out, that's stated. But, it doesn't say what it even is or tell you anywhere to go and get help for it or anything.

So, it's really, I think, such a poor way to get at some of the frailest and sickest and most disadvantaged and health legacy disadvantaged folks that we have.

CO-CHAIR MONSON: Okay, we have time for one more comment.

MEMBER LAKIN: You allow proxy

comments sort of I gather you leave that up to the program.

I just wonder to what extent you've studied the sort of the inter-rater reliability between primary service respondents and those who respond as proxies?

The evidence in -- with some surveys that have looked at that suggest they're not at all equal and they're somewhat idealized by the proxy as compared with a service recipient.

Is that something you've looked at and put -- if so, do you give that advice to people who might be making a decision about using proxy respondents?

I'm particularly concerned with 50 percent of the respondents for people with intellectual and developmental disabilities are not service recipient.

DR. MALLERY: This is Coretta Mallery.

And, we did look at that. And, I could give move specific details on, I believe, but I'm just going off recall, that only one of the measures was higher for proxy respondents. But, I'll need to dig back and find which report that was.

But, I agree that there are, you know, there's a lot of research out there that shows that proxies may respond more positively than the beneficiaries.

So, what we would recommend is absolutely taking a look at that if you were including proxies and we would also recommend including it as a case mix adjuster. So, you know, including whether or not a proxy responded on behalf of the beneficiary. So, but the scores could be adjusted for that.

CO-CHAIR MONSON: Clarke, did you have something burning to say?

(OFF MICROPHONE COMMENTS)

CO-CHAIR MONSON: Okay. Got it.

All right, well, thank you Kerry and team. We appreciate you sharing with us today and, you know, it'll obviously be very helpful for us as we consider this from a voting perspective.

And, so are we not voting now?

MS. MUKHERJEE: So, no. What we'll do is we'll vote via SurveyMonkey and it's going to go out, so please look at your email and there'll be two questions.

One for inclusion into the duals family of measure.

And, two, for inclusion to the starter set.

So, instead of voting now, just for the sake of getting everybody's voices and some people had to leave, we're going to send out a SurveyMonkey voting.

MEMBER ROSS: I did want to highly encourage us to endorse this measure. This is a CAHPS trademark. This is already National Quality Forum endorsed.

The MAP report to CMS, two of the six high-value measure areas are patient-reported outcomes and patient experience.

This is a high priority of not only

the Consortium for Citizens with Disabilities, but the Disability and Aging Collaborative that Joe Caldwell chairs with NCOA.

So, this is a big priority that's been in the works for quite a few years that the advocate and provider community-based provider community thinks is really valuable.

And, so, I encourage you all to vote yes.

CO-CHAIR MONSON: So, even though we're not voting, I do think -- do other people want to chime in about perspectives on the CAHPS? Pro? Con? Yes, go ahead.

MEMBER LUTZOW: There might be 11 votes here right now.

(OFF MICROPHONE COMMENTS)

CO-CHAIR MONSON: We can -- okay, we can vote. So, therefore, just -- but still, again, does anyone have any more commentary on the measures themselves before we vote?

Yes, Charlie?

MEMBER LAKIN: You know, in general,

I'm in support of this. I'm not sure -- what are voting for an instrument in this case? Are we voting for composite measures? Are we -- I don't know quite what we're endorsing.

MS. MUKHERJEE: So, it's the measures. It's the measures.

CO-CHAIR MONSON: All of them?

MS. MUKHERJEE: All of them, yes.

MEMBER LAKIN: The ones that were endorsed not --

MS. MUKHERJEE: Endorsed.

MEMBER LAKIN: -- the whole survey, you need to go back and show them what --

CO-CHAIR MONSON: Yes, can we see which ones?

MEMBER LAKIN: -- need to clarify.

MS. MUKHERJEE: So, do you guys want to screen share 2697, the --

(OFF MICROPHONE COMMENTS)

MS. MUKHERJEE: Yes, so, we're voting on all the measures that came through and were endorsed. And, under 2967 and we're screen sharing. And, there are a lot of measures, but this is not a survey, these are sort of questions.

