THE NATIONAL QUALITY FORUM

+ + + + +

PULMONARY TECHNICAL ADVISORY PANEL

PATIENT OUTCOMES PROJECT

+ + + + +

Thursday, December 3, 2009

+ + + + +

The Pulmonary Technical Advisory Panel met, in Suite 600, in the Homer Building, 601 13th St., N.W., Washington, D.C., at 8:30 a.m., Barbara Yawn, Chair, presiding.

PRESENT:

BARBARA YAWN, MD, CHAIR
MARK MILLARD, MD
MARGARET NEFF, MD, Msc
RICHARD D. O'CONNOR, MD

ALSO PRESENT:

MEASURE DEVELOPERS:
GERENE BAULDOFF, MD, American Association of Cardiovascular and Pulmonary Rehabilitation
FRANCOIS de BRANTES, CEO, Bridges to Excellence (via phone)
R. ADAMS DUDLEY, Philip R. Lee Institute for Health Policy Studies, University of California
LARRY HAM, MD, American Association of Cardiovascular and Pulmonary Rehabilitation
AMITA RASTOGI, MD, Bridges to Excellence

STAFF:
ALEXIS FORMAN
EMMA NOCHOMOVITZ
KAREN PACE
REVA WINKLER

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

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

8:24 a.m.

CHAIR YAWN: Good morning, everyone. Welcome.

I hope all of you know, if there are any attorneys in here and you think this is for depositions, you are in the wrong room. We are in an attorney's office, a law office.

(Laughter.)

But this is the TAP for Pulmonary Outcome Measures, and we are very pleased to have everyone here for our meeting this morning.

I think we are going to start, basically, with introductions. So I am going to ask everybody to go around and tell us your name, obviously, where you are from, what kind of work you do that you think would be related to outcome measures and quality improvement, and things like that.

And maybe after lunch, we will do another quick round and you can tell us one exciting thing about you that you think everybody should know, but they don't.

(Laughter.)

But we won't start quite that way this morning. It is probably too early.

I am Barbara Yawn. I am a family physician. I am from Rochester, Minnesota. No, I do not work at the Mayo Clinic. I work at the other group in town, which is a group of 140 physicians, mainly primary care.

I do research full time and have for several years now. So I am the Director of Research there. Have been involved in lots of guidelines, panels, and was involved on the expert review panel for asthma, and am becoming a part of the gold group, I think in -- they haven't decided if it is June or July, but sometime like that in the next year.

So I am very excited to be here. I get the name "Chair", which is one of those figurehead things that I just sit up here and smile, and Reva and Alexis and Karen do all the work, but we will try to move things along.

The other thing I just wanted to mention, while I am doing the introduction, is, as you know, when we develop something, all of us have tremendous ownership of it and it becomes part of us. So we will be the same as people are when they are on study section, necessarily critical, but always positive.

(Laughter.)

So that is how our comments will be. So we will work on that.

Why don't we go ahead, Alexis, since everybody knows you, but tell us who you are anyway?

MS. FORMAN: Good morning.

I am Alexis Forman. I am the Project Manager for the Patient Outcomes Project, and you have received lots of emails from me and lots of information. So now you

know; you are able to put the face with the name.

Thank you all for coming this morning, and thank you for all your hard work thus far.

DR. O'CONNOR: I am Richard O'Connor. I am from San Diego. I am with Sharp Rees-Stealy Medical Group. We are a group of about 400 physicians. I am Chief of the Division of Asthma there, and I am also head of the Department of Quality Management and have been involved in quality improvement and quality management improvement for many years now. I am also a member of the NCQA's Respiratory Measurement Advisory Panel.

DR. MILLARD: Well, you go from a highly-integrated healthcare system in San Diego to Dallas, Texas, which is the bastion of the last standing angry individual, I think, a Lone Ranger.

(Laughter.)

Healthcare is where I am at Baylor University Medical Center, which is sort of the flagship hospital of the Baylor Healthcare System, different Baylor in Houston. I am the Medical Director of a tertiary care referral center for asthma and COPD. We do pulmonary rehab.

I have done a lot of work with outcomes in pulmonary rehab in terms of water and land, developing water- and land-based programs, as well as worked with the Dallas

Public Schools, recently published an article on children with asthma in Dallas in 2003 who did not miss more school than kids without asthma. That certainly was a sea change of opinion, but I think points out the possibilities of aggressive and proactive school nurses.

But, anyway, I get to take care of patients and to dabble in pulmonary rehab and asthma, and have a lot of fun at it.

CHAIR YAWN: Thank you.

DR. RASTOGI: I am Dr. Amita Rastogi. I am with Bridges to Excellence in PROMETHEUS Payment System, with a grant.

We are developing a payment reform system in which we are differentiating typical and reliable care from what we call potentially avoidable complications. So, when the call for patient outcome measures came out, somebody recommended that maybe our potentially avoidable complications could serve as some patient outcome measures. So that is what I will be presenting.

My background is I am a cardiothoracic surgeon by training, actually at Mayo. I trained to be a heart transplant surgeon, apprenticed in bypass surgery, but my main focus now has been in patient quality and outcomes for the last 11 years.

DR. HAMM: Good morning.

I am Larry Hamm, and I am here representing the AACVPR, which is -- are you

ready for this now? -- the American Association of Cardiovascular and Pulmonary Rehabilitation, which is why we call ourselves the AACVPR.

My day job is just down the street teaching at George Washington University. I am a professor there in the Department of Exercise Science in the School of Public Health and Health Services, and before teaching, have about 25 years of clinical experience out in cardiac and pulmonary rehabilitation programs.

CHAIR YAWN: And our goal in the future is to have you say, "pulmonary and cardiac rehab".

(Laughter.)

DR. HAMM: Okay. I will go with that. That's fine. I am not offended in the least.

(Laughter.)

CHAIR YAWN: Thank you.

DR. HAMM: Yes.

DR. BAULDOFF: Hi. I'm Gerene Bauldoff. I am an associate clinical professor at Ohio State University. I am also a member of the Board of Directors of AACVPR and served as one of the coauthors with others on the measures that we brought forward to the Committee today.

My clinical background, I have been a nurse for almost 30 years now. My clinical background includes pulmonary rehabilitation,

coordination, lung transplant coordination, home healthcare nursing, and I served as the Nursing and Allied Health Representative on the AACVPR Pulmonary Rehabilitation Guidelines that were published in Chest in 2007.

CHAIR YAWN: Thank you.

Emma?

MS. NOCHOMOVITZ: Hi. My name is Emma Nochomovitz. I am a Research Analyst at NQF and am looking forward to hearing the conversations today.

CHAIR YAWN: And before any of us leave today, we have to be able to say your last name three times backwards.

(Laughter.)

MS. NOCHOMOVITZ: And there's a trick. It's not that hard.

(Laughter.)

CHAIR YAWN: Okay. Tell us the trick.

MS. NOCHOMOVITZ: Oh, the first syllable is "knock", like you're knocking, "uh", "mauve", like the color, although it depends on how you pronounce it, "its", like "it's".

CHAIR YAWN: Nochomovitz?

MS. NOCHOMOVITZ: Yes, yes.

CHAIR YAWN: Okay. And you wore mauve today to help us?

(Laughter.)

MS. NOCHOMOVITZ: Yes.

CHAIR YAWN: We appreciate it.

Margaret?

DR. NEFF: My name is Margaret Neff. I am a pulmonary and critical care physician at Harborview Medical Center, which is the academy hospital for the University of Washington.

I do most of my clinical work in pulmonary and predominantly critical care. I have done ARDS and substance clinical trials. Because of that interest, I went on and did some master's work in epidemiology. So, in the distant kind of cobwebs in my brain are some wonderful statistician sort of skills, but mostly now doing critical care.

Then, for the last couple of years, have been serving as Associate Medical Director for Critical Care. So, in that role, have really been building on quality improvement protocols, you know, sort of building consensus throughout the whole hospital. So it has been pretty exciting.

MS. PACE: Hi. I am Karen Pace, and I am on NQF staff. I am one of the Senior Program Directors.

The reason I am here is I was Director for the Hospital Outcomes Project that is winding down, we hope.

(Laughter.)

And also, I have been working with our Consensus Standards Approval Committee on the evaluation criteria and measure submission forms, and those kinds of things.

CHAIR YAWN: Reva?

MS. WINKLER: I am Reva Winkler.

Welcome, everyone.

I am the Senior Advisor for the Patient Outcomes Project, of which this is a part. My background is I have been at NQF for almost nine years now. So I have done any number of projects in a number of topic areas, and it is quite possible -- I keep running into old friends as we regroup committees all the time.

But thank you all for coming.

I will be helping guide you through the process within the context of a larger project. We are going to explain to you how this fits in with the larger project, as Alexis started as an intro.

Just a couple of things, comments I would like to make. Due to the expertise on this Committee and the measures we have in, this Committee is actually sort of pulmonary/ICU. It seemed to be the best fit. So just keep in mind it was because of you that that combination happened.

And then more practical details, just recall -- I don't know if you have seen Donald and our transcriber in the back, but this meeting is both being recorded and transcribed. Both the transcript and the recording will be posted on NQF's public website. So realize you are definitely on the record, as well as summaries that will be

made. So just fair warning on that.

So, again, my thanks to all of you for coming, making the trek. I know some of you have come a long ways to come be with us. So, hopefully, we can make this a very productive meeting and draw on all of your expertise to help us get the best outcome of this project possible.

CHAIR YAWN: And in addition to the more ICU, we do have, as you have probably noticed, we do have from primary to tertiary care. That was also very intentional because it is the full spectrum of care, too.

So you are going to go ahead, and Alexis is going to give us a background.

MS. FORMAN: Yes. Okay. So the goals of the meeting is to, one, get an orientation, a background on what NQF is, what do we do, and what is currently going on at NQF.

We are also going to go over the Outcomes Project as a whole, so you can understand how you fit within this large project and the work plan.

Then we are going to discuss the evaluation criteria that you all have been working on when you were reviewing the measures.

We are also going to start to begin to review the seven candidate measures that you all were assigned. Then, hopefully, if we have time, and if not, we will have to have a

conference call, to discuss the gaps in outcome measures as far as pulmonary conditions.

CHAIR YAWN: And I should tell everybody before, and I am going to apologize immediately, I got a call this morning and my flight, the one I was going to go on at 7:00-something was canceled. So I am leaving on a four o'clock flight, which means I am leaving here at 2:15, and I apologize ahead of time, but I couldn't stay overnight another night.

So we will be moving briskly through and probably will have to do some things by telephone, perhaps some of the gap identification and other things, but hope we get through all of the seven measures before I have to leave.

MS. WINKLER: Please excuse us. Donald is working on trying to get the phone line, and we are expecting some folks calling in, particularly some of the measure developer representatives to have available for us. But we are having some technical issues in terms of making that phone connection. So, that background, I apologize for, but he is trying to fix it so we can have those folks join us.

CHAIR YAWN: So, sorry, after all these interruptions, would you like to go forward now?

(Laughter.)

MS. FORMAN: So NQF is a private, nonprofit, voluntary consensus-studying

organization. We have over 400 members currently, and those 400 members are organized into eight Councils, which represent the stakeholder perspective within our healthcare system.

All right. NQF structure, we have a Board of Directors which oversees the entirety of the project. We have a Consensus Standards Approval Committee, which approves the same Committee's proposed standards. So, when the Standards Committee decides to recommend certain measurements for endorsement, they, then, go to the CSAC, and the CSAC approves that endorsement.

The CSAC also acts as sort of like an assistant committee to the Board of Directors, so the Assistant Board of Directors, with policies and procedures within NQF.

We also have a National Priorities Partnership, and that was convened in 2008 by NQF. We currently have, I think, 32 organizations that sit on this Committee, and their goal is to improve our healthcare system. So they have come up with priorities and goals and come up with some action items and ways to improve our healthcare system.

Then we also have a Leadership Network which consists of our eight Council Chairs that represent the stakeholder perspective of the healthcare system.

So that is just some brief

background on our current structure at NQF.

This slide, we just wanted to show you our new website. When you come to our website, this is the actual first page that you will see. So I don't know if some of you have been going back and forth to our website, but we recently got a new website. So we just wanted to show you our new face on the web.

MS. WINKLER: The other thing about the website is, see how it says over on the side "Enroll now"? Anybody, any public person can enroll. What that does is allow you to set up your own dashboard of things of interest within NQF, such that if you just want to follow this project, you log in and this stuff all pops up, as opposed to scrolling through everything else we might be doing. So that allows you to individualize it. So I would encourage you to check that out.

MS. FORMAN: And also, our Department, the Department of Performance Measures, if you look at the second tab, "Measuring Performance", that is how you would find out information about what is going on in our Department, our current projects, and you would also see our Patient Outcomes Project page by clicking on that tab.

Okay. So NQF has a threefold mission. Our mission is to improve the quality of American healthcare by setting national priorities and goals for performance

improvement, endorsing consensus standards for measuring and public reporting on performance, and promoting the attainment of the Nationals Goals through education and outreach.

Some of our strategic goals is to become the primary standards used to measure quality of healthcare in the U.S.; also, to become the principal body that endorses national healthcare performance measures, quality indicators and/or quality-of-care standards.

NQF will increase the demand of high-quality healthcare as well as be recognized as the major driving force for and facilitator of continuous quality improvement of the American healthcare system.

So this slide talks about our growth in our portfolio of measures. So we are looking for measures that are needed for pay-for-performance programs and, also, measures that are addressing the gaps. We will go into more detail about that when we look at our criteria as far as importance.

We are also looking at disparity-sensitive measures as well as measures of patient experience in multiple settings and, also, cross-cutting areas, which is actually part of our Patient Outcome Projects. We are looking at cross-cutting measures, so non-condition-specific measures.

Some key issues for our current NQF portfolio: do we have too many? Do we have

too few? And do we have the right measures currently for all of the conditions?

Our availability of data sources as well as the transition to electronic health records, which is currently a big issue at NQF. Our current Health IT Department is working hard, and Reva could talk more about that as far as our quality datasets.

MS. WINKLER: Right. I just want to mention NQF over the last 10 years has endorsed now well over 500 measures. This is not a static set, but a growing and evolutionary set. So we are constantly trying to look at the measures in the portfolio to ask, which ones still belong there? What are the new ones? How have things progressed?

Measures that were okay maybe five years ago probably aren't as good for us today. We are looking for other things. That is probably the underpinning for this entire Outcomes Project, is there has been an evolution in thinking. The idea of patient outcomes, measuring patient outcomes as quality and performance measures has been a little unsettling certainly in the early years of NQF. So we saw growth of a lot of process measures.

But there has been incredible change in thinking in all stakeholder groups about the benefit of outcome measures. So moving into that realm of outcomes is part of this evolutionary process of finding the

better measures.

So we are constantly remodeling that portfolio to add better, more robust measures and to weed out the ones that either no longer perform, never did much to help drive quality improvement, or are just not as good as perhaps other measures. So it is an ongoing process, that this is very much sort of in the forefront of.

DR. NEFF: One question for you along the lines of sort of that strategical being sort of the primary driver. What percent, just at a gut level, would you say of those 500, or even the last, say, year or two worth, have really taken hold to the point that it is the primary driver? Where do you think you are on that spectrum?

MS. WINKLER: That is a good question. It is actually something we are doing a very formal evaluation of to find out the amount, the measures that are being used. There are a variety of them. A lot of our measures are picked up in CMS's PQRI project. Some of our perinatal measures are now being implemented by the Joint Commission. Most of the measures or a lot of the measures coming through our hospital project you find posted on CMS's Hospital Compare.

So I can't give you a percentage or numbers, but the thing that I think is the real unknown is the fact that we get lots of questions from people, sometimes our members,

sometimes not, on projects because they are implementing within their hospital, their system; they have questions about them.

So there is a lot of use out there that is a little hard to track because it isn't something that is big and maybe done in a very local way. So we are struggling with trying to figure out the best way to get that information, so we have a better understanding of how widespread the use is.

But I do find it amazing, the questions I get from folks saying, well, we're are using your measures, but we've got a question on "X". So we have to figure out a better way to keep track of that less formal use, if you will, or that local use.

MS. FORMAN: So, at NQF, we are driving toward higher performance and we are also looking more at submitting or getting measure developers to submit composite measures. We do have a few composite measures submitted for this project, none under pulmonary specifically, but we do have a couple of composite measures.

CHAIR YAWN: Define a composite measure for us, please. What is a composite measure?

MS. WINKLER: Let me do it.

A composite measure is some combination of individual measures combined in -- you know, there are a variety of methodologies for combining them. One of the

classics is all or none. But others may be weighted averages, so that measures are in some fashion combined to have a summary score.

CHAIR YAWN: So they could be across either intermediate or true outcomes, or do you have any that are across conditions?

MS. WINKLER: I don't believe so at the moment. It would be nice if we could.

MS. PACE: They are more condition-specific, but some have combined process outcome.

CHAIR YAWN: Okay. Well, that is what I wanted to define: did composite ever include more than one condition? Because when you are taking care of people with COPD, for example, they don't have one condition.

MS. WINKLER: Right. It is not restrictive. It is just I don't think we have seen any of them, these kind of measures presented to us. There is probably some significant complexity in developing a measure like that, but it certainly would not be out of bounds.

CHAIR YAWN: Okay. Thank you.

MS. FORMAN: We are also looking at harmonization with our measures, with our current measures that are in our current portfolio, as well as measures that we will put forth for endorsement. So harmonization around age, things of that nature. So we are trying to do a better job of making sure that our measures are consistent within certain

parts of the specifications.

We are also looking to promote shared accountability and measurement across patient-focused episodes of care, as far as outcome measures, appropriateness measures, and cost and resource measures.

So our quality and disparities measurement -- and, Reva, if you can help me out with this? -- we are looking at an assessment of quality by race, economic status, ethnicity, primary language. We want that to become a part of our performance measurement.

We would like to explore direct methods for collecting this information that are efficient and effective. We are also looking to identify disparity-sensitive measures that I mentioned earlier.

MS. WINKLER: I just want to comment. One of the things that is an important issue around the topic of disparities is, when we look at risk adjustment methodologies and risk factors, what is included and what is not.

There are some fairly strong opinions among the NQF membership of whether not to include some of the classic race, ethnicity, SES kinds of things that could sort of get buried that could sort of get buried in the midst of an adjustment methodology.

So we will talk more about that. That will be an important thing, consideration

to look at and note many of the risk adjustment methods.

MS. FORMAN: And here we have an episode framework that we recently did at NQF. This is just an example of acute MI.

MS. WINKLER: Yes, the episode-of-care framework is something that is growing that NQF has been -- it started with an overall framework, and then is applying it to various very common conditions.

This is known as the NQF bubble diagram. I actually go to conferences now and see our own bubble diagrams presented by other people. So it is making its way out there.

In terms of trying to look at what is an episode of care, looking at populations at risk, patients that actually have acute and then post-acute and secondary symptomology, and where the episode would begin might be different for different conditions, whether like acute MI or chronic like diabetes.

But this is a concept that a lot of folks have embraced for a lot of different ways of trying to describe something more than the point in time, single-visit kind of approach to measurement and assessment of quality performance.

CHAIR YAWN: At some point in time, and not today, but you might want to explain why you don't use what has been for many years the standard epidemiology terminology of tertiary prevention and choose two kinds of

secondary. But we won't go there today, please.

MS. WINKLER: I will let Ellie explain that to you.

CHAIR YAWN: Yes, good.

MS. FORMAN: As I mentioned earlier, NPP, NPP has several priorities that they are looking to work with different stakeholders within our healthcare system to improve our healthcare system.

So the first one is engage patient and families in managing health and making decisions about care.

Improve the health of the population.

Improve the safety and reliability of America's healthcare system.

Ensure patients receive well-coordinated care across all providers, settings, and levels of care.

Guarantee appropriate and compassionate care for patients with life-limiting illnesses.

And eliminate waste while ensuring the delivery of appropriate care.

This was a part of the evaluation form that we all had you fill out. Staff looked at these priorities and the specific goals under each of these priorities to see if the measures actually fit the goals. It is okay if the measure doesn't, but NQF will be working to make sure that we endorse measures

along those NPP Priorities and Goals.

CHAIR YAWN: We would prefer that asthma is not a life-limiting one.

MS. FORMAN: I would hope so.

CHAIR YAWN: Especially in children. No, it is one of our measures we looked at, is asthma in children.

MS. FORMAN: And this is the same framework. What we did, we took those priorities and matched it along the different phases within this framework.

Now a little bit more about our Patient Outcomes Project. It is being funded by the Department of Health and Human Services, and we are focused on the top 20 Medicare conditions in which 95 percent of the expenditures for Medicare are being spent on these specific conditions. So we are looking to improve the outcome for patients, whether it be to reduce re-admissions or to improve the health of the patient.

We are also looking to expand NQF's current portfolio of outcome measures.

MS. WINKLER: Just to mention, when HHS came to us to start this proposal going, their focus was the top 20 Medicare conditions. However, in response, we broadened it a little bit in some areas, asthma being one of them, because that doesn't hit the top 20 Medicare list, but certainly is a huge thing for everybody else.

So it is not just those. It is

those plus a few things. We have tried to expand some of the boundaries that were logical and made sense. So that is where the asthma comes in for this particular group. There are a couple of others, but just in case anybody was wondering about asthma.

MS. FORMAN: So these are our conditions for this project. They are broken up into three phases, since there are so many conditions that we are focusing on. You are part of phase one, the pulmonary path. We will be looking at asthma measures, COPD measures, as well as some ICU-related measures.

We currently have eight TAPs total for this project, and we have three total standing committees. For phases one and two, they share one steering committee, and for phase three, mental health has a steering committee and child health has a separate steering committee. But phases one and two have eight TAPs total. So you are the first TAP to actually meet. So you are a trial-and-error group.

DR. O'CONNOR: We didn't sign a consent form. I am not sure we should be experimenting.

(Laughter.)

MS. WINKLER: Exactly.

CHAIR YAWN: They mean no IRB. Physicians and nurses are humans. We don't count.

(Laughter.)

MS. WINKLER: We'll get a waiver.

MS. FORMAN: So we also have a cardiovascular TAP that will be looking at CAB, MI, heart failure, stroke, afib. We have a diabetes metabolic TAP.

Then, in phase two, we have a bone and joint TAP, a cancer TAP, a GI TAP, infectious disease, and eye care TAP.

Then, in phase three, we will be looking at mental health, depression, Alzheimer's, and then child health as a whole. Child health is pretty broad. It is not limited.

So our project goals are to endorse additional measures suitable for public reporting and quality improvement, and we are looking at, again, cross-cutting measures, so non-condition-specific measures, as well as the measures within those conditions that I just named.

We are also looking to identify gaps in measurement. So we are going to look to you all as our TAP members, as well as our Steering Committee, to come up with or recommend potential outcome measures to fill those gaps.

MS. WINKLER: Just to mention, these goals actually are fairly equal. This particular group has a reasonable number of measures to evaluate, but there are certainly other types of outcome measures of interest,

and the Steering Committee has sort of created a bit of a framework of the different types of outcome measures. Once we get through the measure evaluation part of it, we will want to have a discussion about potential -- you know, what might an outcome measure look like around functional status or COPD for asthma? How might an outcome measure look for adverse outcomes, things like that?

And that is an important part of it because it will form an agenda that is very eagerly sought by some of the other activities we have under this very large HHS contract that we have, but also HHS itself. I talk with them twice a month for fun. This is an important aspect of it because they are in a position to direct some of the resources in the federal government to create some of these measures that are desirable. So they are looking for that agenda as well.

So your input will be very important as we sort of build that. You know, what are the measures that are really needed, but we don't have yet?

MS. FORMAN: Okay. All of our projects go through what we call a consensus development process. It is to make sure that we are looking and receiving views from all stakeholders within our healthcare system. So that is why we have a multi-stakeholder membership. What we try to do in our steering committees and TAPs, where possible, we try to

have at least one representative from each stakeholder within healthcare. So consumers, providers, health professionals, purchasers, community and public health agencies, supplier and industry organizations, health plans. We try to make sure that we get everyone's opinion because we want it to be a consensus.

Also, as you know, our formal endorsement is voluntary. We are not saying you have to use these measures. They are pure voluntary consensus measures.

So this is part of our consensus development process. As you can see, you are highlighted in yellow. Technical advisors, our panels, and workshops.

So what happens is we get a project. We start our project up by having a call for intent. So this call for intent is mainly geared toward measure developers. We say, "Hey, we have a new project. We would like for you to let us know if you plan on submitting to this project, if you have any measures that fit within the scope of our project."

Then we have a call for nominations and a call for measures. The call for measures for this project was broken up into two call for measures, one for phase one and one for phase two. That was, again, because we had so many conditions.

Our call for nominations is how we created this task, as well as our seven TAPs

and our main steering committees.

Then, after we had our call for nominations and our call for measures, you received our measures. We selected our TAP and Steering Committee, and here we are today. We are viewing the measures that were submitted.

So, after this meeting, we will meet with the other seven TAPs. Then the main Steering Committee will come together and review your recommendations and rationales behind the measures that you reviewed, and they will come to a consensus as a Steering Committee and propose consensus measures to be approved or endorsed by NQF.

CHAIR YAWN: And I am the liaison to that Steering Committee. So it won't be just what the written things are; I will be there to be able to give them some context, in addition to the staff being there.

MS. WINKLER: Right. Barbara is not a liaison. She is actually a full member of the Committee.

CHAIR YAWN: Okay, I am a member of the Committee.

MS. WINKLER: With voting power.

CHAIR YAWN: Wow. Okay.

(Laughter.)

MS. FORMAN: So, once the Steering Committee puts forth the measures that they think should be recommended for endorsement, we will draft a report and we will go out for

NQF member and public comment. The draft report will be listed on our website, and anybody can comment. That is a 30-day period.

Once we receive all of our comments from the recommendations from the Steering Committee, the Steering Committee will then meet together for a conference call to review those comments. At that time, we will go through the major concerns or the major comments that we have, and we propose action responses to each of the comments that were submitted.

Once we do that, we then draft or we edit a report, if there were any changes. Sometimes what happens is there are comments saying, "Well, oh, you didn't recommend this measure, and these are the reasons why you should have recommended this measure." "These are the reasons why you shouldn't." Either it is based on scientific evidence or guidelines of the nature that the Steering Committee or the TAPs didn't think of or didn't necessarily know.

So, then, after we look at the comments, we have the conference. We then go to voting. This is a 30-day voting period where NQF members vote on the measures that the Steering Committee has recommended for endorsement.

Then, once that closes, after that 30-day period, we then go to our Consensus Standards Approval Committee that I was

talking about earlier. Then they approve the proposed endorsed standards that the Steering Committee would like NQF to endorse.

MS. WINKLER: CSAC is essentially a subcommittee of the Board, the action of putting the final endorsement on something that was directly a Board action, but just sort of overwhelmed them with the amount of work. So they created the subcommittee to do sort of a lot of the heavy lifting for them. They, then, ratify the recommendations of the CSAC.

MS. FORMAN: Then, once we have that Board ratification, we have a 30-day appeals at that time, where anyone can submit a letter to NQF, if they didn't have a chance to get their voice heard or if they have some concerns about the endorsed measures.

So your role as a TAP member, you will provide technical input to the Steering Committee regarding the criteria within that evaluation form and within that evaluation criteria, and that is what you all have been working on. You were assigned measures, and there is a primary reviewer and a secondary reviewer, and we will be going over that information.

Also, our Chair, Dr. Yawn, will sit on the Steering Committee, and she will represent this TAP. She will be the voice of the TAP.

And for this project, as we said

previously, we will work with you to come up with suggestions on gaps in measurement and recommendations on what measures that aren't out there that should be out there and should be endorsed.

Our role as NQF staff, we are here to achieve the goals of the project and to make sure that we do the consensus development process. We will organize all meetings and conference calls. We will make sure you get the information that you need in order to review the measures. We will make sure that you get through the steps of the CDP as well as adhere to NQF's policies and procedures.

We will draft all of the reports. We will make sure that we send everything out to you before we post it, to make sure that we have collected your voice.

And we will also ensure communication amongst all project participants, including the Steering Committees, the measure developers.

Now we will look at the measure evaluation criteria. Karen has been deeply involved in this.

So our new criteria was approved by the Board in August 2008. This new criteria, it strengthened our endorsement criteria as well as clarified some issues.

So what we are looking for now, like I said earlier, is having a greater measure harmonization. We are also looking

for more outcome measures, and with our process measures, we are looking for that link between the process and outcome measures.

So this is a table that shows our old criteria and some of the new changes from our new criteria. Our importance to measure, with our measures now, you must pass, the measure must pass the importance criteria in order to move forward to be reviewed for scientific acceptability, feasibility, usability. If it doesn't pass the importance criteria, the measure stops there, and it will not be reviewed for potential endorsement by NQF.

Our feasibility, we now have a greater emphasis on health IT. Our usability, we have a greater emphasis, again, on harmonization.

Karen, did you want to add anything?

MS. PACE: Not right now.

MS. FORMAN: Our conditions for consideration, and these are the four steps that the NQF staff completed before handing out the measure to you. So we looked to make sure that we have an intellectual property agreement signed, and if it in the public domain, they, of course, don't have to have that agreement.

We also look to make sure that there is someone responsible for this measure, and that the measure will be updated and

maintained.

We also have to make sure that, no matter what, the measure both is used for public reporting and quality improvement, not one or the other, but both.

And we look to make sure the measure submission form is complete and the information is there that is needed for you all to be able to review it to the highest potential.

Also, generally, we like for our measures to be fully developed and tested. However, it is okay if they have not been tested. If the measure moves through the process, the CDP, and it is recommended for endorsement, it is only allowed a time-limited endorsement because it has not been tested, because we don't have those test results.

And the measure developer must complete testing within 24 months. So we do have a couple of measures within our Patient Outcomes Project in which they will only be eligible for a time-limited endorsement because that testing has not been completed.

MS. PACE: I am just going to say that is changing a little bit in the very near future. We will have to see what impact that has on this project.

But, generally, the Board I think discussed this week not making time-limited endorsement available for outcome measures because outcome measures are so complex,

outcome or composite measures.

So we will have to give you an update on that, but if there are -- are any of the measures today untested? I don't think so. So we will have to provide you an update.

CHAIR YAWN: But it doesn't seem to apply since we don't have any untested measures.

MS. PACE: Yes, right.

MS. FORMAN: And this is just a brief, overview view of our timeline. Our selection of our TAPs is still ongoing. We are wrapping that up. So we should have our final proposed slate out for our 14-day comment period within the next week.

The main Steering Committee met in October, the 19th and 20th. The scope of that meeting was to come up with a scope of the project, to get familiar with the project, to get familiar with our measure eval criteria, and our measure evaluation form, as well as, at that time, we were still doing some outreach to receive measures for this project. So they did an excellent job of providing us with suggestions and avenues on how to solicit more measures for our project.

In phase one, we have three TAPs. So the TAPs will meet from December through January. Like I said, you are our first TAP to meet. Then, for phase two, we have five TAPs, and they will meet from January to March.

Then, once all of our TAPs meet, and they have come up with their recommendations and rationale, the main Steering Committee will meet to review your strengths and weaknesses in your rationale behind your ratings for each of the subcriteria. They will meet on the 20th and 21st of April.

Then, if they get through all the measures at that time, we will begin to get ready for our comment period. But they will, between April, either that meeting or conference calls later, they will decide on which measures they would like to recommend for endorsement.

We hope to have our comment period begin in June of 2010 and then our member voting in August, with the Board ratification in late October.

This timeline could be changing, but, as of right now, this is how we would like our project to go.

So let's talk about --

CHAIR YAWN: So how are you doing? We've got about 10 minutes max. Okay?

MS. FORMAN: Ten minutes max.

CHAIR YAWN: I think that if you have something like that, they can read a lot of it.

MS. FORMAN: Okay. So we can skip through some things?

CHAIR YAWN: Yes.

MS. FORMAN: Okay.

CHAIR YAWN: Well, I am not saying you should skip some right now.

MS. FORMAN: I mean we can move to the evaluation process.

CHAIR YAWN: That would probably be okay, I think.

MS. FORMAN: Okay. So, as we all said, the TAP members will evaluate the subcriteria for the condition-specific measures, and the full Steering Committee will evaluate the measures and vote.

For measures that pass importance, which for this project it kind of seems like, because you have so many conditions, that it has already passed important. We know that the conditions that we are looking at are valuable within our healthcare system. So, again, the Steering Committee votes on the recommendations for endorsement.

So our four main criteria are important to measure and report, scientific acceptability, usability, and feasibility.

So important to measure, we are looking for, is this measure important enough for resources for measurement and reporting? Is there opportunity for improvement? Is there a current gap or is there a high impact within our healthcare system? And do we have the evidence to support why this measure is important?

And again, in order for this

measure to move throughout the process, it must pass the importance criteria.

For scientific acceptability --

CHAIR YAWN: And the fact that the staff decided it passed the importance criteria doesn't mean that we have to say we agree completely. We can say we think it is pretty important, but maybe not the highest level, because you are going to get the grades, is that correct? Okay.

MS. PACE: And let me just clarify. Staff don't usually make that decision. So I think what Alexis was saying, that, in general, these outcome measures and the fact they relate to the priority conditions probably indicates that they will pass that criterion, but it is your review and decisionmaking.

CHAIR YAWN: Okay. So you have looked at it and said, well, blood pressure control is not a good outcome criteria for asthma; you would have thrown that out, for example? I am just using a wild example.

MS. PACE: I don't know that we would have thrown that out. That probably still would have come to the TAP for review.

CHAIR YAWN: Okay. Well, we would have thrown it out.

MS. PACE: Right.

(Laughter.)

CHAIR YAWN: Okay. Especially with children.

Okay. Go ahead.

MS. FORMAN: With scientific acceptability, we are looking at the specifications. We are also looking at the reliability testing and validity testing, as well as risk adjustment, which is huge when it comes to outcome measures and exclusions.

So, for exclusions, the evaluation criteria requires that evidence is presented, that measure results would be distorted without specified exclusions. And if a patient preference is a consideration in the numerator and denominator exclusions, the measure should be specified so that the effect of patient preference on the measure is transparent.

Karen, did you want to add anything about it?

MS. PACE: Well, the reason we have this extra slide on exclusions is exclusions has been a growing issue at NQF. There are some measures that we get where there seems to be a tendency to try to identify every potential exception that somebody may have seen in their practice over the last 10 years. So we need to include it as an exclusion.

So the work on the evaluation criteria really stressed that exclusions should be limited. They should be evidence-based. There should be a good rationale for it, rather than trying to think of every possible thing that could happen in a

particular situation.

DR. O'CONNOR: I am trying to think of examples. Pediatric immunization, for example, if parents refuse an immunization, would that be considered an exclusion?

MS. PACE: Well, the discussion about patient preference, and this came up specifically in an immunization project that we had a year or so ago --

DR. O'CONNOR: I'll bet it did.

MS. PACE: That a patient preference is going to be one of those issues that it really should be transparent. So, in our immunization project, how that came out, the committee actually recommending standard specifications, is that would be a numerator category. So the numerator actually included patients that were offered the immunization and refused, patients that actually received the immunization, and patients --

DR. O'CONNOR: So it would be part of the numerator rather than --

MS. PACE: Right, exactly.

DR. O'CONNOR: -- subtracting from the denominator?

MS. PACE: You know, patient preference is one of those things, and there was a lot of sentiment on that particular committee that it is easy to kind of check that box or lean in that direction. So they just want it to be transparent, if it is really an issue for a particular measure.

CHAIR YAWN: And it also frequently you can make feasibility much more complex if you start putting a bunch of exclusions in.

MS. PACE: Right.

CHAIR YAWN: So I think that is another reason that I am sure you are thinking about exclusions. The more you put in, the more difficult it is to --

MS. PACE: Right. The more data you have to collect.

CHAIR YAWN: -- to operationalize that measure.

MS. PACE: Exactly.

MS. FORMAN: Okay. When looking at our usability criteria, is this measure meaningful? Can it be used for public reporting and quality improvement, not solely one or the other?

Then feasibility, can this be done without undue burden?

CHAIR YAWN: Of course, you define undue burden by saying, "hmmm" --

MS. WINKLER: Actually, one of the things that is particularly timely on this is the adaptability or the ability to use HIT in this in terms of feasibility, either existing electronic data systems or what is your plan to embrace or transition to EHR use and data from readily-available electronic sources. That really is a major focus because the idea of paper chart review is pretty much no one is ever going to do it. So we need to move on.

MS. PACE: And actually, the emphasis on importance to measure and report also has some reasoning behind it related to feasibility. I mean there's only so much resources to go into data collection and reporting. So the idea is to really try to focus on those things that are going to have the biggest impact on overall improvement in healthcare and health.

So it is not just, is it important to do in your everyday practice? I mean there's thousands of things that people have to do. So we are really trying to focus in on resources used for data collection, data reporting, to those that are going to make the biggest difference.

MS. WINKLER: Barbara, just one thing Donald is saying. We have folks on the phone, and the question is, who is there and can you hear us? So I heard a couple of folks.

Francois, are you on the phone? Can you hear me?

MR. DE BRANTES: Yes, I am.

MS. WINKLER: And you can hear me?

MR. DE BRANTES: Yes, I can.

MS. WINKLER: Thank you.

Is anybody else on the line?

(No response.)

Not admitting it or can't hear me.

CHAIR YAWN: Francois, can you introduce yourself, please? This is Barbara

Yawn. Everybody else has introduced themselves. We would appreciate a short introduction. We have part of your name anyway -- who you represent, and a sentence or two about your background.

MR. DE BRANTES: Well, it is sort of broad. I am the CEO of Bridges to Excellence at PROMETHEUS Payment. We are here today to present a couple of measures on complications of care.

My background is I have been working on payment reform and incentives for quite some time, starting a few years at GE as a leader for a healthcare initiative and, more recently, on a full-time basis, in this not-for-profit organization.

I have worked with the NQF before, in particular, as a member of the Steering Committee on Efficiency in Episodes of Care. That is about it.

CHAIR YAWN: That is great. Thank you very much.

Go ahead. Alexis, do you have more?

MS. FORMAN: No.

CHAIR YAWN: Okay.

MS. WINKLER: At this point, what we need to do is just allow each -- we have measures from three different measure developers. We have representatives from two.

MS. FORMAN: Two.

MS. WINKLER: I don't see CSF. He

is not on the phone. Okay. Well, then we will go to two of them.

As an introduction to the measures that you are going to be looking at this morning, the measure developers can provide you some background on how they developed them, why they developed them, what was the circumstances around it.

Who do you want to have go first, Alexis?

MS. FORMAN: Francois can go first since he is on the phone.

MS. WINKLER: Okay.

MS. FORMAN: Because I know his schedule is pretty tight.

MS. WINKLER: All right. Why don't we let Francois continue, and he can explain the background to the measures that they have submitted to us.

Francois, are you there?

MR. DE BRANTES: Yes, I am. Okay.

CHAIR YAWN: Okay. So could you briefly discuss that?

MR. DE BRANTES: Just a few words of background because I know that talking on a cell phone over a conference line is not ideal at all.

So, about three and a half to four years ago, we started this work around definitions of episodes of care for various chronic conditions, procedures, and acute medical events. That process turned into what

is now known as the PROMETHEUS Payment Model.

As a part of that, the charge really of the team was to look at and understand the different components of costs of care, and to base episodes on what we could, and not just us, but through clinical working groups and teams of medical experts, delineate the appropriate, typical, normal care for each one of these episodes.

As we did this, one of the charges of each one of the working groups, clinical working groups, was to identify the potentially avoidable complications of care that would occur within a disease stage procedure, acute medical events.

Since then, we have turned these definitions into formal, delineated, complication-of-care measures, which is what my colleague, Amita Rastogi, is going to present later today.

And we had an opportunity to run the definitions of these complications of care on several national and regional commercial claims databases in order to ascertain both the feasibility of the methodology, as well as its reproducibility in different datasets. At each step, we have gone back to either physicians, hospitals, and communities or the working group of members to look at and validate the outputs. So that, ultimately, the original definitions around what constitutes the potentially avoidable

complications of care were, in fact, illustrated by the data modeling.

So, while we certainly and I certainly don't pretend that we have had an opportunity to thoroughly statistically set, validate, and analyze the robustness of the measure, I do think that we have at least gone through a series of feasibility testing and field testings, if you will, the results of measuring these definitions around complications of care, and to ascertain both their prevalence within the delivery system and their underlying cost.

So that is a broad brush. Of course, there is a lot of work underlying the definitions of these complications-of-care measures. That really is what my colleague is going to focus on.

But let me pause here and see if there are any specific questions that the Technical Advisory Panel would like me to address.

CHAIR YAWN: I have a feeling those will be much more specific questions as we look at the measures.

So does anyone have any overall questions now?

(No response.)

I don't think so.

So the next step, Reva?

MS. WINKLER: Yes, I think we are ready to kind of move on.

Do we have anybody from San Francisco on the line?

MS. FORMAN: No.

MS. WINKLER: No? Okay. So I guess our friends at the end of the table who are here with us get the first shot.

What are the numbers of the measures?

MS. FORMAN: Nineteen, OT1-10-09.

DR. NEFF: Can I ask one question, just about the sort of expectations for today as well?

MS. WINKLER: Yes.

DR. NEFF: Since we have all been, obviously, working in our own little silos on these measures, and now we are going to get together, undoubtedly, there will still be questions that we have and have to maybe dig a little bit deeper.

Is the expectation, then, sort of to get as much as we can out through today and then still do conference calls back and forth to finalize? I mean just to get a feel for it because it will be complex to nail it all down.

MS. WINKLER: It will depend very much on how these conversations go.

DR. NEFF: Okay.

MS. WINKLER: But one of the reasons we asked you to try to get the information to us is so that we could present like both reviewers' perspectives and discuss

where they are the same, everybody agrees that they are different; what are the issues? How might we resolve it?

At the end of the day, what we want is one of these evaluation forms for each measure reflecting the evaluation of the subcriteria for the TAP as a whole. So the first step was two folks got to really look into each measure to present, to discuss, and lead the discussion. Then we will come up with sort of a final version out of the TAP.

Whether it will be totally completely today or not remains to be seen, but, certainly, there is plenty of time for us to do followup as needed.

CHAIR YAWN: And I think that you will find several of the measures are really quite similar. So, when we look across, I think that we will be able to sort of say, yes, what we said before applies to this one also. So I don't think it is quite as onerous as it looks like from having that many different measures.

So I have to step back, and I'm sorry, I should have asked you on the phone the other day. So how much do we ask the developers to give us before we start our review? Because we did have the materials. So I am just asking.

MS. WINKLER: Yes, I think it would be nice, just as we did with Francois, to ask if they want to give a few minutes of just

introduction and background of the measures. Then we will start discussion.

CHAIR YAWN: Okay. So you were talking about the other group who is present. That was my question. You are not asking, because you are Bridges to Excellence also, you are not asking her to tell us more about it? You are asking the other group?

MS. WINKLER: Right.

CHAIR YAWN: Thank you. I was confused.

MS. PACE: I mean, just in general, we usually ask the measure stewards or developers to give a brief introduction, but this is the TAP's meeting, and they are here to respond to questions or provide more information that you might ask for.

CHAIR YAWN: Okay. No, I misunderstood. I thought you were asking for just Excellence to give us even more information.

MS. PACE: No, no, no.

CHAIR YAWN: So I understand now.

Could you please give us --

MS. PACE: Give us your measures.

CHAIR YAWN: Yes.

DR. HAMM: Yes, I would be happy to. Thank you.

I am just going to open up with a few general remarks, and then Gerene is going to more specifically address some remarks about the measures.

AACVPR is a group of about close to 3,000 multidisciplinary healthcare providers who work in the field of rehabilitation. I just want to thank you, on behalf of the Association and on behalf of all of our members, for reviewing these measures.

They are very important to us because I am sure you realize that the rehabilitation is not the pizzazz and upfront area in healthcare. But, to us, it is very important, and we are pleased to see that our measures have gotten to this point in the process.

Also, for another reason, it is important to us. Actually, for two reasons. The first being that last spring we got two time-limited endorsements for some cardiac measures relative to rehabilitation, and we were very pleased about that. This sort of is the bookend for us on the pulmonary side of things, and we hope that we are successful here as well.

I think something else that might be of interest to you to know is that there is new legislation going into effect January 1 that makes pulmonary rehab a guaranteed benefit for Medicare subscribers, which was not the case prior to this new legislation.

So we expect quite an uptick in participation in pulmonary rehabilitation programs around the country, and it would be very nice to have some quality measures in

place right from the very beginning of this increase, what we anticipate to be increase in activity for our services.

So, Gerene?

DR. BAULDOFF: We have submitted two outcome measures to be evaluated. The first is capacity in COPD before and after pulmonary rehabilitation.

And the instrument of measurement that we propose to use in this is the six-minute walk. The six-minute walk, we can get into more detail, is a very well-validated, very well-tested, long history of use in patients with COPD and in pulmonary rehab.

The second measure is health-related quality-of-life outcome measure for patients with COPD who participate in pulmonary rehab. In that, we selected a single instrument, the chronic respiratory disease questionnaire, for our description, although we have others, if there is interest in us expanding that program.

The rationale for these two outcomes comes out of multiple sets of guidelines that have been generated that are all evidence-based. Of these, these are the two outcomes that have the strongest evidence behind them for evaluation.