CO-CHAIR MONSON: And, they were endorsed by which committee?

MS. MUKHERJEE: The PFCC, Person and Family Centered.

CO-CHAIR MONSON: Was there any -- do we know, was there any commentary from that group that we should be aware of?

MS. MUKHERJEE: Also, for a quick reference, it's one on the slide in the PowerPoint, so if you want to look at the PowerPoint, if you have it one your desktops.

DR. ROILAND: So, while we're getting everything pulled up there, I can give you a brief summary of the issues that the Committee went through when they reviewed the measure.

The first round of evaluation at the in person meeting of the PFCC Standing Committee, there was a lot of questions around the which -- how the measure would be applied to an organization. Basically, who would be held accountable for the measure.

And, so, the developer provided clarity around that saying that -- I'm sorry, let me pull this up here, I don't think we're loading.

CO-CHAIR MONSON: Okay, so, which one -- when we're looking on your screen, is it the recommendation measures that we're looking at or the global ratings measures or all of them?

MS. ROILAND: so, all of these measures are considered NQF endorsed underneath the NQF Number of 2967. There are 19 measures, I believe. You might need to scroll down a little bit, Madison. Sorry, can you scroll back up, let's start from the beginning.

So, there are scale measures. Those are the measures that they went over in the first part of their presentation. So, they're a very number of items that are grouped together to create a scale measure. There are seven of those.

Like they said in the presentation, they are, beginning with number one, staff are reliable and helpful, staff listen and communicate well, case manager is helpful, they're able to choose the services that matter to them, they have access to transportation to medical appointments, another scale related to personal safety and respect and then planning your time and activities is another scale measure.

Then, they move on to three global ratings measures and these are global ratings of personal assistance and behavioral health staff, global rating of the homemaker staff and global rating of the case manager.

We had some discussion around two of those three.

Then, there are three recommendation measures. Would you recommend your personal assistant or behavioral staff to family and friends? Would you recommend homemaker staff to family and friends? Would you recommend case manager?

Then, there are five measures related to unmet needs. They're on dressing and bathing, preparation -- meal preparation, medication administration, toileting and other household tasks.

One physical safety measure, if you can scroll down a little bit, Madison, talking about harm from the staff.

And, so, those are the 19 measures. So, if these measures -- these measures are derived from the items on the HCBS experience of care survey, or the CAHPS survey, I'm sorry.

So, the measures are derived from that survey and each of these are endorsed measures under the -- under NQF's approach to endorsing measures derived from the survey.

MS. JUNG: And, I can read the summary of what the Committee said. Would you like that?

So, this for the CAHPS survey. This new PRO-PM is a package of 19 different measures calculated from data from a newly developed experience of care survey focusing on HCBS programs.

Numerous challenges were identified with this measures submission including level of accountability and variation in the types of programs and services offered both across and between states.

The developer noted that the survey and reporting of the measures are being introduced for voluntary use by states and relevant programs and would help programs identify areas for quality improvement interventions.

Committee members with experience in this area noted that what matters to consumers is that their needs are met, not who is meeting them.

The Committee decided to vote on evidence all together and then split the measure into five measure domains and vote on each of the domains separately for performance gap and the remaining criteria.

The performance and testing data submitted for these measures were limited due to the pilot testing of the survey so the Committee found it challenging to understand the opportunity for improvement performance gap and reliability of some of the domains.

The Committee provided recommendations to the developer on the opportunities to address some of the data challenges. However, ultimately, two of the measure domains failed performance gap and the remaining measures failed on reliability.

The Committee encouraged the developers to determine if alternate testing procedures might be better -- might better differentiate programs and better support the reliability of the metrics.

And, there is a lot more. Let me see if there's a summary. I can keep reading.

DR. ROILAND: Would it be better just for us to share this with folks and then we can still vote via survey so you have time to review it better or what would everyone be most comfortable with at this point?