Also, as part of the certification process for programs through AACVPR, these are the type of data that is collected by our certified pulmonary rehab centers that provide

that information.

Finally, to address usability, there are plans in developing a pulmonary rehabilitation registry that is currently being started. We have already begun work and have our cardiac rehabilitation registry ready to be up and running. We are also now doing, as Larry mentioned, the bookend for pulmonary rehab.

So neither of these instruments we consider to be tested. So we just wanted to clarify that measure for you.

MS. PACE: Neither of the measures or the instruments?

DR. BAULDOFF: Neither of the measures. All the instruments have both been very well-tested and very well-validated.

DR. NEFF: So using this as an endpoint for rehab has not been --

DR. BAULDOFF: It has been described in the literature, but I guess that I am uncertain as to what you would call tested. There is lots of literature to say both of these, we are able to show these improvements in both of these across multiple studies. However, in this format, it has not been tested yet.

CHAIR YAWN: I know that is a very difficult thing to say, has it been tested or not? Because I know this is an outcome of many of the studies which are used to give the evidence behind recommending pulmonary rehab.

DR. BAULDOFF: So would you consider these measures to be tested?

MS. WINKLER: I think that is the guidance from you all.

CHAIR YAWN: Yes. As we go through the criteria and look, I think we will have to decide whether we think they have been. Sometimes something has been tested 40 times; it has just never been called "apple pie" before. So that is maybe what we are deciding, and other things we don't even know what apples are. So we will have to decide that.

But thank you for being as conservative as possible in your definitions. We appreciate that. We may be less conservative or more conservative. Who knows? We'll see.

All right.

MS. FORMAN: Do you want to get started --

CHAIR YAWN: Sure.

MS. FORMAN: -- with our first one?

CHAIR YAWN: I think it is time.

MS. WINKLER: Who are the two reviewers for 019?

MS. FORMAN: For 19, we are going to start with Lewis, and I have his because he is out of the country. So his evaluation is posted. And Dr. Millard.

DR. MILLARD: All right, and this is for HLQR --

CHAIR YAWN: Okay. Well, now you just threw me a curve.

MS. FORMAN: It is health-related quality of life in COPD patients before and after pulmonary rehab.

DR. MILLARD: Right. You know, what is interesting is that, when I looked through all the current NQF-endorsed pulmonary and respiratory stuff, this is almost putting the cart before the horse, which is nothing to say because I think AACVPR, our group, just got certified again by the program. So we like it.

(Laughter.)

But there is no guideline that says when pulmonary rehab should be used. In many ways, if you want to -- I mean the leading question of importance is, when should pulmonary rehab be utilized? Because that is really more important in the long run in terms of health outcomes than whether or not we use the six-minute walk, constant low endurance, SGRQ, CRQ, whatever.

So I would add that as an initial sort of statement. I really think the most important issue is the initiation point of rehab because that will change, alter the equation of outcomes. Because if you don't get into pulmonary rehab, it doesn't matter what happens to either quality of life or exercise tolerance.

MS. PACE: Unfortunately, this

project was focused on the outcomes.

DR. MILLARD: But I think that should be said upfront because that is really --

MS. PACE: Sure.

CHAIR YAWN: So that would be a processing issue, actually.

MS. PACE: Right.

CHAIR YAWN: For COPD. The number of people at each stage initiating pulmonary rehab, but that is a process measurement, and we are now trying to --

DR. MILLARD: Yes, but I mean to reduce potential avoidable comp PSEs will be, pulmonary rehab will be one of the tools.

CHAIR YAWN: Yes.

MS. WINKLER: As you make all of these comments, there are places for us to capture that and put that in. So we can say this is great, but the most important thing is the selection criteria, who goes into pulmonary rehab.

DR. MILLARD: Right, and I just needed to say that.

MS. WINKLER: Got it.

DR. MILLARD: Now you all need to guide through the sort of format of this process. I have not used that syntax before.

So we go to first --

CHAIR YAWN: To the importance of the measure and the report.

DR. MILLARD: Yes, the percentage

of patients enrolled in pulmonary rehab who are found to increase health-related quality of life. I mean that is an interesting outcome because that requires, that I think assumes that pulmonary rehab is going to increase the quality-of-life score in a certain percent of people.

We know that, in general, it does. I just don't know what the standard deviations are in terms of what percent of people actually -- I mean, in our program, we use SGRQ, the average increase is eight points, which would correlate with a 1 point CRQ score, which is what the guidelines note. But I don't know what the confidence limits are to saying, okay, we are going to specify a percent.

CHAIR YAWN: So you think that this measure should have a percent attached to it?

DR. MILLARD: No, I don't. I don't, no.

CHAIR YAWN: Okay.

DR. MILLARD: But, in fact, it has one. It says the description is the percent.

CHAIR YAWN: Correct, but it doesn't say 50 percent, 75 percent, 3 percent.

DR. MILLARD: But we are going to base quality on what the percent of -- are we setting a quality score at some point and saying, well, if you don't have "X" percent of people, then you have not reached your goal?

MS. PACE: So one of the things we

probably need to do is we really need to kind of go through these criteria, and that will be kind of a measure construction issue. So the first question about this is measuring the outcome of health-related quality of life in COPD patients.

So the first question we would like you to look at is, does this meet our criteria for importance to measure and report?

DR. MILLARD: Yes, right. The answer is yes, but I would say partial because I think my own bias is -- and I am not sure that CRQ has been shown to be superior to SGRQ -- this SGRQ central respiratory questionnaire, and in the pulmonary rehab world I always understood most pulmonary rehab programs actually use the Saint George's Respiratory Questionnaire and not CRQ.

DR. BAULDOFF: Actually, that is not true. The reason I say that is that the SGRQ is so difficult to score and interpret. It requires extra programming, and CRQ is much more straightforward in its utilization.

Actually, out in the clinical programs, most programs probably use the SF-36, but that is generic --

DR. MILLARD: Well, yes, but we gave that up.

DR. BAULDOFF: And the other issue with the SGRQ is that, when pulmonary rehab is being measured over a three-month period, the SGRQ truly was originally designed by Dr.

Jones to measure changes over a full-year period. And it has been found in other studies not to be quite as sensitive. So that was why we went with CRQ.

DR. MILLARD: It wasn't sensitive when it was sent home with the patients at home, as opposed to when it was at the site.

And the CRQ requires -- there are different methodological ones. My only point would be I think that I would like to equalize the syntax between SGRQ and CRQ.

DR. O'CONNOR: What percentage of your rehab programs use these two instruments, would you say? I mean because that is an important consideration. If you have got a significant proportion of your certified programs not using the measure that you propose, there is going to be some difficulties.

DR. HAMM: I think there's two ways to look at that question. One is that, if we are dealing primarily with certified programs, that requirement can be put into the certification requirements and, in a sense, sort of push the issue to help increase the data collection.

The actual percentage of programs that use the questionnaire, I don't have that number, either.

DR. NEFF: But, in a way, you could almost say, if the goal here is the monitoring of this in pulmonary rehab, that you could

almost allow, if you think there's some equivalence, either/or, I mean to some extent, I mean that would be another way to swing this. If you think it is okay, then focus on that goal.

CHAIR YAWN: You have to have a validated measure.

DR. NEFF: Right, but --

DR. MILLARD: Both of them are.

CHAIR YAWN: I know they are, but that is what I am saying. If you have two equally-validated measures, then you can say either/or as opposed to SF-36, which none of us would say is a validated measure for this.

MS. PACE: Right, but the issue here is this is a measure of outcome, meaning looking at the change in scores.

CHAIR YAWN: Right.

MS. PACE: So, in order to construct this measure that you could use any validated instrument, you would first have to show that this change in scores that you would get by using any of those are similar. I mean we are talking about measurement here, so it has to be standardized.

So, as soon as you start saying you can do this or you can do that, you are getting away from a standardized measure, unless you can prove equivalency.

CHAIR YAWN: But they have both been used in the studies that show improvement and have been calibrated in this particular

case.

MS. PACE: So the two instruments you can --

CHAIR YAWN: You can use those two, I believe.

DR. MILLARD: You can almost do the slash. If you say, "CRQ/SGRQ", then you are going to include --

CHAIR YAWN: In this particular situation, I think that is true.

DR. BAULDOFF: And in fact, the SGRQ was the health-related quality-of-life instrument that was used in the National Emphysema Treatment Trial. So that has been the largest study in which there has been randomization for surgery, but in which pulmonary rehab was used.

MS. PACE: But, also, just to back up, that is getting into the specifics of the measure. So, under importance, what we want to know is, you know, is this a high impact area? Is there a performance gap in terms of patients achieving health-related quality of life, and the evidence that this can be impacted?

Certainly, it is an outcome measure which is something that we are interested in. Is it relevant to this particular patient population? And hopefully, there's some things that can actually influence that.

So those are your kind of first set of questions. Then we get into the

reliability and validity of the instruments.

CHAIR YAWN: So I think that is what we do. Let's go way back to the beginning. Can you tell us your strengths and weaknesses that you developed under this very first question of importance, please? Did you list strengths and weaknesses?

DR. MILLARD: The strength is that both the quality-of-life measurements are well-validated in the literature of pulmonary rehabilitation. The weakness is that the definition of pulmonary rehabilitation is not uniformly assumed.

CHAIR YAWN: Okay. Although there are guidelines specifically for it.

DR. MILLARD: There are, but we don't have any -- I mean we are assuming the guidelines. Do you see what I mean?

CHAIR YAWN: Uh-hum.

DR. MILLARD: In other words, we are assuming that we all agree on what pulmonary rehab is, and we haven't, but --

CHAIR YAWN: Okay. So it is way back to the beginning of --

DR. MILLARD: Yes.

CHAIR YAWN: -- the importance of pulmonary rehab, as you said, has not been widely available, but will be more widely available, and we are not entirely sure that we have a definition that is universally accepted of what it is that is mainly of importance. Okay?

DR. MILLARD: Although the new guidelines, I mean CMS finally deciding, well, Congress deciding to fund pulmonary rehab, that is now -- the new, what look it is, is out. It was published and it will be in --

CHAIR YAWN: Yes.

DR. HAMM: I would just quickly add that CMS has now completed their announcement of final rules for what is going to be paid for.

CHAIR YAWN: Right.

DR. HAMM: Which is going to drive program models.

CHAIR YAWN: Right. Okay. So you are saying that you believe this has high importance potentially?

DR. HAMM: Yes.

CHAIR YAWN: And then health-related quality of life is a very important outcome to patients?

DR. HAMM: Absolutely.

CHAIR YAWN: Okay. And is one worth assessing for improving quality of care, as to whether or not we improve the patient's quality of life?

DR. HAMM: Yes.

CHAIR YAWN: Okay. Is that sort of a summary, then, under importance? And you said it meant partially -- were you the secondary reviewer, Margaret?

DR. NEFF: No, not on this one.

MS. FORMAN: No. It was Dr. Lewis.

CHAIR YAWN: Oh, this is, I'm sorry, Dr. Lewis.

MS. FORMAN: What I have on this screen is his review. For high impact, ensuring a high impact of healthcare, for 1a, he said completely.

CHAIR YAWN: Okay. Did he give us strengths and weaknesses?

MS. FORMAN: Yes.

CHAIR YAWN: Okay. Could we look at those?

MS. PACE: What do we do for the other criteria? Just running through the --

CHAIR YAWN: I mean this is all what the staff did. So I want to see what --

MS. FORMAN: No. No.

CHAIR YAWN: Oh, I'm sorry.

MS. FORMAN: That is the reviewer.

MS. PACE: No, staff hasn't done evaluation.

CHAIR YAWN: No, I'm sorry. You are absolutely correct. I just wanted to see what he gave us.

MS. FORMAN: So he has got "C", again, completely, for opportunity for improvement, demonstrating a performance gap.

CHAIR YAWN: Okay.

MS. FORMAN: Then evidence to support, he also has "C".

CHAIR YAWN: Okay.

MS. PACE: And he had a comment?

MS. FORMAN: Yes, down at the

bottom.

CHAIR YAWN: Yes, that is what I am really interested in, is his comments.

Okay, so here are his comments, the strengths: "These new requirements include provisions for number of sessions, required elements directing physician supervision, and other components."

So he is saying he believes that CMS's new criteria will help define what pulmonary rehab is, at a minimum, anyway. Okay.

"Clearly have major impact on a large segment of the pulmonary disease population. Known therapeutic performance gaps will be impacted. Compelling high-grade evidence to support the benefits of well-designed and performed rehab programs."

The weaknesses: "Likely to be a limited number of programs that could help close this performance gap." And I am reading this because I think it might be hard way back there in the back.

"Cost of implementation is really unknown at this point. So, even though CMS has reimbursement plans, we all know it won't be for everyone."

"Logistics. In terms of requirements for greater physician involvement, which is a good measure, but may deter new program development."

So he has some concerns about --

that is really the requirement of CMS. That is not your requirement in the measurement. It is not a specification of the measurement.

DR. NEFF: But I guess by us saying that we concur --

CHAIR YAWN: Yes.

DR. NEFF: -- those are valuable things to have in terms of assessing pulmonary rehab. It would lend weight to that or not, if we said, oh, we don't care. Throw in a better program to do. Do you know what I mean? We are sort of buffering up what CMS is saying, actually.

CHAIR YAWN: So we are saying that we are accepting CMS's definition?

DR. NEFF: Yes. On our own terms.

DR. MILLARD: The problem is accepting CMS reimbursement.

(Laughter.)

Ironically, CMS has, in a wonderful review of pulmonary rehab and the effectiveness of it, two years ago, two or three years ago, they said there's no question that pulmonary rehab is an effective intervention; it is just not a covered benefit. So Congress, finally, when somebody wasn't looking, passed it.

CHAIR YAWN: All right.

DR. HAMM: No, no, a lot of hard work.

(Laughter.)

CHAIR YAWN: So we have someone

saying, one reviewer saying complete and one reviewer saying partial for the importance. Should we go on --

DR. MILLARD: Mine is complete.

CHAIR YAWN: Okay.

DR. MILLARD: My only partial had to do with the consideration of using only CRQ or the emphasis on CRQ.

CHAIR YAWN: Okay.

DR. NEFF: And would those sort of concerns, actually, just in terms of this process and structure, be more in the scientific, in the second set, since we are going to have --

CHAIR YAWN: So the importance is really complete, and we agree --

DR. MILLARD: We are all in agreement.

CHAIR YAWN: Okay.

DR. MILLARD: Although the one thing, the opportunity for improvement, the citations for data on performance gap, at least that first one, co-morbidity and mortality, COPD-related hospitalizations, there is no discussion of performance gap in that reference.

I am not sure that there is the literature on performance gaps in pulmonary rehab. I mean it is to be determined because, historically, it was such a hot sort of scattergun of who could get pulmonary rehab, that we have no way of knowing what the

previous performance gaps were.

CHAIR YAWN: Well, except I thought the performance gap was really sort of between people who didn't get any and people who got some or got not some, who got it and who didn't get it.

DR. MILLARD: It was all driven by whether or not CMS reimbursed in the LMRP or not.

CHAIR YAWN: Well, except the randomized control trials were not on that.

DR. MILLARD: Right.

MS. PACE: But this is about patients in pulmonary rehab.

CHAIR YAWN: Yes.

MS. PACE: So is there a performance gap in --

DR. MILLARD: We have no way of knowing.

MS. PACE: So this really would not be "completely", for this particular 1b then, because there is no data that was provided, right? Or are you aware of any data that there is variability in achievement of --

CHAIR YAWN: We do not know, yes, I agree, because nobody did a randomized control of halfway-done pulmonary rehab. So it is a little hard to do that one. Okay.

DR. O'CONNOR: I think, from your viewpoint, the performance gap is all or none.

DR. MILLARD: Yes. Referral and completion.

DR. O'CONNOR: Yes.

MS. PACE: Is that what this is measuring? Because this is measuring those in it.

DR. MILLARD: Exactly.

MS. PACE: So it is not that question.

DR. NEFF: So that is the question. Should this measure of improved health-related quality of life be restricted to those getting into pulmonary rehab or should it be for all pulmonary patients? I am just asking. It is not my field.

CHAIR YAWN: Not all pulmonary patients. All patient --

DR. NEFF: I mean the COPD patients.

CHAIR YAWN: All COPD patients of a certain stage is who should be the comparator group, is what I think people are asking.

DR. NEFF: Right.

CHAIR YAWN: Do we compare people only in pulmonary rehab and say this is a good pulmonary rehab program because 90 percent of their patients achieve this improvement in quality of life versus this pulmonary rehab program where only 30 percent do, or is it this is a 90 percent improvement, and when they don't get anything, 3 percent improvement?

So that is the question, is: can this become a broader measure than actually

what you have said? Could it include patients in pulmonary rehabs, to say something about the different kinds of pulmonary rehab and something about not getting it at all? But that is not what you proposed.

MS. WINKLER: I was going to say that sounds like a different measure.

CHAIR YAWN: Yes.

MS. WINKLER: Maybe a desirable one, but a different one.

CHAIR YAWN: But a different one. So that might be a gap for the future.

DR. NEFF: Well, because you guys have already sort of bought into the concept that pulmonary rehab is a value for these people. Then it is a matter of tracking how well the program works in terms of outcomes.

So you've bought Part A. Then you wouldn't be looking at it in Part B. So you would have to restructure the whole shooting match.

DR. HAMM: Well, it is sort of interesting because what you are talking about right now, basically, referral, too, I mean that is the denominator, is all that population out there. Then the numerator becomes those people who get referred and participate.

DR. NEFF: Yes.

DR. HAMM: That is exactly where we went with the cardiac side of things, referrals from inpatient hospital programs as

well as outpatient physician offices.

As you can imagine, there are quite a few pitfalls around trying to get that denominator. But, yet, with cardiac, there is that acute event usually; whereas, with pulmonary there isn't.

CHAIR YAWN: All right, but let's take a step back. Now we are saying, is this measure important to look at for people who at least are referred or begin pulmonary rehab? That is the only group we are now discussing. Is this an important measure?

DR. MILLARD: Again, not referred.

CHAIR YAWN: Again, that is fine. That is fine.

MS. PACE: In a rehab program.

DR. MILLARD: Yes, that you are in it. If you are referred --

CHAIR YAWN: Yes, that is fine. I'm sorry. People who are in pulmonary rehab, is this a valuable measure for everyone in pulmonary rehab? Does this meet the complete importance?

DR. MILLARD: Yes, but there are -- these citations do not really support that there is a performance gap. There has been very little work around that.

CHAIR YAWN: Right.

MS. PACE: So there is really no information about performance gaps?

DR. MILLARD: There is limited information.

CHAIR YAWN: So we believe this will provide a lot of potentially useful information, but we can't be sure. Okay. So that is why you might have said partial instead of complete, just because --

DR. MILLARD: Yes, on that, yes.

CHAIR YAWN: -- we truly believe, but we don't have a lot of evidence out there. Okay.

Let's go on next to the second measure, which is --

MS. WINKLER: Section.

CHAIR YAWN: Yes, section, that is what I meant, the scientific part. I'm sorry. We are having trouble with words today. I am having trouble with words today.

So the scientific?

DR. MILLARD: The scientific part is, again, now the numerator and denominator is the first one.

MS. PACE: Right.

DR. MILLARD: Is that the first issue?

MS. PACE: Exactly.

CHAIR YAWN: Yes.

MS. PACE: The specifications, which is quite long, but yes.

DR. MILLARD: And I think that, again, my concern primarily was CRQ versus SGRQ, that we needed to broaden that to include both CRQ -- and I am much more familiar with SGRQ than CRQ because we were in

that trial rehab center. So we did lots of SGRQs and had to decide what to do, and chose, actually, the SGRQ over the CRQs because it was, my staff said, easier.

CHAIR YAWN: Your staff would be among some of few, I believe. From a primary care perspective, we would never choose --

DR. MILLARD: SGRQ?

CHAIR YAWN: No.

DR. NEFF: But it sort of factors into that sort of feasibility issue --

DR. MILLARD: Yes.

DR. NEFF: -- that you allow --

CHAIR YAWN: Yes, but you can expand this numerator to say one point on the CRQ or I thought it was five on the SGRQ, but you are saying it is --

DR. MILLARD: Well, .5 on the CRQ is the minimum clinical difference.

CHAIR YAWN: Right.

DR. MILLARD: In terms of the numerator, I would like to see, define a positive improvement as .5, not as 1, simply because I would like to define it as the minimum.

CHAIR YAWN: We can't really change measures that way. We have to accept what they are suggesting when we rate it. I mean you can't change some of these measures --

MS. PACE: Right. I mean I think that there is certainly room for discussion around that, but this should relate to what is

the evidence that it should be one or the other.

DR. NEFF: Well, but wouldn't we be in our ultimate recommendation saying this is what we don't -- we don't accept it like this, but would like this? I mean, wouldn't that be basically providing the feedback --

MS. WINKLER: Yes.

DR. NEFF: -- and then they kind of know what path they are on?

MS. WINKLER: Right. Exactly.

CHAIR YAWN: So you are suggesting that you would like to see it, rather than saying 1.0, say the determined clinically-minimal difference?

DR. MILLARD: Yes.

CHAIR YAWN: Because that isn't, then, specific to any one of them, to either of the two --

DR. O'CONNOR: Right.

CHAIR YAWN: -- validated measures.

DR. O'CONNOR: I am not sure if this is the right point for this question. How easy is it going to be to collect data in terms of electronic retrieval?

CHAIR YAWN: I think we need to wait for the feasibility phase, if you wouldn't mind.

DR. O'CONNOR: Feasibility? Sure.

CHAIR YAWN: Okay?

DR. MILLARD: And the numerator is just the number who participated in the PR and

found an increased healthcare quality-of-life score by the minimum significant clinical difference, as being an end of PR, regardless of --

CHAIR YAWN: Well, it says the time period should be no more than three months.

DR. MILLARD: Right.

CHAIR YAWN: And we know that those scales can measure change within three months.

DR. MILLARD: Yes.

CHAIR YAWN: So we are okay with that? All right.

DR. MILLARD: Numerator details I guess would be with just the CRQ?

CHAIR YAWN: Which can be expanded to add the SGRQ for the same way.

DR. MILLARD: The denominator statement: all patients with COPD during the reporting period who are enrolled in a PR program. So, again, I think that is enrolled, not referred to, which is different from what you have done in cardiac rehab, is that correct?

DR. HAMM: Yes, that is correct.

DR. MILLARD: Then I agree.

Now the next target population range, and this also relates to harmonization later on down the road, is persons greater than 20 years of age. This pops up, several different parts, on all the COPD processes. Most guidelines talk about 40 and above or above 40, and this is 20. There is no one

that is greater than 18, I think, on the PAC reduction.

So there needs to be a uniform age --

DR. NEFF: Harmonization.

DR. MILLARD: Yes, target population range harmonization.

CHAIR YAWN: Well, and it needs to be clinically-relevant. I think if we started telling most physicians we are going to look at COPD pulmonary you have at age 20, they would probably think we might have lost out minds.

DR. NEFF: And I guess you could have different age cutoffs with different allowances for chronic diseases that are in young adults.

CHAIR YAWN: Yes, but that is going to be risk adjustment, I think.

DR. NEFF: Yes. What was the rationale for the 20 --

DR. BAULDOFF: The rationale for such a low entry age was to be able to include those patients who have very early onset. I appreciate it is 1 percent probably of the population.

CHAIR YAWN: But having early onset COPD --

DR. BAULDOFF: Right.

CHAIR YAWN: -- it is even less than 1 percent.

DR. BAULDOFF: I can tell you

clinically that I have had patients over 26 years old --

CHAIR YAWN: Oh, of course, yes.

DR. BAULDOFF: -- in my rehab program.

CHAIR YAWN: I understand that, but aren't they kind of the exception perhaps?

DR. BAULDOFF: Yes, they are the zebra, yes.

CHAIR YAWN: Obviously, as a primary care physician, I am looking for horses.

DR. O'CONNOR: Have these instruments been tested in alpha 1 antitrypsin deficiency?

DR. BAULDOFF: I believe they -- I don't have the literature in front of me.

DR. MILLARD: I think for the greater purpose of harmonization, I would say 40 and above. That would be my recommendation. And you will capture 95, you will --

CHAIR YAWN: Oh, you will be much more than 95, I suspect.

DR. NEFF: And what is the process for allowing for case-by-case exception? I mean I know you brought up the whole exception issue, but if this were an adopted guideline, and then recognizing that there would be some legitimate fallout, you know, sort of misses where we are setting the bar, is going to miss some people that would legitimately, just kind

of do that or --

MS. WINKLER: Remember that these are measures of performance of the facility. You are not talking about taking care of every single patient.

So, if, indeed, it is the exception, how will that really impact the overall assessment of quality of care provided by that facility if they are not included?

DR. NEFF: Right. I guess you sort of feel the weight of any sort of guideline recommendation because you know how easily they get adopted as gospel, right, which then can exclude people that you really wouldn't care to exclude, because you wouldn't mind if they were in the mix. But I don't know where we are --

MS. PACE: And these really are not guidelines. They are measures. So the guidelines are developed by the clinical specialty group.

DR. NEFF: But isn't one of the strategic goals that these become, you know, essentially, looked to as --

MS. PACE: Well, they are what we call consensus standards for measuring quality, right. Right.

CHAIR YAWN: But, I mean, do you have any reason -- I will turn it around the other way. Do you have any reason to believe that, if we only measure 40 and above, this will negatively impact the quality provided to

people 20 to 39?

DR. NEFF: If people really consider these just measures and sort of recommendations, no. I think that the risk is how things get applied.

CHAIR YAWN: Yes, but I don't think that we can, again, go out for all exceptions in the world. I mean Medicare is not covering the patients 20 to 39 already.

MS. PACE: So I guess to maybe put your question another way, what is the risk of including the broadest population?

CHAIR YAWN: I think there is a big risk of including the 20 to 39. They are quite different than people age 40 or 60 and over. I think that, if you had a large number of them, which some sites might, they could adversely affect your outcomes.

DR. NEFF: For the site?

CHAIR YAWN: For the site; that is what I am concerned about. But there may be so few that they will never be a statistical even blip on anybody's radar. It is face validity, too, though. For most physicians, they look at that at 20 and they say, what are you thinking? And nurses and everybody else. I am not picking on --

DR. MILLARD: I think if we set 40, then we are consistent with other --

CHAIR YAWN: It also becomes quite difficult sometimes to tell asthma from COPD, and somebody who is 28 and has severe

asthma --

DR. MILLARD: Okay. So, now, the denominator details, one of the questions is PR program entering completion who have completed at least 10 PR sessions within a three-month period, 90-day period. Is that CMS language? Is that what their definition is? What do they define?

DR. BAULDOFF: No.

DR. HAMM: No, it is not CMS language.

DR. MILLARD: What have they used as the minimum? Because I really think we need to be consistent, parallel with what CMS --

MS. PACE: What does the evidence show of pulmonary -- and is that how the CMS --

DR. MILLARD: I am not sure how that 10 PR within three months got there.

DR. BAULDOFF: I believe the 10 came out of the interim rule that relates to lung volume reduction surgery, that they had a specific number of sessions that were required prior to lung volume reduction surgery, to be able to indicate some kind of change. That certainly could be modified to the current CMS language.

DR. HAMM: These were written before those rules were published.

DR. BAULDOFF: Yes, before they came out.

CHAIR YAWN: I guess I have a question about this. Suppose you have a program that nobody ever completes more than three? Because your program is just impossible to get to; it is not interesting; it is not anything. I don't understand why you have to have completed the 10 sessions, because I think adherence is the biggest problem we have across all of healthcare, and this ignores the adherence issue entirely.

So my take would be, and I know this changes when they change the specification, anybody who starts and attends one session is in.

DR. MILLARD: But you may not get post-program data. If they drop out, you don't know.

CHAIR YAWN: Well, but you can say that you have that many you have no -- a lot of these people on the telephone give you that. I think you should at least try to follow up with everybody who starts.

DR. NEFF: Do we think there would be a minimum set, though, that you would need to have exposure to, to then see a benefit in these health-related quality measures? I mean, would you actually envision that one session would then change your CRQ or SGRQ?

CHAIR YAWN: Absolutely not.

DR. NEFF: Right.

CHAIR YAWN: I would not anticipate it, but it says to the program there's a

problem here because you have so many people dropping out that you don't get this improvement. So, again, it is measuring the quality of the program, not quality of completion of the program. That is different.

Which one do we want? I mean we can say we are looking for the quality of the program, if you complete the program, or from primary care, it is always if you start the program.

DR. NEFF: Yes. No, that is fair.

CHAIR YAWN: Well, I mean I don't know if it is fair. I am just asking.

DR. NEFF: No, no, no. I mean it should reflect the whole program, which would include your ability to hang onto people, follow them up. I mean the whole real deal. I mean that, I think, is what you are getting at, rather than just say we are going to look at you if you finish the whole shooting match.

MS. PACE: Because you are narrowing and narrowing and narrowing what you are measuring here.

CHAIR YAWN: Yes, because the rate of completion is not anywhere close to 100 percent.

DR. HAMM: Absolutely not. But I would just sort of respectfully suggest that, as probably a program outcome, in terms of how your adherence is, as opposed to a quality-of-life outcome that is patient-centered --

MS. WINKLER: Right. A different

measure is what you are saying?

DR. HAMM: I believe it is, yes.

MS. WINKLER: But maybe a desirable one.

DR. HAMM: Oh, absolutely. I am thinking of probably 20 that would be --

CHAIR YAWN: I am going to respectfully disagree. I think that is a patient quality-of-life outcome measure. If they start the program, they don't complete it, their quality of life is not improved one iota, or maybe it has improved greatly. I don't know. I do still think it is a patient outcome measure because, if we start ignoring adherence, when we talk about outcome measures, I am really concerned about we are going to be measuring the outcome in this tiny, little group of people.

But this is a group, and I am going to be willing to listen to everybody. Don't let me drive it. I am just asking difficult questions.

DR. MILLARD: You almost would like two different denominators, which just makes it too complex. One is people who complete the program. One is people who enter but don't complete.

In our experience, we enroll about 10 patients in pulmonary rehab every six weeks, six to eight. I mean it is a six-week program, and they let 10 come in. So we will lose, routinely, two or three of them.

CHAIR YAWN: Yes, I would expect you to lose 30 percent minimum.

DR. MILLARD: And usually, the reasons are either they get sick or life happens. So that would significantly impact the data if you didn't have a -- one of the reasons to put it down to the minimum clinical significance, as opposed to a higher one, is to allow it.

But I think these are not fixed in stone. I think if you say the denominator is people who enter the pulmonary program, you have built in drive to adherence. If in retrospective review you find that that denominator is too big a denominator, then that can always be modified.

DR. NEFF: And you can build that language probably into this measure even now, where you are actually trying to look at the overall program's effectiveness, but then, also, the specific health-related quality of life among the completers. I mean because it is a different question. I mean both are true, but you don't want to lose one for the other.

CHAIR YAWN: That it affects -- it goes back to my level of importance. If it is only measuring completers, I think it is a less important measure than if it measures all beginners, all people who initially --

DR. MILLARD: So you would take out the entire 10 sessions in three months? You

would say --

CHAIR YAWN: No, it is fine to have both. I mean that is fine. I am just saying the measure as proposed, that is what we have to keep looking at. It is the measure as proposed. This, to me, then, says there is a gap in this measure. We can talk about gaps later.

MS. PACE: Because I think what you are suggesting is, regardless of the program, if the people complete, they are probably going to have this improvement in health?

CHAIR YAWN: Well, no. It is just that, if you don't complete the program, you will have very little chance of having any benefit from the program.

MS. PACE: Right, right. But I mean, will there be any variability across programs?

CHAIR YAWN: Yes, there will.

MS. PACE: For the completers? If you only measure the completers?

CHAIR YAWN: Yes, there will, I believe.

DR. MILLARD: If you measure the completers, they will have a much higher quality of life than if you measure those who --

DR. O'CONNOR: But what I think Karen asked, if you look at people who complete in Dallas versus the people who complete in Seattle or Rochester or New York

or New Haven or something --

CHAIR YAWN: Yes, but the ones that complete in Podunk, Louisiana -- well, Podunk, Minnesota, we'll say; I'll pick on Minnesota -- may have different rates of improvement in quality of life than the ones who improved in Rochester. It may be higher.

MS. PACE: Yes, that was my question.

CHAIR YAWN: Yes, yes.

Okay. So we have said that you think this is acceptable, but the gap is that we are not doing anything about non-completers. We are not looking at adherence in this measure.

DR. MILLARD: And under strength and weakness, that would be weakness.

CHAIR YAWN: Okay, thank you.

The first one always takes a long time, and I apologize because we have to think through all of these things in context.

Okay. Do you have other strengths or weaknesses? We are going to go on with the denominator --

MS. PACE: Can I just ask, the pulmonary rehab program, does that need to be defined? Or is that going to be as defined by the CMS regs? Or is there other definitions? Or how do you know? I mean I just don't know.

DR. MILLARD: I would say it just follows the guidelines of the Joint -- what is

it now, ATS, ACCP, AARC, AACVPR? I mean everybody is onboard with the guidelines.

CHAIR YAWN: There are guidelines.

MS. PACE: So it probably should at least just reference what that definition is.

CHAIR YAWN: Uh-hum. Okay. Stratification --

DR. MILLARD: I can't see a word on the board.

Stratification details. I don't have any -- I mean stratification, risk adjustment, all these things, I think we have discussed them.

CHAIR YAWN: And you are comfortable with the no risk adjustment? I mean we have to have said, we have to have mentioned that because it is going to get a lot of pushback.

DR. MILLARD: If there is no risk adjustment?

CHAIR YAWN: Yes. And it is okay. I am not saying it is bad. I am just saying --

DR. MILLARD: I can't see how risk adjustment enters into that.

CHAIR YAWN: Oh, somebody who has cardiovascular disease or very severe arthritis who has more trouble participating and gaining some of the functional improvement might not have as big an improvement in quality of life. And those are quite common for people who are depressed.

I am not saying we have to do it. I just know that we are going to get pushback from other people. So it is okay to say -- I'm taking all comers.

MS. PACE: It is a very big issue that outcome measures, in general, should be risk-adjusted or a very good rationale for why not. So, if there is evidence that patients' achievement in this area varies by co-morbidities, severity of their COPD, then what is the justification for saying you don't risk-adjust?

So those would be the general questions that will come up as this measure continues through.

CHAIR YAWN: And is the justification that people -- we are still talking about a difference in quality of life and improvement. So people who have all of those co-morbidities start quite low and they go up .5. People who don't have all those co-morbidities start much higher, but go up .5.

DR. BAULDOFF: And the other issue is that including both the SGRQ and CRQ, these, again, are disease-specific questionnaires. So they are going to focus primarily on the pulmonary symptoms, and that is what we are using to calculate score. So they are not going to focus so much on the symptoms that we would see as part of the co-morbidities.

CHAIR YAWN: Well, that is not true

for cardiovascular disease. Dysemia is dysemia is dysemia.

DR. BAULDOFF: Right.

CHAIR YAWN: So you do have that issue, and almost all of these people have cardiovascular disease because they were long-term smokers.

DR. NEFF: It may come up more with the non-completer issue.

DR. MILLARD: Yes. Well, the other issue that really will come up with risk adjustment is when we do Richard's, which is a physical metric, as opposed to an emotional one.

CHAIR YAWN: Yes. So we are saying that we do not believe we have to risk-adjust because we are using a disease-specific outcome measure and because we are using the individual patient's change in score, which already accounts for their difference in initial scores.

DR. MILLARD: Right.

CHAIR YAWN: So that is our justification for that risk-adjusting?

(Interruption from phone recording.)

I think that Francois may not be there anymore.

MS. WINKLER: Or anyone else.

(Laughter.)

DR. NEFF: The thing that people may not fully reconcile is --

(Interruption from phone recording.)

The only thing that may kind of be seen as a conflict there is, if we are saying the reason we don't need to risk-adjust is because we are using individual changes, but then we are saying a weakness of this current measure is that we are not capturing the quitters, so to speak, you know, or the people that can't finish the program, well, that sort of rationale for the non-risk-adjusting wouldn't work for the people that come and just stop, right?

Because, then, that is going to affect the whole program's scoring, so to speak, if you have people that have that arthritis come once, can't walk, stop. If you don't risk-adjust for them, which you recognize that they had a higher likelihood of not completing --

CHAIR YAWN: Yes. Well, right now, the measure as proposed is only for the completers. So, when we talk about the gap, and talk about the gap is that you need to measure non-completers, then there may be a different comment on risk adjustment for that measure. But that is a different measure than this one.

DR. NEFF: Oh, okay. I thought we were going to sort of encourage or propose that they be together. No?

CHAIR YAWN: I don't think we can.

DR. NEFF: Okay.

CHAIR YAWN: I mean we can say they didn't do that and we think it is a gap.

DR. NEFF: Oh, okay.

DR. MILLARD: The 10 sessions is the key. That is the denominator, is the 10 sessions.

DR. NEFF: Then the delta is fine; there is no risk adjustment.

CHAIR YAWN: Yes.

DR. NEFF: Okay.

CHAIR YAWN: And that is, I think --

DR. NEFF: For now, I just have to close my brain to what we just said about the other thing. It is okay.

(Laughter.)

In a way, assessing the risk adjustment is only based on what they wrote, not what we are saying is a problem with it, and we would actually advise.

MS. WINKLER: One of the things you are doing is both evaluating this and trying to make a better measure at the same time. You are welcome to do all that, and we will capture it as, "gee, it would be nice if....", but the actual evaluation is what we've got here.

CHAIR YAWN: How many of you sit on study section? Have you ever been on a study section? Okay.

Yes, in study section we say, "This

is the proposal. You can only evaluate the proposal. You cannot rewrite it."

At the end, we will by saying the gap analysis, gee, we don't have a measure that measures the quitters, and we would like to have a quitters measure. We won't call it that, obviously.

MS. PACE: Also, in the weaknesses that you identify in this measure.

CHAIR YAWN: Yes. Yes.

DR. MILLARD: So we are at now testing analysis, reliability testing?

CHAIR YAWN: Yes.

DR. MILLARD: Again, I said "P" simply because I wanted to balance out CRQ and SGRQ.

CHAIR YAWN: Okay. Otherwise, you would have said "C"?

DR. MILLARD: Otherwise, it would be "C".

CHAIR YAWN: Okay. Is that the last --

DR. MILLARD: Validity testing.

CHAIR YAWN: There's more here.

DR. MILLARD: Content validity had been reported. Again, comments are ditto, which would be "P", just because of the SGRQ.

CHAIR YAWN: Okay.

DR. MILLARD: I want to make sure that that also -- yes, and the earlier caveat about what -- and I have done no -- if this enters into what you had to say earlier about

you want to suggest something different that is on the side or the study section about this is as it is, because this proposal has a 1.0 score. I think we all agreed that it really should be the minimum clinical difference as opposed to --

MS. WINKLER: I would ask you to expand on that because, No. 1, why was one chosen? Right, that is what I am saying; I would like to hear a little bit around that discussion.

Because it sounds like this measure, being a measure has established a certain threshold to achieve the positive credit, if you will, and you are differing with that. But I haven't had a handle exactly on why that is.

DR. BAULDOFF: I think we just reselected the moderate change. We certainly can go with the .5. There would be no argument on making that modification whatsoever. It is all out of the same article. It is all out of the Jaeschke article.

So what we did, though, is that we were looking for that whole point difference. We just went with the moderate change because we expect to see a larger change than that actually, because we are looking at completers.

CHAIR YAWN: Right, and the evidence behind recommending pulmonary rehab

was first based on it had to achieve at least the minimal significant difference for it to be recommended as valuable. But then it wasn't a huge decline in the percent of people that received moderate versus the minimally-significant difference.

DR. BAULDOFF: Right.

CHAIR YAWN: So that what you said was you chose the 1 because you didn't think it was that much different in all the evidence from the number of people that achieved .5?

DR. BAULDOFF: Right. So that is a modification that would easily be made.

MS. WINKLER: I guess we will really need to grapple with the intent of the measure. If you are basically trying to say we want to identify the really good programs because they are able to achieve a higher change score, and that is reflective of the quality of the program provided, then 1 may be your choice.

DR. MILLARD: Then why did you say 2?

MS. WINKLER: Yes. Well, I know. That is what I am trying to find out.

DR. MILLARD: That is what I am saying is, I don't think there is any data that --

MS. WINKLER: Well, that is what I am trying to get at, is there a --

DR. MILLARD: What we are really going to have a problem with, and we need to

look forward, is that when we look at physical measures, pulmonary rehab programs don't even meet, when you look at six-minute walks, they don't even meet the minimum clinical difference.

CHAIR YAWN: It seems to me right now the question is, are we trying to identify good programs or the really good programs, or are we just trying to understand the variability across the country, which we don't even know that yet? We don't know the variability of outcomes.

DR. MILLARD: The minimum clinical difference, that at least sets the goal for whatever you try to achieve. Then you have to look back and see if you have reached that goal.

DR. NEFF: Then you would have that binary component like you did yes or no, and that could get you a pool of programs. Then you would still be tracking the amount of interval change. Then you could grade the programs, if you wanted to, against each other by who had small, medium, and large changes.

MS. PACE: That is only if the measure is constructed to do like an average instead of the percent that achieved this minimal. You could have the mean change, but that is a different measure and it is not going to happen unless it is actually specified in the measure.

CHAIR YAWN: That has to be

translated into all kinds of things.

So, right now, are we willing to accept the 1.0 instead of the minimally-significant clinic difference from the data that has been presented, saying that in the studies there was not a lot of difference between the percent that achieved 0.5 and 1.0 because -- I can't remember -- it was 2-3 percent difference is all in the number of programs. So that is why you chose a higher standard. Is that putting words in your mouth?

DR. BAULDOFF: No, that sounds right on.

It was coming out of trying to figure out the highest, the whole quality. So perhaps we looked at this in the wrong way. We should have looked at this as looking at quality in a starting point rather than a higher-level --

CHAIR YAWN: Well, but even with 1.0, it seemed to be a reasonable starting point from what the research data is.

DR. BAULDOFF: But, again, I think it would be appropriate. Being the one that wrote this measure, I am almost embarrassed to admit right now, I think the .5 really is the better one to go with. But I appreciate your reviewing what was on paper when it came in.

CHAIR YAWN: So what do we do about that?

MS. WINKLER: I mean I think it is

a negotiable point, but I think, ultimately, this speaks to the question of the tools that have been used, have been tested and validated, but the measure as specified I think is where you are all having your questions about. It sounds like we don't have enough data of how that will perform when it is put in place to evaluate the quality of various programs.

So the question, then, I would ask you, we go back to, has the measure been tested or not?

DR. MILLARD: Not at 1.0.

CHAIR YAWN: Well, not across lots of different programs. We don't have data using this across lots of programs. But the measure has not been tested. So what do we do about that?

MS. PACE: So you don't have any data on this measure? So you haven't done any kind of program scores using this measure? You don't have data?

DR. BAULDOFF: We don't have that data. We are just starting a registry.

MS. PACE: Oh, then it isn't --

MS. WINKLER: Yes, well, the question I would ask, though, of Dr. Millard is, is the research data using .5 essentially a test of this, where you got results of comparing programs or that you can see how the measure performed at the .5 level? Is that what the research data shows? Because that,

potentially, tested that measure, that version of it, if you will.

DR. MILLARD: Well, then the programs measure improvements, and a minimum clinical -- I mean the goal is to see improvement.

MS. WINKLER: Right.

DR. MILLARD: And what's improvement? It's .5 or above.

DR. NEFF: So then it was .5, right?

DR. MILLARD: Well, but that is a minimum clinical difference.

DR. NEFF: Right. That is what I mean.

DR. MILLARD: Because that is how we can say, yes, we did something.

DR. NEFF: So, in a big trial, that minimally-clinically significant difference was used to then establish benefit?

CHAIR YAWN: But it was used to compare the sites. It was aggregate data. It did not compare different sites. It was a study proving that this tool --

MS. PACE: Intervention.

CHAIR YAWN: -- that this intervention -- thank you -- this intervention is beneficial, but it did not measure the quality of sites. It did not differentiate among sites. It was aggregate data and pooled data.

So I don't believe they

presented -- I mean we could probably go get it. Well, they may not give it to us, actually, because they may have promised not to give it to anybody.

DR. NEFF: Although is this measure actually saying this has to be able to compare sites or is it just saying this is what we are going to use to identify significant difference? And whether you say it in one program or multiple, you are not actually speaking to comparing it. I mean that is probably how it will be used.

MS. PACE: That is the reason for NQF endorsement, is public reporting and quality improvement. So public reporting implies that someone could look at a variety of program scores and make some conclusion about which one has the better quality. It is part of the mission of NQF-endorsed measures.

CHAIR YAWN: And it is just that this has never been used to do that.

MS. WINKLER: The comparison can be done in a couple of ways. You can have absolute numbers, percentages.

The other question, I think it is a more focused measure, maybe less robust, but to ask the question, what percentage of patients do hit the minimum? Perhaps that is useful information or not. I don't know.

DR. MILLARD: I think that is more useful.

MS. WINKLER: So it would be, yes,

you did versus, no, you didn't.