MEMBER ZLOTNIK: Could you just clarify what you're asking us to vote on to add it to our family of measures?

DR. ROILAND: It would be -- so, we're asking you vote on Measure NQF Measure 2967 and that measure number encompasses all 19 measures from the survey that are created from the survey.

DR. ZLOTNIK: Right, but are we voting on it to add it to our family of measures?

DR. ROILAND: Well, yes, we're voting on that. But, then, yesterday, the issue also came up that someone wanted to propose also adding it to the starter set, so we're going to ask you to vote on those two things.

DR. ZLOTNIK: Thank you.

MEMBER PARKER: Could you also just say something about the scaling? Is it typical CAHPS? Is it, you know, it says yes and no and then there's something in between and where is the cut off?

DR. ROILAND: It's different for each item. So --

MEMBER PARKER: Yes, it's something like never, sometimes, usually, always.

MEMBER HAMMEL: But, then they have an alternative.

CO-CHAIR MONSON: So, do people -- are there -- do people have concerns about these measures? Let's just start there.

Yes, Joy?

MEMBER HAMMEL: I didn't have concerns until you read off that entire report which doesn't sound very positive.

CO-CHAIR MONSON: So --

DR. ROILAND: So, in the initial round of evaluation, there were some concerns around the testing of the measure, I think, particularly in terms of they had a lot of subgroups, they had a larger sample, but a lot of subgroups that they tested the measure on.

And, so, the Committee asked the developer to bring back additional testing which they did during the post comment call which happens after the draft report is released.

With that information, the Committee was satisfied that the measure met reliability and validity and also demonstrated performance gaps.

MS. JUNG: Yes, and the following five paragraphs, it got a lot better.

MS. ROILAND: So, ultimately, it was endorsed this past October as I believe Kerry mentioned.

CO-CHAIR MONSON: Joan?

MEMBER ZLOTNIK: My concern is really on the workforce issues. They -- without knowing how the survey is used and then how the information is then reported, it has the potential for giving -- providing very unclear information.

So, like, in general, like, the information that's included in those 19 areas that have been obviously functionally put together, makes sense. But, it just concerns me that the surveyors are using words for staff that multiple meanings.

And, while someone can be filling it out, I just feel like I'd be sitting in some congressional hearing and, you know, someone would say, oh well, these, you know, case managers aren't doing their job or whatever it is. These things happen.

So, that's my concern.

CO-CHAIR MONSON: Yes, D.E.B.?

MEMBER POTTER: In the interest of disclosure, I'll say I've served on the Technical Expert Panel for this survey for many years.

What we didn't see the full scope of was all of the up front questions. You know, it starts with a program. Does someone come in to your house and who do you call them?

And, because it's a CAPI and a CATI survey, the interviewer can insert the respondent's concept of what that is. Oh, Mary helps me make my bed and get dressed in the morning.

And, so, the subsequent questions say, and does Mary do this and does Mary do that?

That's one of the advantages of having it be a CAPI or a CATI survey to try and deal with all of those legitimate issues that you bring up.

And, as we heard earlier, in some states there are multiple versions of this. But, because the accountable entity is a program, it's not like everybody in a state. You start with, well, what does the program call it? You know, and builds from there.

I don't know if that helps but --

CO-CHAIR MONSON: Wendy?

MEMBER FOX-GRAGE: I am all for this. We are so in need of an experience of care for people in home and community based services programs. So, I'm a fan.

My only question is this, and it's because I'm new, but I don't know this, and that's the question of, you know, being in the family versus the -- and also being in the starter set.

So, I just want to make sure, then, in the family, do we have any -- I know we've spent all day yesterday voting on measures that were really more for an institutional setting.

So, I just wanted to ask what that -- the part that we didn't cover which is the whole family of services, is there anything else at all that's experience of care survey on HCBS? There isn't, is there? Okay, that decides it for me.

CO-CHAIR MONSON: Clarke?

MEMBER ROSS: So, right, Wendy. So, this is a priority of the Disability and Aging Collaborative because we don't have these measures.