DR. MILLARD: So it is more useful than what percentage because we don't know what the difference between 0.5 and 1 is. We know what .5 is. We know that .5 is the minimum clinical --

CHAIR YAWN: Well, isn't that what this measure is, the percent who hit 1.0 or greater?

DR. MILLARD: Right, but --

CHAIR YAWN: I mean that is what this is.

DR. MILLARD: But I think it should be .5.

CHAIR YAWN: No, I hear what you are saying.

MS. PACE: But, yes, you're right, it is a percentage.

CHAIR YAWN: Right, and that goes back to the question it is an untested measure.

MS. WINKLER: Yes, I think so. I am trying to find a way around it.

CHAIR YAWN: Well, I mean I think we have all tried several times. That isn't bad because COPD rehab is so early that we can understand. It is so early in its history, unfortunately, for widespread use. We understand why it is not yet tested. That doesn't mean we can waive all the requirements.

MS. PACE: I am just kind of

quickly going through this, but do you have that information about the minimally-clinically significant difference? Is that mentioned in the information you provided in the submission form?

DR. BAULDOFF: That is the Jaeschke article.

MS. PACE: The what?

DR. BAULDOFF: I don't know --

MS. PACE: But you didn't extract that information? It is just in one of the articles? I am just asking if you --

CHAIR YAWN: Well, you need to go back to the validation for the two measures now that we are talking about, and you have to go way back to those. I don't think you quoted those as references, is what I was saying.

DR. BAULDOFF: Well, actually, the Jaeschke article from 1989 is one of the earliest on the CRQ.

CHAIR YAWN: Okay.

DR. BAULDOFF: That was reliability and validity testing. For the SGRQ, that is very simple. I have all of that --

CHAIR YAWN: Yes.

DR. BAULDOFF: -- all of that information.

CHAIR YAWN: So you did give us the reference? You just didn't say --

DR. BAULDOFF: Didn't clarify.

CHAIR YAWN: So we have it. She

just didn't cite --

DR. BAULDOFF: Right.

CHAIR YAWN: Because they used a higher standard.

DR. BAULDOFF: Right. Yes.

MS. PACE: Whenever a measure is based on some kind of benchmark, we can almost predict the question will be, what is the evidence for establishing that benchmark?

CHAIR YAWN: Yes, and we have it.

MS. PACE: Right.

CHAIR YAWN: We just could pull it out in a specific sentence. Okay.

DR. MILLARD: Exclusions justified --

CHAIR YAWN: Let's go down. Could you take us down to the exclusions, please?

DR. MILLARD: And that's neurocognitive psychiatric conditions; you can't read or write.

CHAIR YAWN: Or speak.

DR. MILLARD: Or speak, yes.

CHAIR YAWN: With the language. Okay.

DR. MILLARD: Then it says risk adjustment for outcome measures resources. It says not applicable.

MS. PACE: I think, just as we have already talked about it, it is not that it is not applicable. It is always applicable.

DR. MILLARD: Yes.

MS. PACE: It is whether there is a

justification for not doing it.

CHAIR YAWN: Well, and in our comments we can put --

MS. PACE: Exactly.

CHAIR YAWN: -- the justification we already mentioned.

MS. PACE: Right, right, right.

CHAIR YAWN: Okay.

DR. MILLARD: And identification of meaningful differences in performance. This is where I think we will have significant -- my recommendations would be, I think, what, "M", in the sense of as written, 1.0 change for moderate and 1.5 would represent a large change.

MS. PACE: And this is a little confusing in this context, but what we are really looking at here is difference in performance across programs because the measure is measuring a program. So, again, it goes back to the question we talked about earlier: are all the programs going to end up with 90 percent of their patients achieving this?

CHAIR YAWN: We don't know, but I think all of our expert guesses are, no, that there will be a fairly wide variability.

MS. PACE: Right.

CHAIR YAWN: I guess we will call ourselves expert opinions. We don't have any data.

Okay.

DR. MILLARD: And 2g, comparability of multiple data source methods, I think, unfortunately, you probably asked the wrong person. Therefore, I am saying this is "M" on that as well, but --

CHAIR YAWN: Yes, but you can really, really in your heart, justify saying CRQ is not good?

DR. MILLARD: No. I just can't say CRQ is the, quote, "most reliable validity and feasibility of use in a patients' COPD programs."

CHAIR YAWN: You could say one of the two most?

DR. MILLARD: Yes.

CHAIR YAWN: Okay.

DR. MILLARD: Yes. And then, disparities in care, that talked about, is this the disparity related to how many people -- is it a completed program or not? I think that is what we were --

MS. PACE: This is really intended to go back to the first question about variability and opportunities for improvement and whether care and outcomes vary by what are typically considered disparities, you know, by ethnicity, race, socioeconomic status, sometimes gender. So, if they have identified that gender is an issue in people getting correct care or achieving outcomes, can it be measured?

CHAIR YAWN: Well, they suggested

we could stratify by gender. So, if we did that, we would be able to tell by gender.

MS. PACE: Right.

CHAIR YAWN: The other one, you may have mentioned this under the weaknesses, but one of the big areas of disparity for pulmonary rehab is geographic location. There is no discussion here. Because if you live 50 miles from the program, you start the program -- it goes maybe more to the ones who are unable to complete, but does that lower your ability to improve your quality of life because you have to drive 100 miles every time you go?

I would put geographic disparity under that as a weakness that they didn't look at, but not a deal-killer or breaker, certainly.

DR. MILLARD: I concur.

CHAIR YAWN: Okay. So these are the strengths and weaknesses from our colleague who isn't with us today, because he is traveling internationally.

English-only, we have that. We have said that was a weakness, that we could only do English at this time.

The IPF patients, you don't have IPF patients in here, do you?

DR. BAULDOFF: No, this is specific to COPD.

DR. MILLARD: No, this is just COPD.

CHAIR YAWN: That is what I wondered because I am thinking, well, yes, okay.

Okay, and he is just saying the other weakness is he doesn't think that CRQ deals with emotional things like depression, so that you aren't measuring a full scope of quality of life.

Does the SGRQ do that, do you think?

DR. MILLARD: It does have some depression/anxiety.

CHAIR YAWN: With its 400 questions, it ought to have something.

DR. BAULDOFF: It has symptoms, activity, and impact subscores that go into the total score.

CHAIR YAWN: Okay.

MS. PACE: So I am not sure where you are at with the recommendation that this measure should include both tools. Is that something you have decided on yet or made a recommendation about? Is that a weakness of this measure as it is stated?

CHAIR YAWN: Well, I think that we have one person saying he believes it is because a number of the programs currently use that, and it would require those programs to change what they do from a perfectly acceptable measure of health-related quality of life to another measure of health-related quality of life, which that is a burden. I

was going to put it more under burden and feasibility --

MS. PACE: Okay.

CHAIR YAWN: -- than I was here.

MS. PACE: I see. Okay.

CHAIR YAWN: Does that make sense?

MS. PACE: Yes.

DR. MILLARD: And the good news is they do cross-balance.

DR. NEFF: Yes. It is not between apples and oranges. It is MacIntosh apples and Fuji apples. Sorry.

CHAIR YAWN: Yes. They will be certain to understand that one.

(Laughter.)

MS. PACE: But I think you are saying one includes depression and the other doesn't. That questions in my mind, then, are they equivalent?

DR. MILLARD: They measure different things, but they are similar.

DR. BAULDOFF: Actually, the CRQ does have a subscore of emotional function.

CHAIR YAWN: Yes.

DR. BAULDOFF: So I would respectfully disagree with the reviewer's comment on that.

CHAIR YAWN: And I agree with your comment.

(Laughter.)

I think that I just wanted to read what it said to make sure --

MS. PACE: Yes, I know.

CHAIR YAWN: Yes, I don't think one ignores depression and anxiety and the other brings it out.

Okay. Shall we go on then?

DR. MILLARD: Shall we go on to the next?

MS. PACE: So do we know what has been agreed on for the --

CHAIR YAWN: Okay. I'm sorry. What is the rating for this?

MS. FORMAN: It says "completely" for all of them.

CHAIR YAWN: And for the measure as stated, we know your concerns about the SGRQ, but we are going to try to deal with those more under feasibility.

If we didn't have the SGRQ concern --

DR. MILLARD: Complete.

CHAIR YAWN: -- would it be complete?

DR. MILLARD: Yes.

CHAIR YAWN: Okay.

MS. FORMAN: Except for the performance scale or that was partially?

MS. WINKLER: That was under importance.

MS. FORMAN: Oh, I'm sorry.

CHAIR YAWN: Yes, this is untested.

DR. MILLARD: Well, actually, it would be "P" because, also, this issue of what

level were they putting it at 1.0.

CHAIR YAWN: Yes. So we say keep.

DR. MILLARD: Yes.

MS. FORMAN: For all?

CHAIR YAWN: For the scientific.

DR. MILLARD: Yes.

MS. FORMAN: Okay.

CHAIR YAWN: Okay?

MS. PACE: We have to do each of the subcriteria.

CHAIR YAWN: Yes.

MS. PACE: But you are saying all of them put as "P"?

CHAIR YAWN: Well --

MS. FORMAN: Because 2f, you had as minimal, and 2g?

CHAIR YAWN: Could I suggest that --

MS. WINKLER: That we go back and fill it in --

CHAIR YAWN: -- we fill it in later?

Because I think that -- and I am going to suggest for these TAPs in general -- to do all of those sub-sub is going to slow the discussion down a lot. If we can do the four base categories, that would be very helpful.

MS. WINKLER: Well, we can't --

CHAIR YAWN: No, no, no. I know. I mean we go through them and we look at them --

MS. WINKLER: Right.

CHAIR YAWN: -- and we take notes, but that we don't go back and say, can we do each one? We try to do it from the knowns.

MS. WINKLER: Okay.

CHAIR YAWN: Would that be acceptable to try?

MS. WINKLER: You can try it.

CHAIR YAWN: If it doesn't work --

DR. MILLARD: So what number are we at?

CHAIR YAWN: We are at 3 now.

DR. MILLARD: At 3.0, meaningful, understandable, and useful information. And I thought it was -- extent to which intended audiences can understand the results and are likely to find them useful for decisionmaking, and with the caveats as to sort of what would be the appropriate benchmark, I think it is complete.

CHAIR YAWN: Okay. And he thought it was complete.

MS. WINKLER: Okay.

DR. MILLARD: 3b. 3c, relation to other NQF-endorsed measures.

CHAIR YAWN: That was not applicable.

MS. WINKLER: Right.

CHAIR YAWN: Because there weren't any.

DR. MILLARD: Yes. And again, there is this thing about 40 years and older;

there needs to be harmonization to 40.

CHAIR YAWN: Thank you.

DR. MILLARD: To receive distinctive or additive value, and I would say complete.

CHAIR YAWN: Well, since there are no existing measures --

DR. MILLARD: Yes.

CHAIR YAWN: -- it has to be either not applicable or complete. That would be true.

DR. MILLARD: That is easy.

On the strength and weakness, I think that's self-evident.

CHAIR YAWN: So you are pretty comfortable with the "C" for overall for this one?

DR. MILLARD: Yes. Yes. Yes.

CHAIR YAWN: For 3?

DR. MILLARD: Yes.

CHAIR YAWN: Okay.

DR. MILLARD: Now in terms of feasibility, one of the big issues is going to be in terms of data generation, and this is 4a and 4b, I think both are, how are these going to be redactable in an EHR?

DR. O'CONNOR: I think it is the overarching issue. How can you retrieve the data? I mean, if it is yes/no, is the patient in a pulmonary rehab program, that even has challenges. The new regs are going to put you ahead of that wave. So that is great because

we couldn't do this last year or two years ago. We had no way to retrieve who was even in a pulmonary rehab program.

That solution looks like it is going to get solved, but Mark and I talked about this earlier. Since you are looking at differences in scores, how can you do that? That is the hard part.

CHAIR YAWN: So this would require all of them to do it on baseline and do it at --

DR. O'CONNOR: But somebody has got to go in and extract the data. It can't be electronically retrievable.

CHAIR YAWN: Why?

DR. O'CONNOR: Because these are going to be scanned documents that you are not going to be able to say, "Tell me what the first one was? Tell me what the second one was?" to compute a difference. You can't do that --

CHAIR YAWN: Well, you are assuming that people will not have that ability in their pulmonary rehab programs.

DR. O'CONNOR: Well, we have an electronic health record, and these sorts of data get scanned. So we couldn't do this, and I suspect that, since most people have electronic health records similar to ours, they are going to have the same problems.

DR. MILLARD: Yes, you would have to pull it out separately and enter it, just

so that you could retrieve it automatically later. That would change the current clinical process, to have it be automated.

CHAIR YAWN: Right. We would have to change the EHR for pulmonary rehab programs to make these electronically-entered. I mean this is true for pretty much any of these things.

DR. O'CONNOR: It is not an issue specific to this measure.

CHAIR YAWN: Yes.

DR. O'CONNOR: This is an issue, basically, for anything that is going to require a number other than a yes/no binary situation. This is a difficult problem.

CHAIR YAWN: Well, anything that doesn't have a specific ICD-9 administrative code right now is not easy. Some lab test results we can now pull out electronically.

MS. PACE: So is this typically done in a paper/pencil format to the patient? Then who in your office would go through and score it, so that you get the score? And is that score entered?

DR. MILLARD: It is in the chart. In our chart, our pulmonary rehab chart, I see the SGRQ beginning; I see the SGRQ at the end. I see the six-minute walk at the beginning. I see the six-minute walk at the end. And it is scanned into the electronic health record.

MS. PACE: Right, but you said that someone had to go through those answers to

come up with the score?

CHAIR YAWN: Oh, yes, to get the score. Usually, the person that administers --

DR. MILLARD: But that data is not retrievable electronically.

MS. PACE: But the score --

DR. MILLARD: It exists in the charts.

MS. PACE: Okay.

CHAIR YAWN: So, if you did a chart review, electronic chart review, like you do paper, you can go down and scan through, find the SGRQ --

DR. MILLARD: Or the CRQ.

CHAIR YAWN: -- or the CRQ, either one, and say, "Okay, there it is," and there it is. But you have to go in. It is not like administrative data that I can tell you how many people have an ER visit in the last year.

MS. PACE: So about all data elements being available electronically, that is a no, not a yes?

DR. MILLARD: The answer is no.

CHAIR YAWN: Yes, right.

DR. O'CONNOR: Everything we say regarding Mark's measure applies to the six-minute walk test as well. So we are doing double-duty here.

DR. NEFF: Yes.

CHAIR YAWN: Yes. Well, and the susceptibility to inaccuracies, errors, and

things are the same kind of problems you have anytime you have to have somebody score something by hand.

DR. O'CONNOR: Now the ideal, obviously, would be to chart each one over time. Because if we take a look at laboratory test results, a CBC or a whatever, and that is electronically retrievable, but in Touchworks, which is the system we use, we don't have that functionality for things like this.

CHAIR YAWN: So you have to specifically go in and program a template for it.

DR. O'CONNOR: Exactly.

CHAIR YAWN: That is a barrier.

DR. NEFF: We would be able to electronically know that they have had it done.

DR. O'CONNOR: Yes.

DR. NEFF: So that they had it at the beginning and the end, you wouldn't --

CHAIR YAWN: Yes, but that doesn't help a lot.

DR. NEFF: -- know if it is out there.

MS. PACE: So say your electronic record had a flow sheet, you know like you use blood pressure, or whatever, can you have a spot where you would be recording the overall score?

DR. MILLARD: Unless your program --

DR. O'CONNOR: Somebody would have to have it programmed in and entered on each visit. That is not impossible.

MS. PACE: Right.

DR. O'CONNOR: But, currently, it is not the way it is done.

MS. PACE: Right, right.

CHAIR YAWN: Again, this is not going to be different for pretty much any measure --

MS. PACE: Oh, yes. It is something we encounter at every --

CHAIR YAWN: -- each set of administrative data.

MS. PACE: Right, right.

DR. O'CONNOR: Maybe since we are the first TAP, you should just take this discussion and save time in the future ones, since they know we have done this already.

(Laughter.)

CHAIR YAWN: Well, but some of them are based on administrative data only.

DR. O'CONNOR: Yes.

CHAIR YAWN: And those will be different.

MS. PACE: Right.

CHAIR YAWN: So it is one of those things that we would love to have all pulmonary rehab programs have an EHR that facilitated this.

DR. NEFF: But they don't.

CHAIR YAWN: So let's go.

DR. MILLARD: Now the other question was whether or not that data is available. Accreditation. I mean if that data is going to be available, are you going to publish the improvements in CRQ, or whatever, as to the six-minute walk on the programs?

DR. HAMM: By programs?

DR. MILLARD: Yes, by program. Is that data going to be retrievable?

DR. HAMM: To the best of my knowledge, and both Gerene and I are on the Board of Directors, I don't believe that question has been answered yet.

DR. MILLARD: So has it been asked? Has it been asked?

DR. HAMM: Yes, I think it has been part of the discussions.

CHAIR YAWN: So that we have a potential source to consider, but that would be only for accredited programs.

DR. HAMM: That is correct.

CHAIR YAWN: So that would introduce a tremendous bias and limit the scope.

DR. HAMM: By the way, I don't know if this is helpful or not. Hopefully, it is, but the company that we are using to develop this registry, electronic registry, will be having fields that will be able, data fields that will be able to report scores. I mean it is a custom database. So whatever we ask for

is what, ideally, we get. That may get at some of the questions that you have been struggling with.

CHAIR YAWN: Yes, but, again, it is not available currently. It is a wish of the future. And we don't know, even if it is there, if we will be able to have access or if the reporting is going to have access.

DR. HAMM: It is in development. I mean that much is true. We are under contract to them. It is going to happen, but --

CHAIR YAWN: Right.

MS. PACE: I want to just ask a question back on your submission about use of a public reporting initiative, and you put "NA". Does that mean not applicable or not available under 3a, using public reporting?

DR. BAULDOFF: You said public reporting initiative?

MS. PACE: Uh-huh.

DR. BAULDOFF: What I have is that the CRQ has been used as a measure of --

MS. PACE: No, right above that, you have "NA".

CHAIR YAWN: "NA". Well, if the measure has never been tested --

MS. PACE: Right, but what I am getting at is just the measures are supposed to be intended -- do you see "Use in public reporting initiative"? You have "NA" underneath that?

DR. BAULDOFF: I am not working

from the final copy because I don't have access to that. I am working from our final draft that was on the submission forms.

CHAIR YAWN: Well, it is right here. What she is saying is it is right here. It says, "Use in public reporting" -- I know you can't read it from back there -- "Use in public reporting initiative". And you put "NA". Is that because you don't think it should be, that it hasn't been? What? What does that mean?

DR. O'CONNOR: Well, this measure hasn't been used for public reporting. So it is not available.

CHAIR YAWN: It has never been used for anything.

DR. BAULDOFF: Right. So it is not available.

CHAIR YAWN: So not available? It is not that it's not applicable.

MS. PACE: Okay. All right, yes.

DR. BAULDOFF: I did not enter this directly myself. We had our Executive Director enter information into NQF.

MS. PACE: The only reason I am asking is the public reporting, the intention for public reporting is a big issue at NQF, and the Board recently affirmed that NQF-endorsed measures should be publicly reported.

So, if you were thinking that this type of measure should not have any public reporting --

DR. O'CONNOR: I could see the opportunity here for misinterpretation of the question then, yes.

CHAIR YAWN: Right.

DR. O'CONNOR: Has it been used in public reporting versus could it be used.

MS. PACE: Well, it is not just going to be "could it be?".

DR. O'CONNOR: Will it be?

MS. PACE: It is, will it be? And we are changing the question.

DR. O'CONNOR: Yes.

MS. PACE: Which is something that has been evolving. But I just wanted to get if there is some --

MR. DUDLEY: I can just tell you what we thought when we did that. Since you asked to provide the name of initiatives and locations of URLs, I viewed that as a past-looking question rather than a forward-looking question.

MS. PACE: Right, and that is changing, but the only reason I was asking if she meant not applicable or not available or --

CHAIR YAWN: So it could be used, and did you develop it thinking it should be used for public reporting, this measure?

DR. BAULDOFF: Yes.

CHAIR YAWN: Okay. So you have the answer. It just never has been.

DR. BAULDOFF: Correct.

CHAIR YAWN: So it was not available, not that it wasn't applicable.

DR. BAULDOFF: Right.

CHAIR YAWN: Thank you.

Okay. So where are we with this one now, d? Oh, I'm sorry. Feasibility. Thank you. We are at feasibility.

DR. MILLARD: Right, and we said, initially, how are data elements that are needed to compute scores generated? I think, I mean, it is "C" there, but the electronic sources is where we fall down.

CHAIR YAWN: And also, this is where we were going to talk about programs that currently use the Saint George having to change to this one; if it is specified exactly this way, it could be a problem. So we think that it would be easier for the programs already using Saint George, which is an equally-acceptable and valid measure, they should be able to continue doing what they are doing.

DR. O'CONNOR: Well, let me ask a couple of just general questions, since you were the developers of the measure.

If you take a look in a broad spectrum, what proportion of patients with COPD across the country actually have the opportunity to be enrolled in one of the pulmonary rehab programs? What percentage of COPD patients actually get into a rehab program?

DR. BAULDOFF: Fifteen to 25 percent. It is very, very small.

DR. O'CONNOR: So it is low?

DR. BAULDOFF: Yes.

DR. O'CONNOR: Okay. And are there rehab programs across the country that are certified by your organization as well?

DR. BAULDOFF: Yes.

DR. O'CONNOR: And how many of those are there compared to certified programs?

DR. BAULDOFF: We are still attempting to collect that data.

DR. O'CONNOR: Okay. So you know how many certified programs?

DR. BAULDOFF: Right.

DR. O'CONNOR: You just aren't sure how many uncertified programs there are?

DR. BAULDOFF: Right.

DR. HAMM: It is a fluid number. The denominator is so fluid. I mean it can be calculated, but it is good that --

DR. BAULDOFF: Those programs open and close according to the budget from the last month pretty much.

DR. O'CONNOR: Yes. Okay.

CHAIR YAWN: In Minnesota, over the last year, it has gone from less than 50 percent to about 65 or 70 percent certified now.

DR. O'CONNOR: Programs?

CHAIR YAWN: Yes. But when CMS

starts paying for it, I am going to bet there's going to be a whole lot of uncertified programs.

DR. O'CONNOR: That's where it is going.

So 15 to 25 percent of patients with COPD currently enroll in some sort of pulmonary rehab program, and you envision that going up with the --

DR. BAULDOFF: Yes, dramatically, with the CMS.

DR. O'CONNOR: Beginning in January because of the new regs?

DR. BAULDOFF: And I think we would also expect to see that there will be an increase in number of programs. Because now that there will be a way for funding, even though the funding -- it is free.