This is tested in eight states and this is already endorsed. And, I'm concerned that we would send a message to the entire field, not just National Quality Forum, that, because the status of social workers and, as they say on The Hill, my good friend Joan and I've worked with Joan since the 1980s on nursing home reform, right.

Because the social work phrasing isn't clear, we're going to vote down one of the few beneficiary experiences in home and community based services which we're all about.

And, the point I was going to make to Stacey on the side in response to Pam's question was, I asked at the MMCO stakeholders meeting we had two weeks, this very question, what is the use by MMCO and the duals demos?

And, Tim said there's a question on how to best fit it. And I went to Steve Kaye, who's one of the five researchers at the University of Minnesota. He directs the University of California San Francisco Center and was co-chair of the National Quality Forum HCBS Committee.

And, he provided data on three states that show a significant number of the dual demo participants are also recipients of home and community based services under Medicaid.

So, I'm just getting really uneasy

that, because every one of these things has technical challenges, but I'm concerned about the big message.

And, the message should be, this is needed. It's well overdue, it's important and it should part of not only the duals family, but personally, it's so important, it should be part of the starter set.

So, that's all I can do to plead with.

MEMBER PARKER: And, I just want to echo what Clarke is saying. I think, with all its flaws and CAHPS surveys, you know, have lots of flaws in them, it's the best thing that I've seen and it moves it forward with some of these other little accouterments that I've already talked to you about.

And, I think with the advocacy community being where it is and we just heard that from two of them, I think it would be -- I would feel really bad if this was my first and last meeting of this group and I had been part of putting something like this to death when it's an opportunity to make such a big movement forward.

So, I would just plead for everybody to support it.

CO-CHAIR MONSON: Alison and then Tom?

MEMBER CUELLAR: Yes, I'd like to say I wholeheartedly agree. It's so difficult to try to evaluate these dual demonstrations without an instrument like this or managed LTSS that it's --

When we were looking at this thing, it was still over 60 minutes long. It hadn't been tested and so there was nothing the state could do or that we could help the state with at that point.

But, we made something up. This is a huge improvement.

MEMBER LUTZOW: Yes, I'm in favor, too, only because we need another CAHPS survey to debate over for the -- in the future.

But, I mean, this is a case where, you know, we can't let the perfect get in the way of the good.

You know, in Wisconsin right now,

every MCO is doing its own survey as opposed to using a standard survey. And so, there's no way to, in terms of these community based services, there's no way to compare one MCO to the other.

We send in our survey and they record the results and our survey isn't like anybody else's.

And, so, I mean it's clear the field needs this. It's looking for it and, you know, it can get perfect over time.

CO-CHAIR MONSON: So I would just add that from another -- from a national health plan perspective we would love this. Look, at the end of the day all of us are here to help the individuals who are duals, in this case people with HCBC services, live the best life possible and remain in the community. And this is a critical tool for us to be able to make sure we're doing that correctly.

And if we can't -- if we as a committee can't get this one over the -- this is a very well-developed measure, by the way. I've been following this one for years. This one has gone under a lot of scrutiny. There's been a -- and there's also been a lot of sharing of information along the way, so it's not been hidden. It's very well designed. It's been executed well. If we can't push this one over, then we do deserve to be put on hiatus.

Okay. Are we ready to vote? Are we -- Madison, are we doing your thing or we're doing it by hand?

DR. ROILAND: And, Alice, if you're still on the line, you can chat your vote to us in the webinar platform chat box.

MEMBER LIND: Okay. Got it.

DR. ROILAND: Thank you.

CO-CHAIR MONSON: Okay. Did we say -- I'm sorry. Are we using the clickers? We have. Okay. All in favor?

PARTICIPANT: Wait, wait.

CO-CHAIR MONSON: Oh, sorry. I think we start with the family. We're doing the family first and then we'll do the --

(Simultaneous speaking.)

MS. JUNG: So voting for Measure 2967, CAHPS home and community-based services measures, is now open for the addition to the family. Well, okay. So we'll start off with yes. So raise your hand if you would indicate yes?