CHAIR YAWN: Well, that is what he is asking.

DR. O'CONNOR: Yes. Because I think that, at the Board level, they are going to want to know the commonality of this. If you've got 10 percent of patients in a program, the measure is of limited value. If it is going to be 50 to 70 percent of patients, it --

CHAIR YAWN: Well, I think the new guidelines, and this is not right now reasonable, but the new guidelines are going to move it up to a higher level of COPD patients with higher FAD lungs, as

appropriate.

DR. MILLARD: But going from 60 to 70 is actually probably not going to make much difference because the average nuance, unless we are making a diagnosis a lot earlier --

CHAIR YAWN: But we are working very hard on that part, too. So, over the next five years, this could increase remarkably, is what we ought to say perhaps.

Very good questions. Thank you.

Okay.

MS. WINKLER: Are you done with that question?

CHAIR YAWN: Are we done with this measure? Do we feel like we have a sense of how we will move it on, then, to the Steering Committee?

MS. WINKLER: Just I wanted to bring up, someone entered the room during this and hasn't been introduced. So perhaps you could just tell us who you are.

MR. DUDLEY: Hi. I am Adams Dudley, a pulmonary doc at UCSF and a measure developer for a future measure.

CHAIR YAWN: Okay. Thank you.

All right, well, we are considerably behind schedule, as you might have noticed, but that is okay. We are going to take a short -- can we just do a five-minute break instead of 15, please?

(Whereupon, the foregoing matter went off the record at 10:59 a.m. and went

back on the record at 11:08 a.m.)

CHAIR YAWN: Okay, the next measure we are going to go through is already up there.

MS. PACE: This is measure 20.

CHAIR YAWN: Right.

MS. PACE: And it is functional capacity. This is comparable, it is the same group presented --

DR. BAULDOFF: But there will be problems.

CHAIR YAWN: Okay, there will be many of the same issues. Okay.

DR. BAULDOFF: Challenges.

MS. FORMAN: And our reviewers are Dr. O'Connor and Dr. Millard.

CHAIR YAWN: Okay.

DR. O'CONNOR: Many of the comments we made with the last one apply to this one. I am going to lean heavily on Dr. Millard, who has done these things. As I indicated, I am a pediatric type of guy. So, while I understand measurement, he understands pulmonary rehab issues.

In terms of importance to measure and report, does it affect large numbers? Absolutely, there is no doubt about that. So I would agree with No. 1 as a "C".

1b, opportunity for improvement. "In summary, does data demonstrate a performance gap?" I believe all of the comments we made before apply. We all suspect

there is a performance gap which varies by geographic regions, but there's little data at the present time to be able to point to that issue.

CHAIR YAWN: But they must be really good because they quoted me.

(Laughter.)

DR. O'CONNOR: I did note that.

CHAIR YAWN: You are very smart. You checked that, didn't you?

Uh-huh. Go on. That is not a measure of performance, I have to tell you. Only an opinion paper.

DR. O'CONNOR: So I scored that a "P" on this particular issue because I think that there is, while we believe there is, there isn't data to support it quite yet.

CHAIR YAWN: So expert opinion is high, but --

DR. O'CONNOR: Yes.

CHAIR YAWN: -- expert only gets us "P"?

DR. O'CONNOR: Exactly.

1c, outcome or evidence to support the measure focus. I am going to have to defer to Dr. Millard. I did note here that in the Goldstein paper they quote outcomes, including a 38-meter increase in the six-minute walk. Yet, the measure suggests a 54-meter measurement cutoff for improvement. I wasn't quite sure why the change was made from 38 meters to 54 meters.

DR. BAULDOFF: Oh, sorry.

DR. HAMM: We are having a lot of trouble following.

CHAIR YAWN: Right here it says that the evidence, you summarize the evidence of that 38 meters increase in six-minute walk from this article.

DR. O'CONNOR: Yes, it is a randomized controlled trial.

CHAIR YAWN: Right. Do you have other articles that say 54 --

DR. BAULDOFF: Yes. That's Redelmeier.

CHAIR YAWN: -- is a better measure?

DR. BAULDOFF: Redelmeier is the one that is most consistently cited.

CHAIR YAWN: Okay.

DR. MILLARD: Well, but that is minimum clinical significance.

DR. BAULDOFF: Okay.

DR. MILLARD: That is not outcomes in pulmonary rehab programs. The discussion comes back to haunt us because, previously, in the HQL, whatever, HRQOL, we used the minimum level of significance as the benchmark. Unfortunately, when you look at mean data on pulmonary rehab programs, improvements in the six-minute walk, which points out the problem of the six-minute walk, it is almost all of the programs reported improvements below the minimum clinical significance.

DR. O'CONNOR: Which is 38 meters?

DR. MILLARD: Which is 54 meters.

DR. O'CONNOR: Fifty-four meters?

CHAIR YAWN: But what article does that come from?

DR. BAULDOFF: It comes from Redelmeier. The 38 comes from Goldstein.

DR. O'CONNOR: Redelmeier isn't quoted here.

CHAIR YAWN: Okay. I don't see Redelmeier.

DR. O'CONNOR: One of the issues I struggled with here in this 1c is that I couldn't find evidence for why the 54-meter distance was being recommended, when the only quote was the 38-meter difference as a significant improvement outcome.

DR. BAULDOFF: Right.

DR. O'CONNOR: So I gave that a "P" myself.

All right. The next one, we go to 2a, precisely specified the number of patients --

CHAIR YAWN: Wait a minute. So, overall, are there any other overall weaknesses or strengths that you wanted to comment on 1? Do you think you have covered them all in your comments so far?

DR. O'CONNOR: Yes.

CHAIR YAWN: Dr. Millard, do you have any other?

DR. MILLARD: Well, I mean, the

weakness is that the benchmark, where it is 54, the current published data on pulmonary rehab programs do not meet minimum level of clinical significance, and it is not likely that they ever will.

CHAIR YAWN: So you believe we would assume that only 30 percent or less, perhaps, of people will ever be able to reach this benchmark?

DR. MILLARD: What I don't know is the percent.

CHAIR YAWN: Yes, I don't know, either, but I am guessing. I was just making a guess.

So you believe, if you set this high, that it will always be less than 50 percent? We can say that at least probably? We don't know what the medium is, but we will assume the medium and median are similar. They may not be.

DR. O'CONNOR: The other question is, if you look at our frequency distribution curve for pulmonary rehab units, do you have some rehab centers that are consistently showing superior results in this compared to those who don't? It sounds like none of them met the 54-meter requirement. Is that what you said, Mark?

DR. BAULDOFF: I would say one of the best ones is at UCSD, and if they are not meeting it, nobody is meeting it.

CHAIR YAWN: Well, they aren't

meeting it as an average.

DR. O'CONNOR: Yes.

CHAIR YAWN: And we don't know what the percent of patients that meet --

DR. MILLARD: The frequency distribution. I mean there is probably a better metric that is more sensitive to improvements in pulmonary rehab, as it turns out, which is a constant low endurance time.

CHAIR YAWN: So, again, we have an untested measure, but the literature suggests that the studied rehab programs cannot meet this as a mean improvement.

MS. PACE: So how did 54 get established as the clinically-significant improvement?

CHAIR YAWN: Well, that is totally separate from how you would decide how much you can do.

DR. MILLARD: The six-minute walk distance is used in a lot of other disease entities besides --

MS. PACE: Right, right. So you are saying that 54 is not specific to COPD?

CHAIR YAWN: Oh, I think it is. It is --

MS. PACE: Okay. All right.

CHAIR YAWN: -- specific to COPD. But what they said was, if you can't improve by 54 meters, then it doesn't improve your life outside of this.

MS. PACE: Okay. All right.

CHAIR YAWN: So they looked at things like your ability to shop, your ability to --

MS. PACE: Right. Okay.

CHAIR YAWN: -- do activities of daily living, and things like that. If you couldn't do a 54 improvement, you didn't improve those other things.

MS. PACE: Okay.

CHAIR YAWN: That is how they validated the minimal clinically-significant difference.

MS. PACE: Right. Good. Okay.

DR. O'CONNOR: Okay. Move on to 2?

MS. PACE: Well, I think that is a huge question of whether it should even move on to No. 2. Because if you are saying that is a benchmark that is not relevant, then the question is, should we even care about how it is measured? Because what's the purpose?

CHAIR YAWN: Well, because we don't have the frequency distribution, we really don't even have a good idea if 10 percent of the patients meet it, 30 percent, 20 percent. We know it is less than 50 for sure probably.

So what do you think? Is this a measure that we should go forward and do all of the rest of the assessment on?

MS. PACE: Because the importance deals with high-impact area, which I think you have all agreed. Is there an opportunity for improvement? Which, basically, you are saying

we have no idea.

Then the third area is the evidence. I am hearing that there is a lot of evidence that this isn't even attainable.

DR. O'CONNOR: That is the weakness. I mean she summarized it.

MS. PACE: I'm sorry. Okay. I'm sorry.

All we are asking the TAP to do is point out these things. It will, ultimately, be the Steering Committee --

CHAIR YAWN: Yes.

MS. PACE: -- that makes the decision.

CHAIR YAWN: So we really have to go through them --

MS. PACE: Yes.

CHAIR YAWN: -- even though we think --

MS. PACE: Yes.

DR. O'CONNOR: As the Chair would say, as it is written.

MS. PACE: Right.

CHAIR YAWN: So the importance should go to what then? What level of importance do you want to say, "C", "P", "M", or "N", based on what --

MS. PACE: Well, it is not an overall. That is for the Steering Committee. So what we are asking the TAP to do is to tell us whether it is high impact, which I think you are saying is "C"; opportunity for

improvement, I think you are saying "P" or "M" because there is no information.

CHAIR YAWN: Actually, I am going to go back to the fact, for the overall impact, because we aren't measuring the people who dropped out, I am going to a "P". Would that be acceptable to the rest of you, to say importance is "P"?

MS. PACE: For the high impact?

CHAIR YAWN: For the high impact, it is "P".

MS. PACE: High impact, "P".

CHAIR YAWN: Okay. Then, I'm sorry, go ahead. Go down, please, to 1b, so we can get that.

Opportunity for improvement, what are we saying it is? We don't know.

DR. O'CONNOR: I have recorded that as a "P" because we don't have any data on the performance gap.

CHAIR YAWN: But are you willing to go "P" or do you want to go lower because of the concern that, so far, the programs can't meet that, we don't think?

DR. O'CONNOR: Well, it is just that piece of data.

CHAIR YAWN: Well, the 54 and the 38, we do know there's something.

MS. PACE: Let me explain, too, what the scores mean. "N" means not at all, not addressed, incorrectly addressed, or not demonstrated to meet the criterion.

So, if you think, based on your judgment, that -- so having no information would actually be an "N". But if you think that, from your judgment, that there is some evidence or --

CHAIR YAWN: Well, there is some evidence, but the evidence suggests that it didn't meet that criteria. So that would still push it back to "N", wouldn't it?

DR. NEFF: It is like the concept of studying the six-minute walk is more solid than the goal measurement. If it hadn't really kind of had that high of a reach, we would probably be saying this is all very doable.

CHAIR YAWN: So 34, 30, or something.

DR. NEFF: Right.

CHAIR YAWN: But then we say it is not clinically-significantly different. Then we have that problem.

DR. NEFF: Correct.

CHAIR YAWN: So, either way, we have a big problem.

MS. PACE: So is there any evidence, either demonstrated -- I mean we don't have it here. Opportunity for improvement, the performance gap, we don't know, is basically what I am hearing you saying.

And the evidence for the 54, under 1c, is also --

DR. O'CONNOR: I am not sure where the number comes from, although the ATS article that Mark has -- read that sentence, Mark.

DR. MILLARD: Well, "The clinical relevance of the benefit of pulmonary rehab is illustrated by the improved functional capacity, as measured by the six-minute walk test. The pooled effect size of all randomized controlled studies in the results of pulmonary rehab is 49 meters with a 95 percent confidence of 26 to 72. The minimum clinical importance difference in the six-minute walk test has been estimated to be 54 meters."

And that reference is, to answer the question of where did that come from --

DR. BAULDOFF: Redelmeier 1997.

DR. MILLARD: Yes, Redelmeier 1997. You got it.

CHAIR YAWN: So where do we want to go with this?

MS. PACE: Well, the question is, has it met the criteria, and to what level? Completely? Partially? Minimally? Or not at all?

CHAIR YAWN: Okay. We are going to make you say something.

DR. O'CONNOR: Well, the 54-meter -- and you're looking at me?

CHAIR YAWN: Yes.

(Laughter.)

DR. O'CONNOR: Yes. Oh, I'm sorry.

The 54-meter is an estimate of minimally clinically-significant difference. I am not sure what data that is based upon because it is an estimate, where in the actual study the average was 49 meters.

CHAIR YAWN: Yes, but that is totally different.

DR. O'CONNOR: I know. So what percentage of patients actually achieved 54 meters or greater? Anybody have any clue? Because that is the important question.

DR. NEFF: Whether it is achievable or not.

DR. O'CONNOR: Yes.

DR. NEFF: And is there data to support it?

DR. O'CONNOR: Exactly.

CHAIR YAWN: So, right now, we have no data to support it.

DR. O'CONNOR: I have no data.

CHAIR YAWN: So doesn't it have to be an "N"?

DR. O'CONNOR: I think, given the definition that she described it as, yes.

CHAIR YAWN: Okay. Anybody want to make it something else?

I mean please realize that this doesn't mean that they can't go back and change the measure and submit it again. We are not telling them go away forever. We may just be saying, right now, this one doesn't

work the way it is.

DR. O'CONNOR: Because there is disconnect; 37 meters, 49 meters, 54 meters.

CHAIR YAWN: Yes. So we don't know what's what.

DR. O'CONNOR: Exactly.

DR. MILLARD: You know, I think just the way it was written, again, the earlier one really should have been written as quality of life should be measured, health-related quality of life should be measured only if rehab programs has an outcome, rather than say this is the benchmark for distinguishing between success. Likewise, you should say there should be physical assessment measurements in pulmonary rehab. The six-minute walk, constant low endurance, but that --

CHAIR YAWN: Yes, but that is not really an outcome. That is a process measure.

MS. PACE: That would be a process measure.

DR. MILLARD: I understand. I understand.

CHAIR YAWN: They would require it to be outcome. Okay.

DR. O'CONNOR: If you look at, in measurement -- I mean diabetes is the best example that Mark and I were discussing earlier. Fifteen years ago, we were happy to measure what proportion of patients actually had an A1C measured.

CHAIR YAWN: Right.

DR. O'CONNOR: Then we went to what proportion are well-controlled, adequately controlled. And we are now at superbly controlled. And now we have bundled them, and there are five measures, and you're either excellent or not.

So we have seen this shift. So, starting with COPD as a process measure doesn't bother me very much because we have to start somewhere.

CHAIR YAWN: Well, but that is not what our --

DR. O'CONNOR: I know.

CHAIR YAWN: I think that is part of the context that we will give back to the Steering Committee, is that it is very early in its expansion to being a major intervention.

DR. O'CONNOR: And that being ahead of the curve is a good thing.

CHAIR YAWN: Yes. It gives some time to come back when we are ready.

DR. NEFF: So what is the form for the valuable process measures? I mean we are not getting too far off-topic. I mean I understand the need to be constrained, but, also, you hate to think you are being constrained just by external constraints, and there are good ideas out there that are just getting --

CHAIR YAWN: Well, there are other

times that process measures are called for. Actually, if you go through all of the process measures of NQF, they do have process measures, lots of them.

DR. NEFF: Yes, but I just mean sort of contemporaneously, you know.

DR. MILLARD: And a suggestion, just a suggestion for a better outcome measurement, at least if you read the literature on pulmonary rehab, would be a constant low endurance time. Because that has been shown to be much more sensitive to changes in pulmonary rehab than the six-minute walk.

CHAIR YAWN: Well, and is it, then, directly related to the patient's life?

DR. MILLARD: It has, well --

CHAIR YAWN: Yes. See, that is where we get --

DR. BAULDOFF: I come at it from Barbara's standpoint. To do that is very different than --

DR. MILLARD: Yes, yes, the six-minute walk, yes.

CHAIR YAWN: Yes. Okay. Let's go ahead, please.

DR. O'CONNOR: Okay. We were at 2a then?

CHAIR YAWN: Yes, we are.

DR. O'CONNOR: Is it precisely specified? Sure, COPD, NPR, who have achieved at least 54 meters, it is very precisely --

CHAIR YAWN: Okay.

DR. O'CONNOR: So I would give that a "C".

CHAIR YAWN: So that one is a "C".

DR. O'CONNOR: Yes. Let's see.

CHAIR YAWN: 2b.

DR. O'CONNOR: Reliability testing.

CHAIR YAWN: It doesn't mean we like all of these.

DR. O'CONNOR: Yes, I know. Is it reproducible?

CHAIR YAWN: It just means that it is.

DR. O'CONNOR: Yes. According to the data, you know, looking at what they provided, the developers have provided, it would seem to be a reliable measure that is --

CHAIR YAWN: Yes.

DR. O'CONNOR: It had a correlation of .88.

CHAIR YAWN: That is pretty high ICC.

DR. O'CONNOR: Exactly, yes.

CHAIR YAWN: So, okay. So that one gets a "C".

DR. O'CONNOR: And 2c, validity testing. I've got "C" here. It was based on a study of 60 patients. That is an incredibly small number of patients to hang your hat on. So I will defer to my pulmonary colleagues on this one. I don't know if this really --

CHAIR YAWN: Well, then, they were

also patients with end-stage lung disease.

DR. MILLARD: What are you testing the validity of? The use of 54 --

CHAIR YAWN: I don't know. Tell us what --

DR. NEFF: It is just a physiologic, right? A six-minute walk to -- maximum of two -- that's it.

MS. PACE: So this doesn't do what you were talking about earlier, about connecting that to quality of life or function then?

CHAIR YAWN: No. So "P" or "M"? I don't think it can be a "C" with 60 end-stage lung patients only, and not talking about what it meant for the rest of their life, and the fact that we used pulmonary rehab in patients with other than end-stage lung disease.

So what would you like it to be?

DR. O'CONNOR: I would give it a "P".

CHAIR YAWN: All right. It sounds great.

DR. O'CONNOR: And exclusions justified, I think everything we talked about before applies here. So I gave that a "C".

CHAIR YAWN: Yes, the only exclusion that they do have and they don't say explicitly is the people who don't complete all get thrown out, and we have talked about that. So, okay.

DR. O'CONNOR: 2e, risk adjustment,

that is not applicable to this discussion.

2f, identification of meaning differences --

CHAIR YAWN: And it is not applicable, but -- or it is applicable, but the same things we said before apply this time, too.

DR. O'CONNOR: Yes.

CHAIR YAWN: Okay.

DR. O'CONNOR: Yes.

CHAIR YAWN: As "see below"?

DR. O'CONNOR: Yes, a better way to put it.

2f, identification of meaningful difference in performances. A hundred and twelve patients with stable, severe COPD, half of whom would increase the patient's perception of clinically-minimum increases the data, determined to be 54 meters.

Again, it is 112 patients. So it is not a huge number of patients. And I don't get a sense of what percent of patients actually accomplished the 54 meters. So I gave that a "P".

MS. PACE: Right. And again, this doesn't address the -- overall, these things are still at the instrument level versus the overall measure --

CHAIR YAWN: Yes, the overall measurement. Okay. So "P".

DR. O'CONNOR: And g, comparability of multiple data sources and methods.

MS. PACE: It is probably not applicable.

DR. O'CONNOR: Yes. I would agree with that.

Let me see here.

MS. PACE: What this is getting at is, which doesn't apply to this so much, but say you have a measure where you say you've got these specifications if you do a chart abstraction, and you have these other specifications of you take it from claims data. And the question is, will you get comparable scores?

CHAIR YAWN: Are they comparable?

MS. PACE: But it is not really applicable.

CHAIR YAWN: Right, because you would write down how far they walked, no matter where you get it from. So we can say not applicable and you'll accept that?

MS. PACE: I think so.

CHAIR YAWN: Okay. Disparities in care.

DR. O'CONNOR: The one disparity they talk about is that fewer than half of the COPD patients have been diagnosed, but that is not really relevant because you are only measuring the program, people in a PR program anyway. So I am not sure disparity here is applicable. I gave it an "NA".

MS. WINKLER: Is there any question about equal access to pulmonary rehab programs

that might be along these kinds of lines?

CHAIR YAWN: But if this is a measure of people who had access already, then that doesn't matter. So, if you already have access, you know, are women and men and high- and low-income people or different ethnicities likely to have different outcomes because of those issues?

DR. MILLARD: Well, the disparity of care, ironically, at least in our area, our local CMS, had always approved pulmonary rehab. So we have always had it. But, ironically, private insurance would not.

CHAIR YAWN: Yes. So there was a disparity, but it was backwards.

(Laughter.)

And it was before this measure would be applicable that the disparities should be seen, which is a weakness of this measure. We have already talked about that.

Okay, 3.

DR. O'CONNOR: Three, usability. Is it meaningful, understandable, and useful? The six-minute walk is a six-minute walk. So, from that perspective, yes, it is very understandable? Is it useful information? I think it is currently employed by all pulmonary rehab programs. So I gave it a "C".

CHAIR YAWN: Do you think it is understandable to the average public?

DR. NEFF: You walk for six minutes and you count how far you go. So, I mean, it

is actually one of the few things that people can understand.

CHAIR YAWN: Yes. No, I just wanted to make sure because you had only commented on health professionals. I just wanted --

DR. NEFF: We'll see you in six minutes.

CHAIR YAWN: You can walk farther if you go to this program, and that must be good.

DR. O'CONNOR: 3b, harmonization. I don't think that is an issue. It doesn't apply here. I scored that as an "NA". I don't know whether endorsed recommendations currently create harmonization issues.

3c, distinctive or additive value.

CHAIR YAWN: Since there aren't any measures otherwise?

DR. O'CONNOR: Exactly. I gave that a not applicable.

CHAIR YAWN: Sounds good.

DR. O'CONNOR: And feasibility, everything that we talked about before applies here because of the electronic health retrieval challenges that are faced. It is the same issue that we talked about in --

CHAIR YAWN: So what did we give it? We didn't give the other one actually anything. Do we give them "P"? What do we do?

MS. PACE: On the feasibility?

CHAIR YAWN: Yes.

DR. O'CONNOR: Yes, because they admit here, coding and abstraction are performed by someone other than the person obtaining the original information. I mean, if one of the goals is to move to using measures that only can be retrieval electronically, then it fails that test.

CHAIR YAWN: Well, but if they have their registries and potential -- I mean they have suggested solutions. So that was why I was --

DR. O'CONNOR: Yes. I think that you could probably score it as an "M" with an explanation.

CHAIR YAWN: I have no problem with that.