(Voting.)

MS. JUNG: And she's yes.

MS. BUCHANAN: Oh, she's yes?

MS. JUNG: Yes.

MS. BUCHANAN: Okay. Great. So we have 15 in favor. So that's -- okay. And so there would be zero not in favor. So it passes.

MS. JUNG: It passes? Okay. So Measure 2967 passes with 100 percent for addition to the family of measures.

The next vote will be the vote for addition to the starter set. Please raise your hand if you would like to vote yes for that.

CO-CHAIR MONSON: Yes, hold on just one second. Jen has a quick question.

MEMBER RAMONA: Sorry, being newbieish. Can you define what the starter set -- what that really means for the use of it? Sorry.

CO-CHAIR MONSON: So the starter set is supposed to be that if a state or an entity wanted to pull -- what are the first measures they would use to do dual measures? And this would be in that -- as opposed to a large group of measures. This is the first ones you'd pull out. So there's 10, 15 on the starter set right now?

PARTICIPANT: Seven.

CO-CHAIR MONSON: Seven? I wasn't that -- high priority is the way I think about it. These are high priority measures.

MS. JUNG: Okay. So voting for the starter set of Measure 2967, CAHPS home and community-based services measures is now open. Please raise your hand if you would like to indicate yes.

(Voting.)

MS. BUCHANAN: So I have 14 in the room. So, Alice, yes. We have 15 yeses.

MS. JUNG: So with 15 yeses that's 100 percent and the measure is now added to the starter set.

CO-CHAIR MONSON: Awesome. Well done. We have -- no, in fairness, that's very true. We have accomplished something. I mean, this is important, because now that it's NQF-certified, this also means that as -- if you think about from -- I'll take my parochial view with managed care -- and we know many states are moving to MLTSS. They need to have ratings. They have to have a quality rating system by -- because of the new managed care rule. So now this will be one that they can use because this is an endorsed measure, which is great.

Okay. We're doing the strategic -- yes, all right. So this is our -- this was -- as you will recall from yesterday, this was -- we said we'd have reserved time at the end of today again to provide more thoughts back to CMS for final parting thoughts about dual measures. So I will open it up to the group in terms of those final thoughts. Anything that we really want to -- and I would say things that you want to leave CMS with: recommendations, ideas, this is your forum to do that.

(Pause.)

CO-CHAIR MONSON: And you're not allowed to leave yet.

(Laughter.)

CO-CHAIR MONSON: Yes, and it may be that -- Alice, you -- we have not given you an opportunity to speak. Is there anything else that you want to share? And you weren't -- you didn't join the conversation yesterday.

MEMBER LIND: I'm trying to think if I remember the topics from yesterday, if I would have had something unique. But I'm really comforted with the compilation of new folks that have been added to the work group and don't feel that I have anything that my colleagues wouldn't have said yesterday.

CO-CHAIR MONSON: Tom?

MEMBER LUTZOW: Yes, I think as a

general statement from my perspective I would never want to go back. I would never want to go back to 10 years ago. I think, yes, we can debate this and nitpick that and -- but what's been achieved in terms of performance measures and getting people on the same page and working with evidence-based outcomes and that sort of thing, this is quite amazing. This is quite amazing. And the short amount of time that it's taken to get here.

Now I still want to debate. It's about adrenaline and it's about living longer. And certainly we can nitpick, but life without measures, I don't want to go back. Now I want to debate which ones should apply and how they should be weighted and all of that, but something great has happened here, not just in this room, but within the total healthcare environment.

MEMBER AGUIAR LYNCH: So I just have a -- sort of a plea for you guys at the Duals Office. So I know you guys are focusing on filling in the measurement gaps and working on social risk factors and the Star Ratings program. And as you guys are working on developing your Star Rating system for the MMPs, I think this is such an amazing opportunity for you all. And I think you guys are the office that's perfectly positioned to be able to do it, to think outside of the box as much as you can to really pull in what you've been hearing from this workgroup over so many years and to try to get the measures that are actually accurate and then the way that the system is developed in a way that's -- truly is measuring quality and comparing quality accurately and that it has the right incentives for in this instance plans as they move forward, and not to repeat the mistakes that were made in the original Star Rating system. So I know it's a tall order, but you guys could do it.

MEMBER LYTLE: Can I respond?

CO-CHAIR MONSON: Yes, of course.

MEMBER LYTLE: Just in response to both, I think we totally agree we don't want to go backwards either. It's not -- it's not our desire, our goal, our hope or our strategy. So we are in complete agreement. As it concerns the Star Rating system, I think that you're right. And we're trying to be very thoughtful about it because the current Star Rating system isn't really applicable for the population that we serve. It's what we have, but we recognize that we need something that works better.

And so hopefully -- unfortunately I am drawing a blank as to where we are in the process, but I know that throughout the process we've done public comment and we've tried to get input. And so I would hope that this -- the members of this group will continue to be engaged in that way because your feedback is still very important to us, not just with the Star Rating system, but in everything we do. We do try to incorporate the voices of those persons who we serve, whether they be the people who are receiving services or the people who are administering them. And so we just hope to continue to hear from you from that perspective.

So we're trying.

CO-CHAIR MONSON: Yes, ma'am?

MEMBER PARKER: Of course my everlasting plea to CMS is that as you look at the MMP ratings system that you've solicited comment on and everything, that we look at what is applicable to the integrated D-SNPs, recognizing that that's a larger even platform than the MMPs and has many of the same features and that we're moving -- trying to move that world further and further toward something that can come together in terms of what the MMPs and the D-SNPs are both accomplishing, both working on the same stuff.

And that population also, especially for the integrated ones, of course is heavily dominated by the home and community-based group as well. So just what's good for one should be brought over to the other, and that's -- we talk about that in terms of network materials and network approaches, and also measurement.

And then as we're thinking about that,

especially on behalf of the D-SNPs, we have to think of how many layers are they under. Are they under a whole set of things from the state level, another whole set from Medicare, and then yet another set that's specific to duals and how do we make all that make more sense? And as you work that out for the MMPs I think it's going to be very instructive to then see how that could apply to the D-SNP world and get rid of some of those layers by looking at what you're doing with MMPs. So that's just my own --

(Simultaneous speaking.)

MEMBER LYTLE: We hope so. We've spent a lot of days thinking about how to make things applicable across --

MEMBER PARKER: Right.

MEMBER LYTLE: -- all of the care plan models. And our hope and desire is that whatever we do with the MMPs is instructive and then applicable across populations with nuances of course, because the way they serve and the way they contract are different. But Pam knows all too well some of the concerns and constraints we've had there, but we still press.

MEMBER PARKER: Yes. And then we have to think of the providers, because they might be working under fee-for-service systems as well as all these other layers. And so we have to make it somehow make sense because they can't be expected to do everything that a thousand entities are telling them to do. They've got -- we've got to focus it so that it makes sense at that level, too.

MEMBER LYTLE: And can I just note one more thing? I know I have probably talked a lot about the fact that we have demonstrations and we have things happening, but I also want to be very clear that that's part of what we do. And so we are very concerned with how D-SNPs and fee-for-service providers and all other dual-eligible individuals who are not in plans are receiving care. So I just wanted to emphasize that because I know I've sort of referenced it as our approach, but it's -- our office's goal is not just about demonstrations. It's about the population at large.

MEMBER PARKER: And we wish you well --

(Laughter.)

MEMBER PARKER: -- as they're -- I mean, it's so wonderful to have you all and to have a focus point that we can even have a discussion about this. I know somebody's supposed to be kind of monitoring it, but that wasn't always the case for the 20, 30 years before. And we're just hoping it all continues and we'll be watching for --

CO-CHAIR MONSON: So I would add one thing for you to think about, which would be I think getting to a common assessment form is a highly unlikely endeavor. I would wish it to be true, right, across -- for all duals, however, I do think what is achievable is the idea of some common -- a handful of common questions as opposed to a whole common assessment, but a handful of common questions could be driven by measurement, right, that the measures require certain questions, especially around ADLs and IADLs.