DR. O'CONNOR: As the Chair said earlier, as the measure is written currently.

CHAIR YAWN: Right. Right. Okay.

So is there any more? Usability? Don't we have usability?

MS. WINKLER: We already did it.

CHAIR YAWN: Oh, we did it? Oops, I'm sorry.

MS. WINKLER: Those are the three.

CHAIR YAWN: Oh, those are the three. I will get this down in a minute. One is importance. Two is science. Three is use, and four is feasibility. I'm getting there.

Okay. Thank you. That was very thorough, I think.

DR. O'CONNOR: Dr. Millard was the canary we sent into the mines, though.

DR. MILLARD: The only question I have about feasibility is, when it says, "extent to which required data are readily available, retrievable, without undue burden," and then, "can be implemented for performance measurement". Now we have already decided that performance measurement, if this is a performance measurement, the majority of programs, what we find, don't work. I mean I don't hit the performance measurement.

CHAIR YAWN: Well, but isn't that what we said -- well, you could put again here, too.

DR. MILLARD: I mean that is part of feasibility.

CHAIR YAWN: Uh-hum. So that would take feasibility down to -- what would you like it to take it to, "M"?

DR. MILLARD: Well, if that is --

CHAIR YAWN: Or do you want an "N"?

DR. MILLARD: If it is times zero, it will be "N".

CHAIR YAWN: Okay, let's do "N" then. I have no problem with that specific statement: this is why it moved from "M" to "N".

MS. PACE: But you are saying, your statement is because the majority of programs cannot meet that 54-meter benchmark?

DR. MILLARD: Because the mean

programs don't --

MS. PACE: Right. Okay.

DR. MILLARD: -- don't meet that.

CHAIR YAWN: Well, and these are not the mean programs. These are the programs who are in randomized controlled trials, so tightly controlled nobody can do what they can do. So this is efficacy moves to effectiveness. We know it brought it up by 50 percent.

DR. MILLARD: Eight out of 14 programs reported in the literature did not meet.

CHAIR YAWN: Yes. So, okay. All right.

MS. FORMAN: Okay. The next one is 023, intensive care length of stay.

CHAIR YAWN: Okay. We've got a change in groups. So you didn't get to give us any information before. Were you where they gave their explanation?

Okay. So what we want is a short explanation of where did this measure come from and why do you have it, and on what basis?

Is that close enough to -- yes, please.

MR. DUDLEY: Okay.

CHAIR YAWN: Did you introduce yourself to the people over here?

MR. DUDLEY: I did, but I will do it again.

CHAIR YAWN: Please.

MR. DUDLEY: I am Adams Dudley from UCSF. I am a pulmonary doc out there, but I spend most of my time developing measures of performance and --

CHAIR YAWN: No wonder we don't have any pulmonologists anymore. you guys are all doing -- would you go back to doing pulmonology, please?

Go right ahead. Sorry.

MR. DUDLEY: No problem.

I founded and run the chart program which produces calhospitalcompare.org. As part of what we do there, we have 246 hospitals in California, and we measure ICU performance.

However, ICU performance measures are not new when CHART started. In fact, we have needed risk-adjusted ICU mortality measures and other ICU outcome measures for decades, and they have been around for decades. So the first versions of the models that I am proposing to you today actually came out of the 1980s.

CHAIR YAWN: Thank you. All right.

MR. DUDLEY: So, when we started measuring performance in California, our program was voluntary. We had to get the hospitals together to get them to agree with other stakeholders of what should be measured and that it could be measured adequately.

I, at that time and still now, was

on the Joint Commission's Intensive Care Unit Performance Assessment Committee. The Joint Committee was getting that measure set ready, and we adopted something. We adopted it, basically, what it was at that point.

We, since then, have added the risk-adjusted length-of-stay measure. We today are presenting from that group of measures only the outcomes measures. So it is the risk-adjusted mortality and risk-adjusted length of stay.

The model that we have been using in California is the mortality prediction model. We have now gone up to the third version. That was based on work that we did where we first compared all extant models of APACHE/SIMS and found that, for the purposes of assessing hospital performance, it didn't terribly much matter which model you used; you got the same ratings and rankings for hospitals regardless of model.

But we also addressed the issue of how much time and effort it took to obtain the data, and the model that we used, the MPM model, required about less than a third of the time required to collect the APACHE data and about half the time to collect SIMS.

So, now, it is up. It is publicly reported. We have 246 hospitals that volunteered to do this with us. Not all of them, but almost 200, a few over 200 have ICUs, and they are doing this on 400 patients

per year.

And I'll stop there.

CHAIR YAWN: And they are all in California still? Thank you.

MR. DUDLEY: They are all in California.

CHAIR YAWN: Yes. Now there's nothing wrong with California.

(Laughter.)

MR. DUDLEY: Okay.

DR. O'CONNOR: You, apparently, haven't talked to our budget people, have you?

MR. DUDLEY: Yes. We need you in the legislature then.

DR. O'CONNOR: Yes.

CHAIR YAWN: There is nothing wrong with California assessing its quality of care. How's that?

(Laughter.)

That's what I really meant. I am not going anywhere near their politics, and I'm sure you don't want to, either.

All right. Very good.

I am going to assume, a wild guess, that you might have been a reviewer on this one.

DR. NEFF: I was.

CHAIR YAWN: Even a primary, huh?

DR. NEFF: I was. How about that?

(Laughter.)

CHAIR YAWN: And the secondary?

MS. FORMAN: Is Dr. Lewis, and his

is on the screen.

CHAIR YAWN: Okay. Great. Thank you.

DR. NEFF: And then we are flipped for the other one. He was doing the first, though.

CHAIR YAWN: All right.

DR. NEFF: I think some of the actually input that you have may answer a few of the question as we are going through. So this will be good.

I think, just summarizing, and correct me if I've got any of this wrong, but I think the nice thing about this is it is sort of using ICU length of stay in a way as kind of a surrogate as well for ICU resource use, quality of care, efficiency of care. I mean that is really kind of what we are getting at the heart of it.

This is presenting a modification of a model, as you have described, that is already there and is modified for contemporary kind of mortality assessments, as relevant to really the ICU environment, not just hospital mortality.

Then using the other sort of attachment as the ICU outcomes data collection instrument, which they provided as well, which is what would be the data collection piece. And I don't know if that was already out there for the prior or if this was developed just for this piece, to the round 3 modified,

MPM-III.

MR. DUDLEY: No, that is an evolution.

DR. NEFF: Okay.

MR. DUDLEY: So, when we built it the first time, it was with MPM-II, which we got from the original.

DR. NEFF: Yes.

MR. DUDLEY: Then this is a modification of that.

DR. NEFF: Including the data collection?

MR. DUDLEY: Yes. I don't know if you guys ever consider or talk about two at the same time, but this and the next measure come from that same forum.

DR. NEFF: Yes.

CHAIR YAWN: Yes. Well, you can see, we sort of, the last measures sort of overlapped, and the others that we are going to talk about overlap greatly.

MR. DUDLEY: Okay.

CHAIR YAWN: So sometimes we will spend a lot of time on the first one, and the second one we say, "as above".

Okay. So this one is the length of stay?

DR. NEFF: This is the length of stay, and the reason we are hearing kind of the mortality talk is more because that is the prediction model that was used kind of really to help them risk-stratify for the ICU. So

that this is all about the ICU length of stay, and the mortality is to come.

CHAIR YAWN: Okay. And my first question is, is length of stay an outcome?

DR. NEFF: It is.

CHAIR YAWN: Okay. No, I'm just asking.

DR. NEFF: Yes, yes. No, no, no. Well, no, I actually had to think about first, I think. Yes. Yes, it is not one that is maybe as readily understandable to you, unlike a six-minute walk, which people can understand, but if you do kind of describe it to how long you are in the hospital.

MS. PACE: It is kind of a proxy for, like you say, complications, management, et cetera.

CHAIR YAWN: But I think we have to be able to justify that to the Steering Committee.

MS. PACE: Right, exactly.

CHAIR YAWN: Because I can see several of them saying, wait a minute, that's a process measure.

MS. PACE: Right.

CHAIR YAWN: So we are saying it is a proxy measure for how well the patients do.

MS. PACE: Yes, and I think, I mean at least as I was kind of running through the list and thinking about it kind of in a more rigorous way, it is about kind of resource use. So it is kind of a cost-related issue,

but then it is about quality of care, efficiency of care, I mean all those things that then combine in one some sort of outcome, which would be your length of stay. So all those would come together.

DR. O'CONNOR: Your viewpoint is that it is an outcome measure?

DR. NEFF: Yes.

MR. DUDLEY: Yes.

CHAIR YAWN: And the other is that it is better not to be in the ICU than to be in the ICU from the patient perspective. No, I mean I think that is perfectly justifiable.

MS. WINKLER: Yes, yes, unless you need to be in the ICU. Then it is a good thing.

CHAIR YAWN: Well, yes, I know, but you want to be well enough to not be in the ICU.

DR. NEFF: And you want to get out. When you are there, you want to get out.

CHAIR YAWN: Yes, you do.

DR. NEFF: Out and alive.

CHAIR YAWN: Thank you.

DR. NEFF: Yes.

CHAIR YAWN: I just wanted to make sure we could justify that to the Steering Committee.

DR. NEFF: I think, as we kind of start ticking through the issues in terms of whether, for 1a, whether it is important and is really describing sort of a high-impact

area, I think they did a nice job of actually sort of accounting for certainly the cost perspective. I mean I think, again, it is a surrogate in some ways of quality and effectiveness and all these other things, but really not only number of people are affected, but high percentage of total hospital cost. However you slice this, it ends up being high-impact, whether it is number of people, cost, social toll, and, you know, all those sorts of things.

So I think, for me, got a "C", which is a good thing. I kind of like "A's", but we don't have "A's" on this. "C" is good.

(Laughter.)

This whole thing, my whole brain is going to have to get reworked. Okay.

And then b, sort of the opportunity for improvement. There has been established, certainly, variation in ICU length of stay certainly within regions across the country. So the ability to sort of document that and then allow some public reporting and tracking, so that you could actually benchmark yourself against other equivalent hospitals, whether it is all academic centers, community centers, I think there would be a lot of ability nationally to be able to do that. So I thought that was on target.

CHAIR YAWN: And there is national data, not just California-based data?

DR. NEFF: There is national data.

CHAIR YAWN: Thank you.

DR. NEFF: Yes. About variation and ICU length of stay?

CHAIR YAWN: Yes.

DR. NEFF: Yes, yes. It is just better described in California.

CHAIR YAWN: Oh.

DR. NEFF: Yes. No, absolutely.

CHAIR YAWN: But we didn't want to have to compare all of those against California all the time.

DR. NEFF: No. No, we would like to be able to do a whole range.

CHAIR YAWN: Yes, that's fine.

DR. NEFF: And then, let's see, the part c, outcome or evidence to support measure focus. I think here sort of the ability to say that this outcome, we are kind of getting back I think at the same question, whether it is an outcome or not. I think the ability to invoke efficiency and quality, and then compare it between sites, I think made it a relevant measure and focus.

I am just seeing I actually put it as a "P". I am just trying to remember why I did. Sorry.

I think, actually, because at that point I was sort of thinking about it more as a summation, as opposed to a single sort of getting at this issue as an outcome or a sum of things that ends up being the outcome. So it kind of slid a little bit off of a "C" to

me, but it really is getting at the same measure that we were talking about.

CHAIR YAWN: So you are pretty comfortable with it as a "C"?

DR. NEFF: Absolutely.

CHAIR YAWN: Okay. All right.

DR. NEFF: Then I don't know how we compile here with --

CHAIR YAWN: He had a "C" also.

MS. FORMAN: He had a "C" for everything.

CHAIR YAWN: Yes. He is a "C" guy.

(Laughter.)

DR. NEFF: Yes, I know. It is kind of like how you score, how you do evaluations.

CHAIR YAWN: Although we do want to make sure we talk about the strengths and weaknesses a little bit because those are important.

DR. NEFF: Yes.

CHAIR YAWN: So, yes, he's got the same comments that you made about there is good that it varies and, yes, this is important.

The weaknesses, there are confounders. Okay.

Yes, I had this question, the e, the step-down beds. I work in a smaller hospital. We don't have step-down beds. Our patients have to stay in the ICU until they are ready to go out to the floor. How does that affect this measure?

DR. NEFF: I kind of had this same issue that sort of threads its way through all of this. How are we describing an ICU? How your hospital is set up, that infrastructure, impacts a great deal, whether it is step-down bed availability, boarding because you are too full to get people out of ICU. You can't get people in the ED yet.

It is valid as long as everybody has similar issues or in tracking over time, but is that a chunk of this that there is a way to assess or is this measure hurt by the lack of that, I think is a concern.

CHAIR YAWN: So, within your hospital, tracking over time probably is okay because it may not change too much.

DR. NEFF: Right.

CHAIR YAWN: But comparing my hospital to yours, and I don't know if risk adjustment is going to deal with that issue or not.

DR. NEFF: I think it is always a hard one to actually quantify, even within a hospital, which is probably why it hasn't kind of fit in there terribly well in terms of measures.

You could describe hospitals and somehow try to build it into a model about whether you had multiple levels of care, as opposed to just two levels of care. I don't know if this is something you guys have struggled with.

MR. DUDLEY: So we have a community sample of hospitals. So we have every type of hospital.

DR. NEFF: A range of them, uh-hum.

MR. DUDLEY: A lot of people think California is all urban, but, actually --

CHAIR YAWN: No, we know it's not.

MR. DUDLEY: There are populations in rural areas.

There are issues of how you draw the lines around the walls of the ICU. But, in general, the view of the participating hospitals has been that, if -- and this includes even the small ones, however -- if there is a good reason for the patient to be somewhere else, then it is actually not that hard to create some step-down-ness.

If there are issues of the ER blocking things and making ICU patients ending up sort of being admitted to the ICU, but actually physically in the ER, something like that, then those are flow issues that ought to be worked out in the hospital. So they haven't objected to this being a performance measure for either the entering direction or the leaving direction, because they feel like if there is a problem there, it is a legit thing that they ought to fix, both because the patients -- so if you could create step-down-ness, then you are taking the less sick patients away from the more sick patients, and you are reducing the risk of passage of

infections, et cetera, et cetera.

And ER blockage of the problem, not just for calculating your ICU length of stay, but, actually, for the care of the patient. So they have, in general, accepted this.

DR. MILLARD: So the underlying assumption is ICU care is bad. The longer you are in the ICU, the worse you are. And the underlying assumption, so if I have tally pulsometry beds that I can move a bunch of patients in the ICU out to, and that will lower my length of stay right away, that is good. The assumption is that that is good?

MR. DUDLEY: Clinically, it is better for the patients.

DR. MILLARD: Okay. Okay. But that is the underlying assumption, is, however you get them out of the ICU --

MR. DUDLEY: On the back end, and, also, the front end, however you get them in. So this business of calling them ICU patient, but having them sit in the ER for 24 hours is not as good of care.

CHAIR YAWN: But do you have evidence to say that it is better to get them out when you don't -- I mean we really don't have the ability to hire more staff to have a step-down unit. It just is an economic.

So can you show, do you have evidence that it is better to get them out of the ICU and put them in a non-monitored bed?

MR. DUDLEY: So no one has done

randomized trials of taking people and putting them in different situations. The longer you are in the ICU, or taking patients at different points of time of equal status, the longer you are in the ICU, the more likely you are to get infections, other iatrogenic complications.

I guess what I was trying to say is that, on the whole, while there are some measurement issues related to moving a patient from the ER and also to moving them out, the hospital community has not raised that as a significant issue in terms of feeling bad about my ICU length of stay being measured this way. And that is in a community that is very large and includes all types of ICUs. Because, on the whole, they also feel like they agree with the intent of this measure. Shorten this and things will be better. And if my problem is in the ER, then I want to work on that.

CHAIR YAWN: Okay.

DR. NEFF: For me at least, it tends to keep -- and it is probably my own little internal world, where it feels like that sniff test for me, but my world is so different because I have a step-down or I don't or I'm boarding or this. There may be that sniff test that feels off if it is not included, although your experience with the variety of hospitals would actually speak against that because it doesn't seem to be

wigging them out about that.

But it may come up, actually, it is probably less relevant for the integrity or the high impact of the value, which I think is probably pretty solid, and it may come up a little more that we might be able to just outline that as an issue, whether it ends up being a real --

CHAIR YAWN: Well, usability, it may be an issue.

DR. NEFF: Or in the science.

CHAIR YAWN: Okay.

DR. NEFF: So I think we are solid here at least.

CHAIR YAWN: All right.

DR. NEFF: So "C's" across the board.

DR. RASTOGI: Could I make a comment, even though I am not a measure developer, just from the science point of view?

I had participated in this treatment, health quality choice project, and also the Anthem, Blue Cross/Blue Shield, and it was more for cardiovascular. So it wasn't pulmonary.

But for people doing this, we were doing ICU length of stay or even hospital length of stay as an adverse outcome measure. We would do risk adjustment for patient issues, but then for the hospital-based efficiency issues we do control charts.

So, like what you were saying, Barbara, within a hospital, you could look over time and see. So that is the efficiency piece. And you can't compare one hospital versus the other because they have their own problems.

So the control charts would help us parse out, but it is a quality-of-care issue or it is an efficiency.

CHAIR YAWN: Yes, and I think that it can, but, again, we have to take the measure the way, and it is supposed to be for public reporting, to compare across hospitals, but it is a potential solution.

DR. RASTOGI: Yes.

CHAIR YAWN: Good. Thank you.

DR. NEFF: So on to 2, on to the science. This is probably where this comes up, for me at least, the most about precisely specified. So, basically, the criteria being all eligible patients admitted to the ICU, and basically getting the time, being the time from discharge minus the time of admission, I mean fairly straightforward, using vital signs to kind of track those start and stop times.

And I think this is all pretty straightforward, pretty precisely defined, with the only caveat in my mind being this issue of where you start the ICU, if it is in the ED or in the PICU. You know what I mean? It is a little hazy around the edges.

And would that matter? Again,

internally it wouldn't. If you had a lot of variability in hospitals, and just that hospital infrastructure piece, is there a way to get your brain around it and describe it in some way that helps people understand; whether it can be a model or not, it may be hard to do.

So this is probably, more than anywhere else, I might say I am not sure. For me, for that reason, I put it as a "P" instead of "C", and I might just lay that out as my sort of comment in the weakness piece. It is just that that makes it harder to generalize and be able to say you are really solidly sure about what an ICU length of stay is, if you are not entirely sure how you are defining the ICU.

MS. PACE: So what is unclear about how they defined ICU?

DR. NEFF: Basically, that it is the time -- well, how they defined it is probably clear. That is fair enough. So it is just whether it is accurate for the entirety of the population. That would be all.

CHAIR YAWN: So that you think there might be some systematic bias that certain hospitals would always have shorter length of stay because they've got these patients trapped in the ER for 24 hours before they ever get to the ICU?

DR. NEFF: Or longer, because they

are stuck in the ED or in the ICU and can't get back to the acute care. I mean it could go both ways.

CHAIR YAWN: Yes, and that may be what you are getting at. It is a systematic -- it is not a random --

MS. PACE: But the trapped in ER would shorten.

DR. NEFF: Uh-hum.

MS. PACE: So a hospital could look better if they are holding people in ERs? Is that then --

DR. NEFF: It is just that it feels like there is a hospital infrastructure issue in terms of flow that could impact this outcome measure that internally within a hospital wouldn't matter. Would it be hard to say, Hospital A and B and C, that it doesn't matter within all of those three, even though they have different flow issues? And whether there would be a way to, again, not trying to solve the measure problem, but to either group those hospitals -- I mean you could almost imagine hospitals with similar sort of hospital structures would be compared. You could still compare, but --

MS. PACE: But part, I think, of what has been presented and what has been said is that, if that is the issue, the hospital should fix the flow.

DR. NEFF: Uh-hum.

MS. PACE: And if I were looking,

you know, regardless of what the problem is, I would rather go to the hospital --

CHAIR YAWN: No, you wouldn't rather go to the hospital that has a shorter length of stay because you are stuck in an ED for 48 hours.

MS. PACE: Right. Right. Exactly. So is that the main issue, that people --

CHAIR YAWN: Less is not better necessarily, if that is a main issue. Now I don't know the size of that issue.

MS. PACE: Yes. Right.

CHAIR YAWN: And I cannot say that is a problem --

MS. PACE: Right, right.

CHAIR YAWN: -- in 48 percent of hospitals. If it is a problem in 3 percent or 1 percent of all ICU admissions, forget it. Do you know?

MR. DUDLEY: Well, this is, again, where I don't know if you are comparing things. So how that will play out --

CHAIR YAWN: It has to compare across hospitals.

MR. DUDLEY: No, I meant across measures.

CHAIR YAWN: Oh, okay.

MR. DUDLEY: But we are proposing at the same a mortality measure. So, if you trap the patient in the hospital, I'm sorry, in the ED, that will actually usually play out as worse outcomes.

First of all, it is not a huge issue. It is more of an issue sort of temporarily when the flu season kicks in, and it has been more of an issue with swine flu. But, for the most part, it is not a big issue. It is a seasonable phenomenon, not so much a variation across hospitals as across time during the year, because it is mainly respiratory waves that cause the ICU access problems.

But it has not played out to be a very big issue here. But I think when you have these flow problems, and you add the mortality measure, which we haven't yet discussed, then if you tried to game your length of stay, that would probably play out within your mortality measure.

CHAIR YAWN: But, unfortunately, you know, we have to do that separate because we cannot be assured that the mortality measure will be used by the same people that use the length of stay. Okay? You could make it a composite.

MS. WINKLER: No, not a composite.

CHAIR YAWN: Okay.

MS. WINKLER: Actually, NQF has done in the past paired measures, so that you do the two together. You don't do them separately. You don't do them independently. You say, if you are going to do one, you are going to do them both. That could be a recommendation from this group.

MR. DUDLEY: Can you say that one would be paired without the other? So I don't think you want to say only do mortality if you are also willing to do length of stay, but I do think you want to say only do length of stay if you are also willing to do mortality. Does that make sense?

CHAIR YAWN: A one-way pairing?

MS. WINKLER: A one-way pairing. Okay. You can make that --

MR. DUDLEY: And if I could, Barbara, with respect to your concern that smaller hospitals wouldn't be able to create a step-down because they can't hire more people, if you actually have patients who are step-down-worthy, that actually means hiring fewer, rather than more, people because it is lower ratios of care when you are able to switch people to a step-down.

CHAIR YAWN: That sounds like a good theory. You come and see our 60-bed hospital and tell me about it then.

MR. DUDLEY: Well, you can just declare parts -- we have had hospitals just declare parts of their ICU as the step-down and go four-to-one now.

CHAIR YAWN: All right.