If we had a common way of measuring them, will that allow us to risk-adjust better, both from a financial perspective and also from a quality perspective. So I would highly encourage you all to think about where you could place some small bets from -- and do it from the Medicare side that would apply to all eligibles.

Tom?

MEMBER LYTLE: I think that makes sense. And I just want to say if I'm not looking at you, I'm typing, because if I don't write it down, I might not remember it. Sorry about that.

MEMBER LUTZOW: Yes, I think when you look at how this field is organized now with respect to measures, not just a common -- we don't need just a common set of questions. We need a common set of measures across the entire system. And it is sort of striking that that's not really prominent.

The value of the IMPACT Act is it is going to test what happens when nursing homes, home health agencies, hospitals and others get -- rally around a common measure like readmission prevention. And we're going to be able to tell whether joining all those resources at the hip around a common measure really can make a difference. But I think there's other things within the family of measures that we now have, like diabetes, where if it were a measure that everybody had to subscribe to we would see population health gains that we don't now see.

So why we're resistant to this as an industry: government side, provider side, whatever side you want to consider, why that isn't sort of self-evident as a value I haven't been able to figure out, but it seems so clear that if we want traction, we're all going to have to get behind the same measure and push.

CO-CHAIR MONSON: Other thoughts?

MEMBER LIND: Hi, this is Alice. Can I actually insert something after all?

CO-CHAIR MONSON: Oh, I don't know.

(Laughter.)

CO-CHAIR MONSON: Okay. Go ahead.

MEMBER LIND: So just the conversation about where our attention is focused and who has still kind of been the left-unattended-to group, the population that is the most complicated for me personally as a state bureaucrat right now are the folks who are dual-eligible with behavioral health needs.

We are using integrated managed care to cover 80 to 90 percent of the folks in a completely integrated package of medical and behavioral health services, and the duals are the ones that really are still in this kind of bifurcated model where the Medicare and the Medicaid benefits around behavioral health are really so different. And the providers have been just kind of cobbling together benefits for these clients as best as they can, and our inserting the managed care plan really has not helped this situation at all, just for those duals.

And so I'm really resonating with that thought about the folks who are not in the demonstrations are the ones we really are going to have to pay attention to, and for us right now it's the folks with behavioral health needs.

CO-CHAIR MONSON: Yes, Charlie?

MEMBER LAKIN: Well, we could go through this. We've focused on the dependent variables. We could spend just as long talking about the independent variables. And I think those are discussions that are really important as we move to some consensus that we might be able to measure outcomes for people in the same way despite very different characteristics. And that becomes increasingly important I think when we talk this way.

I'm struck having worked mostly in the area of intellectual disability how we have never come up with a better predictor of outcomes for people than their level of intellectual disability. And so often we just -- we're satisfied with saying they have an intellectual disability. That is really meaningless.

So at some point we really do need to think about the independent variables that really are important. And I just think it's going to be increasingly important as we develop data systems that encompass just more and more of these fields that have lived in silos, need to try and get out of their silos. But the characteristics of people in a broadly framed disability community are hugely different and we need to accompany our consensus around the dependent variables with some consensus about the independent variables, I think.

CO-CHAIR MONSON: All right. Well, I think that brings us to a close. I think it's important for us to first of all thank CMS for -- oh, did you want to --

MEMBER LYTLE: I just wanted to say thank you. Should I do that now?

CO-CHAIR MONSON: No, I was trying to say thank you first.

(Laughter.)

MEMBER LYTLE: It's not fair.

(Laughter.)

CO-CHAIR MONSON: Go ahead.

MEMBER LYTLE: Well, I get to go first. Thank you very much.

But I do want to thank the group again for all your hard work over the years. I think that many of the strides that we've made as an office and an agency are in large part of your contributions, and that -- we can't say that enough. Even though I think some people said this is their first meeting and they're new, that is -- I know that feeling. And it's still valuable because your work in the industry stands for itself and still contributes to the ongoing discussion about people who are Medicare and Medicaid-eligible. And so we appreciate you and we appreciate NQF for your work with the workgroup over the years.

We aren't all happy and excited about where we are right now in terms of not being able to see you in a meeting again, but we are excited about the way the work is going to continue to move forward, and hopefully we'll be able to have these discussions again in another forum not too long from now.

CO-CHAIR MONSON: And I was just reminded we need to do public comment.

DR. ROILAND: Hi, Shawn. If you wouldn't mind opening up the lines for public comment?

OPERATOR: Okay. At this time if you'd like to make a comment, please press star then the number one.

(Pause.)

OPERATOR: And there are no public comments at this time.

DR. ROILAND: Thank you.

CO-CHAIR MONSON: So I do think we should thank CMS, because I think it was very forward-looking to put this entire infrastructure together and then to call out the duals in particular and allow this group to do its work. And that's no small measure to Stacey and all her colleagues, because we do know that the MMCO is a true partner and collaborator in improving the care and the lives of all dual-eligibles regardless of their payer source. So thank you for that.

I think we should thank the NQF staff who do so much work.

(Applause.)

CO-CHAIR MONSON: And it's kind of unbelievable, and we couldn't have done any of this without them. And I think that we have left ourselves -- while we are going on hiatus, we have left ourselves I think with a very good family of measures and a great starter set, especially with the addition of the CAHPS surveys. And so I think if part of our purpose is to make the world a little bit of a better place, then we've accomplished that.

And thank you, everybody. Most of you have been on this longer than I have, so thank you to everyone who was doing this before some of us newbies got on here, but it's been a pleasure working with everybody and hopefully we will have another opportunity to convene under a different banner.

So I think we are adjourned.

MS. JUNG: We need to --

CO-CHAIR MONSON: No? No.

(Simultaneous speaking.)

MS. JUNG: -- next steps. And we have just --

(Laughter.)

MS. JUNG: We just don't want to let you guys go, yes. Just a few closing housekeeping things.

So in terms of -- we have taken copious notes and have noted all of your concerns and everything, and those will be summarized in our -- in the report that will be finalized in August, but before it is finalized it will go out for a 30-day public comment period. That will be from June to July.

There is our contact information. And you guys have the slide deck that was also emailed to you. The project page is linked here as well as the Committee SharePoint page with the slides, family of measures, duals, Excel, agenda. All the meeting materials from today will be posted. And you will be receiving a very long email of the laundry list of resources and follow-up items that you have all requested.

Listed also here is the project staff contact information. There's our inbox, which I'm sure you've received many emails from and the individual emails from our staff team.

And we thank you for your patience with all the technical difficulties and everything, the clickers, and really appreciate your work. And turning it over to Debjani for a few closing remarks.

MS. MUKHERJEE: So again, I would like to thank all of you, the longstanding and sort of the new members, and especially Michael for sort of shepherding us to the bittersweet end and sort of stepping in. For any workgroup committee to work really well, it's a lot on the chairs to sort of manage and sort of keep up with the flow and sort of manage the flow.

So thank you to the NQF staff, the team, the duals team. I think a lot -- as you can imagine, a lot of work has gone into this and with the news a lot of sort of rejiggering the different sort of aspects of the presentations and things and that were all sort of done last minute. So it takes a lot sort of from everybody to sort of rally together and hope -- and to CMS for allowing us to work with them for this long and sort of in this capacity. I think it's been very interesting.

And with that, I want to sort of wish everybody a safe journey, safe travels. Clarke, with his surgery and being in recovery coming here, thank you.

And so with that, I think we're officially adjourned.

(Whereupon, the above-entitled matter went off the record at 4:13 p.m.)

CERTIFICATE

This is to certify that the foregoing transcript

In the matter of: Measure Applications Partnership
Dual Eligible Beneficiaries Workshop

Before: NQF

Date: 03-30-17

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R Gross

Court Reporter