DR. NEFF: I mean I think that every hospital has struggled with this in some way or another. I think, to a great extent, as you are saying, I mean you kind of need to know if this is a problem within your hospital

and try to figure out what, if anything, you want to change and at what end of the flow you need to change that. It is just when you start, then, getting into the hospital-to-hospital comparison, that that gets a little trickier.

So, I mean, I think maybe the pairing is a possibility, if we wanted to put that as some way to get it to kind of all balance out in the wash.

CHAIR YAWN: So we would say for this one we recommend it is always paired with mortality?

DR. NEFF: If you are going to do this, you do mortality.

CHAIR YAWN: Okay. Good.

DR. NEFF: Yes. And I think from just the way I think I wrote this, I ended up putting it as a "P", just because that was sort of gnawing on me a little bit.

CHAIR YAWN: Are you comfortable, with the paired, now moving it to a "C" or not? That's fine. You don't have to. I don't want to take forever.

DR. NEFF: Yes. I guess, because I still, even with that, honestly --

CHAIR YAWN: Sure.

DR. NEFF: -- I probably still feel like there are issues --

CHAIR YAWN: Fine.

DR. NEFF: -- but I think it would solve the issue in terms of its being

functional by having it paired.

Then, just in terms of where I stuck all this stuff, it ended up being kind of in the descriptor of the pros and cons or weaknesses and strengths, so just that ability to assess this sort of hospital infrastructure's impact on length of stay.

Then, let's see, 2b -- or not to be.

(Laughter.)

Reliability testing I thought was quite solid. It was large patient population, a number of different hospitals, large range of time, well done with random sampling and auditors, and yadda yadda ya. So I gave that a "C". Yadda yadda ya.

CHAIR YAWN: Good.

DR. NEFF: So why the yadda yadda ya?

(Laughter.)

MR. DUDLEY: The story of my life.

(Laughter.)

You can do that, as long as my wife doesn't.

(Laughter.)

CHAIR YAWN: Your kids will, if you ever have any, I can assure you.

(Laughter.)

MR. DUDLEY: I've got two, but they are not there yet.

CHAIR YAWN: Oh, okay.

DR. NEFF: And then validity, did

it in a 40 percent sample. So, basically, of the group, sort of tested this in 40 percent of the population, which is a pretty reasonable chunk of patients.

I can't actually give you a reason why I had it as a "P". I think I perfectly fine with that as a "C". I must have had a little slip of the click.

CHAIR YAWN: Okay. So we get the "C" there.

DR. NEFF: Because I was looking at the rest of my notes, and I didn't have anything else that was bothering me there.

CHAIR YAWN: Okay.

DR. NEFF: Then the justifications for the exclusions, the issues here were excluding patient populations who had well-established other risk stratification and adjustment models. So burns, trauma, post-MI, post-CABG, and readmission. And that all made sense.

I think the only reason I put it as a "P" instead of "C" was that it would still be -- and this maybe goes into more of a feasibility maybe than here. Okay, that's fine, but then what about the hospital that has all those patients? How do they, then, equate their length of stay if they are doing it in a smaller subset of their population? It seemed like it would be a little harder to use, particularly if you had hospitals that had large volumes of these patients.

CHAIR YAWN: So that might be more feasibility?

DR. NEFF: I think it is going to be more feasibility, now that we are sort of in the rhythm of this whole thing.

CHAIR YAWN: Yes.

DR. NEFF: So I would be "C" here.

CHAIR YAWN: "C"? Okay.

DR. NEFF: I think this all made sense. It was well-supported for why they kept those people out of that.

CHAIR YAWN: Okay.

MR. DUDLEY: If you could bring that up again when you get the feasibility, I will address it.

DR. NEFF: Okay.

CHAIR YAWN: Okay.

DR. NEFF: Let's see, the risk adjustment, actually, issue for the outcomes measures, they are using all the variables. I think my only question just was really kind of -- and I think I understood it better as I then got into the model paper that you are saying in terms of what you actually added to this. So I would be actually up at a "C" now.

You added sort of the contemporary information. Then you added the code status and the ICU time, time prior to ICU. So a couple of new measures that, then, helped make the model work better in terms of mortality. So I actually was fine there as well.

CHAIR YAWN: And you have cut out

the people that were less than four hours and you cut out the over 30 days. So you've got the outliers on both sides.

MR. DUDLEY: We didn't cut out; we truncated.

(Laughter.)

CHAIR YAWN: Yes. I am not a surgeon. I can cut them out.

(Laughter.)

MR. DUDLEY: Okay.

MS. PACE: So could I ask, the information you put in validity testing is actually about your risk model? Is that your risk model performance?

MR. DUDLEY: Do you mean back on 2c here?

MS. PACE: Yes, 2c, right.

MR. DUDLEY: If you don't mind just showing me what's in 2c, just so I can make sure I am stating it correctly.

Yes, this is about how we validated the model. So we built it on 60 percent of the sample of roughly 11,000.

MS. PACE: Right.

MR. DUDLEY: Then we validated it on 40 percent. Those were randomly sampled before we got started.

CHAIR YAWN: So a split sample validation?

MR. DUDLEY: Yes.

MS. PACE: Okay. So, when you came down to risk adjustment testing, and you put

testing results not applicable, and I am just looking at this quickly, but did you report like discrimination in calibration statistics for your model?

MR. DUDLEY: Yes, those are up in the reliability --

MS. PACE: In the 2c?

MR. DUDLEY: Yes, in the reliability and validity sections, yes.

MS. PACE: Okay.

MR. DUDLEY: And they are in the paper. So the c statistics are .83, for instance, and -- I'm sorry -- oh, and the calibration, we can show you the calibration terms. They are in the reference.

MS. PACE: Okay.

MR. DUDLEY: But they look okay.

MS. PACE: That is for the future. We realize we are going --

MR. DUDLEY: We actually struggled a bit with what goes where.

MS. PACE: -- to have to provide more specific guidance, but --

MR. DUDLEY: Well, part of that is that so many different kinds of measures with different kinds of validation can come in.

MS. PACE: Right.

MR. DUDLEY: So you might consider, oh, well, if this is a mortality model, then this approach.

MS. PACE: Right. That is what we need to do, I think.

CHAIR YAWN: Yes, and I think, also, making sure we have the references attached whenever you have the curves and other things will be something you should be able to let people try to do.

MS. PACE: Right. Well, we did have -- did you submit those, the risk model information?

MR. DUDLEY: I think so. I think so. We certainly -- it definitely was in there.

MS. PACE: So it is in the attachments?

DR. NEFF: And there are two more articles that are in there.

MS. PACE: Right. No, they are here.

CHAIR YAWN: She was able to look at them. That is the important part --

MS. PACE: Right.

CHAIR YAWN: -- is that whoever is assessing it can look at them.

DR. NEFF: Right. There was an MPM-III LOS model and a MPM-III model which was the original.

MS. PACE: Okay. Great.

DR. NEFF: And then, I think in terms of the meaningful differences, the length of stay and then the adjusted length of stay is actually well-described and compared to SAPS and APACHE. So you kind of have a range of what might be expected. So I think

that was actually reasonable.

And I didn't see anything that was actually off there. Then, when we get down to comparability, basically, there's really nothing for this to compare to that I could find, either. So I think that was reasonable.

CHAIR YAWN: Reva?

MS. WINKLER: Yes, a question. You said it is a CHART measure, right? Is it being publicly reported?

MR. DUDLEY: This one is not yet publicly reported.

MS. WINKLER: Not yet, but intended to?

DR. NEFF: But the earlier version, I mean the earlier use of the MPM is.

MR. DUDLEY: For mortality --

DR. NEFF: For mortality.

MR. DUDLEY: -- it is publicly reported.

CHAIR YAWN: But not for length of stay?

MR. DUDLEY: For length of stays, you get to look at it and work on it for a year, sometimes a year and a half, before we do the publicly reporting, when we develop a new measure.

And this one, as you pointed out, we had to develop ourselves.

MS. WINKLER: Right. So you actually have data from different hospitals on this in terms of --

MR. DUDLEY: Yes, that data you just saw there.

MS. WINKLER: Yes. Okay. And I guess the question would be, there is a plan to publicly report that down the road?

MR. DUDLEY: Uh-hum.

MS. WINKLER: Okay. And the discriminatory characteristics of it? Of the results?

MR. DUDLEY: Are you asking if there's variation?

MS. WINKLER: Yes, exactly.

MR. DUDLEY: So, if you could go back to that data, the one that had the -- go up one.

DR. NEFF: 2f.

MS. WINKLER: Yes, I just didn't know how to interpret.

MR. DUDLEY: So the range, originally, we were deciding which models. So, for APACHE IV and MPM-III, too, the range of your standardized length-of-stay ratios is from, roughly, .4 to, roughly, 1.6 across our hospitals. So a really big range.

MS. WINKLER: Yes, okay. Got it. Okay.

CHAIR YAWN: Let's go on.

DR. NEFF: Then disparities in care, actually, I think, in reality, it is sort of not stratified on it. There's not any reason to suspect there would be any differences, and there's not anything

currently reported with this or to compare it to. So I think that I am not sure if that would actually fit into more of a not applicable than a complete.

CHAIR YAWN: You don't think there are any racial differences?

DR. NEFF: Well, I guess it is whether it is in the model. I am not certain. Let me see.

I get a little lost with this when there's not --

CHAIR YAWN: Well, I guess what we say, is there the ability for --

DR. NEFF: Yes.

CHAIR YAWN: -- this measure to say, look, it's always a longer length of stay for African-Americans after you have done all of the other risk adjustment? Can this model do that? Or is race one of the risk adjustment factors?

MR. DUDLEY: No, we never risk-adjust by race.

CHAIR YAWN: Yes.

MR. DUDLEY: So the model does not include that variable.

CHAIR YAWN: But it could?

MR. DUDLEY: You could then; we do not.

DR. NEFF: No.

CHAIR YAWN: I understand that you do not.

MR. DUDLEY: No, we do not

currently plan to, but you could say here's the LOS for African-Americans, for Asian-Americans, for Hispanics, and so forth. You could do that.

CHAIR YAWN: So, right now, it does not measure any of those types of disparities, but it potentially could?

DR. NEFF: Built into it, but not reported.

DR. O'CONNOR: So, then, why was it based on hospitals with 400 admissions? Isn't that correct data?

MR. DUDLEY: Oh, a random sample of 400 of their admissions.

DR. NEFF: So it is in the data, but it is not reported out by race? I mean it is there, if somebody felt inclined or wanted to --

MR. DUDLEY: Yes, you could do that, yes. So we do ask race.

DR. NEFF: Yes.

MR. DUDLEY: The way that it works, actually, though, is somebody looks at you.

DR. NEFF: Yes.

MR. DUDLEY: So, if it is really obvious, then --

CHAIR YAWN: Well, we know that that is not a good measure. We understand that.

MR. DUDLEY: Yes, exactly.

CHAIR YAWN: But we don't know what are good measures of race or ethnicity anyway.

So it's fine. But it could be by age; it could be by gender. We can usually tell gender, usually.

MR. DUDLEY: Yes.

CHAIR YAWN: Not always.

MR. DUDLEY: We had four of those in California.

CHAIR YAWN: Yes, I know.

(Laughter.)

DR. NEFF: Okay. In the strengths and weaknesses, I think, actually, we addressed any of the other issues that I had, which is really just about feasibility. So I don't --

CHAIR YAWN: In this, the only thing I see here that you haven't measured or you haven't mentioned is patients' treatment decisions based on family goals and values. You did do something about their code status. So you have dealt with that. And I am not sure how you would ever do that, but --

DR. NEFF: I mean, on the assumption that that decision is based on family --

CHAIR YAWN: Yes.

DR. NEFF: -- and family wishes, it is built in.

CHAIR YAWN: Yes.

DR. NEFF: Full code, DNR, limited interventions, comfort care.

MR. DUDLEY: What we do is there are so many choices there because there are

really so many levels that could play out.

CHAIR YAWN: Right. It is gradation.

MR. DUDLEY: But what we actually do is we only use the full code. So it is binary full code, yes/no.

DR. NEFF: Or it is something else.

MR. DUDLEY: Because the gradations have too many different meanings, it gets dirty when you go down into the lower stuff.

CHAIR YAWN: Oh, that's fine. So you have addressed this. He mentioned it as the weakness. You have addressed it to some extent, the only extent you probably know how to?

MR. DUDLEY: Yes. We actually tried other ways, and it just got too messy.

CHAIR YAWN: Okay.

DR. NEFF: So it is basically the full code and everything else.

MR. DUDLEY: Versus others, the full code versus others, versus limitations on care.

DR. MILLARD: One of the issues we face in our hospital is when the treatment team says stop and the families say continue. That is actually not captured here because they are full code. Yet, that makes a significant difference in terms of total cost and total length of stay. Because if you have a number of -- and sometimes that is culturally-driven. So that would be a

potential relative weakness, is if there is a discordance, and I don't know how you track that, between treatment team decisions and family decisions, that could be a significant influence on length of stay, although you cut it off at 30 days.

But what we will find is that it adds, often family conflicts in decisionmaking in critically-ill patients often will lengthen, add one or two weeks to the resolution of the case.

DR. NEFF: There would be sort of like the team length of stay and the family length of stay or something.

DR. MILLARD: And I don't know --

MR. DUDLEY: We haven't been able to deal with that. The thing is everyone thinks they face that. So there hasn't been a lot of --

DR. MILLARD: So you feel like it washes out?

MR. DUDLEY: Yes. I mean that definitely happens. It happens all the time.

DR. MILLARD: But maybe culturally some areas it happens --

DR. NEFF: Yes, and it would be almost like if you had a really active palliative care service, you would get them involved in kind of cultural family/team communication stuff, and you could have a quantitative variable for that. But that is more qualitative-based --

MR. DUDLEY: So is this the point where I would be addressing your earlier concern about feasibility? What if you have a lot of trauma patients or burn patients? I mean earlier you said --

CHAIR YAWN: We are not on feasibility yet. We are still on usability.

MR. DUDLEY: Okay. All right.

DR. NEFF: One more step.

MR. DUDLEY: I didn't want to miss it.

CHAIR YAWN: Go on. Okay. Usability, go for it.

DR. NEFF: Usability. So, currently, being used already for QI purposes, the mortality risk prediction model is already in use. No current usage, although anticipation of using reporting for the ICU-I to study.

So it already has some track record, essentially, with its use. It is just in a slightly different realm, but the same overall model that has now just been tweaked, I guess. A little tweaking.

MR. DUDLEY: Okay, yadda yadda yadda, tweak.

(Laughter.)

DR. NEFF: Okay. For the relation to other measures, you mentioned the PICU length of stay, which already has an NQF member unit assigned to it.

There is sort of the link to the

mortality, although it is a different measure as well. That kind of brings up the harmonization category. Really, I mean adult is adult, and the only difference here, obviously, with the other measure is peds; the ages of the those are appropriately different.

CHAIR YAWN: And we are not going to say, "peds' ages". We are going to say, "children".

DR. NEFF: Children. Okay.

CHAIR YAWN: Parents don't have pedias. They have children.

(Laughter.)

DR. NEFF: Kids have two feet, though.

(Laughter.)

MR. DUDLEY: You're discriminating against the one-foot children.

(Laughter.)

CHAIR YAWN: You are also discriminating against family physicians. They are children, please.

(Laughter.)

DR. NEFF: And then added value to the other measures is really not actually applicable because there aren't other measures that are equivalent for this at this point.

MS. WINKLER: In fact, that, actually, is of some value, is the fact that, since there aren't measures --

DR. NEFF: Right, right.

MS. WINKLER: -- and you have

declared it to be very important --

DR. NEFF: Right. Yes.

CHAIR YAWN: So it is not a weakness.

DR. NEFF: Right. I thought, from a strength, that what is being used is understandable, currently in use. The concept of length of stay is one that is familiar, certainly, with the healthcare community, and you can describe to people.

I guess, assuming this is factoring in survival, so it is length of stay among those who survived or just overall length of stay?

MR. DUDLEY: It is overall length of stay.

DR. NEFF: Overall length of stay.

MR. DUDLEY: But, again, there is the pairing thing.

DR. NEFF: Yes.

MR. DUDLEY: You wouldn't well on your length of stay in the waiting room but die on the first day.

DR. NEFF: Yes.

MR. DUDLEY: But that is why we pair it.

DR. NEFF: Yes. So another good spot in here for the pairing. Because, otherwise, it should be length of stay among survivors, but it is not, because they are pair it.

CHAIR YAWN: Yes.

DR. NEFF: But we are going to encourage it for whatever.

CHAIR YAWN: Yes, we are going to say it shouldn't be done separately.

Go ahead.

DR. NEFF: There you go.

CHAIR YAWN: Yes. You've got it.

DR. NEFF: I was trying not to --

CHAIR YAWN: No. It is good. It is very good. Thank you.

DR. NEFF: So that is another great example.

CHAIR YAWN: Okay.

DR. NEFF: All right. So we are through 3. So now we are on to feasibility.

CHAIR YAWN: Now we are on feasibility. Now your issue?

DR. NEFF: Right. So hang on.

Data generated as a byproduct of the care processes, yes, you still need to actually have somebody abstracting. Right. So it is not 100 percent, actually, totally electronically available, unless I am interpreting that wrong.

MR. DUDLEY: That is correct. We were thinking that you meant you have to do something to the patient to get the data.

DR. NEFF: Ah, okay.

MR. DUDLEY: And the answer there is no.

DR. NEFF: Okay.

CHAIR YAWN: And in some sense,

this is very soon will be totally electronic.

DR. NEFF: Yes, but --

MR. DUDLEY: I wouldn't say very soon.

CHAIR YAWN: Okay.

MR. DUDLEY: It could be today, but the case of EHR adoption in the real world is --

CHAIR YAWN: Okay.

MR. DUDLEY: But, yes, there are places that have this electronically.

MS. WINKLER: But the idea is, with an EHR, these can --

MR. DUDLEY: This would be incredibly easy, yes.

MS. WINKLER: Right. Okay.

MR. DUDLEY: Yes.

MS. WINKLER: So it is very compatible with EHR?

DR. NEFF: Right. So it is not there, but it will be good when it is.

MR. DUDLEY: There actually are commercial products out there that have this in it. So there are vendors who are building it in.

CHAIR YAWN: Yes. Yes. Well, I just happen to know ours does. So, I figure if ours has it, everybody should.

MR. DUDLEY: You're a 60-bed hospital?

CHAIR YAWN: Well, they are part of -- yes.

DR. NEFF: So I think we are probably in a similar situation where the electronic piece kind of gets put down because it is not by virtue of the design of the study, but by just --

CHAIR YAWN: Scalability.

DR. NEFF: Yes.

CHAIR YAWN: How far down do you want to go?

DR. NEFF: I had it as an "M", actually, because I think you have to have an individual person to abstract data. I think there is just no way around that.

A solid "C" for the electronic piece, but then, because there is a chunk that you have to do by hand --

CHAIR YAWN: Okay.

DR. NEFF: I mean we should try to probably, from a harmonization standpoint, at least try to be consistent.

CHAIR YAWN: Or did we do "M" on the last one?

DR. NEFF: That's what I can't remember. It is the same issue.

CHAIR YAWN: Whatever we did on the last one, we are going to do on this one.

DR. NEFF: Because it is the same across the board.

CHAIR YAWN: Was it "M"? Okay, then we will do "M".

DR. NEFF: That can be a little SOP, as far as future taps, because it is not

speaking to the study as much as it is just a reality.

DR. O'CONNOR: I am curious --

MR. DUDLEY: I am sorry. I missed that. What is "M" then?

DR. NEFF: Because that is the electronic sources, whether all the data is available electronically.

MR. DUDLEY: Oh, okay. So that's "M".

DR. NEFF: Yes.

DR. O'CONNOR: That is more of a statement about the hospital.

DR. NEFF: It is a statement about --

CHAIR YAWN: Yes, it is not the measure.

DR. O'CONNOR: But, of the 200 hospitals or so that you deal with in California, what percent do you think don't have this electronic capability? I mean, certainly, in San Diego we --

MR. DUDLEY: Unfortunately, still the large majority don't have --

DR. O'CONNOR: Really?

MR. DUDLEY: So there are products out there where you are running your ICU and the blood pressure is being recorded electronically, but it is not, then, put into a risk-adjusted system that pops out a mortality calculation.

There are also products that do

that last step, eICU, VisICU, et cetera, but most hospitals, the large majority of hospitals do not yet have those things. So I guess, in that sense, yes.

But almost everything is an "M" then, right, on this one?

MS. WINKLER: No. The data generated is a byproduct of care is what they are putting the "M". The electronic source is still "C" because it is available.

CHAIR YAWN: Oh, yes.

MS. WINKLER: People will start using it.

MR. DUDLEY: Oh, I see what you are saying.

CHAIR YAWN: Yes.

MR. DUDLEY: Oh, okay.

CHAIR YAWN: It is just that, for people that don't have it, they don't have it.

MR. DUDLEY: Okay. Yes.

MS. WINKLER: Yes, that is one of the issues. I mean you have kind of got a dichotomous situation for the first one. If you've got it, then it is fully electronic. If you "don't got it", then it's not.

CHAIR YAWN: Then it is a problem. So that is why, that's what gets the "M", is the fact that not everybody has it, and we are not close to 70 or 80 percent having it yet. So that is why it gets an "M".

MR. DUDLEY: Has there ever been anything, then, that isn't an "M" on that?

CHAIR YAWN: Pretty much no.

MR. DUDLEY: Oh, okay.

MS. WINKLER: Not much. Not much, no.

CHAIR YAWN: No. I mean it doesn't bring it down --

MS. WINKLER: Pure admin data.

MR. DUDLEY: Okay.

CHAIR YAWN: Yes, but there aren't any. Well, alive or dead.

MR. DUDLEY: Yes.

CHAIR YAWN: You know, you ca pretty much get alive or dead.

MR. DUDLEY: Yes.

DR. NEFF: So I may have had mine switched in here, but I agree at this point that the "M" is only relating to the availability of electronic.

CHAIR YAWN: Yes.

DR. NEFF: Okay. Then we get exclusions do not require additional data sources. So that was a "C". We were fine with that.

Then I think this might be a place we can sort of talk the sort of susceptibility to inaccuracies, errors, or unintended consequences. You were bringing up sort of the issue of people might try to game things, so that they got lesser sick people, so their length of stay was lower. I mean that is always a possibility.

But I think, in reality, hopefully,

there would be enough hospitals involved that it wouldn't just be a regional phenomenon.

MR. DUDLEY: Yes. And again, that is true, that is applicable to every outcome measure.

DR. NEFF: Yes.

MR. DUDLEY: Then the specific issue that you had raised was, well, what if I have a lot of trauma or burn patients? Or the other big group is coronary bypass. Those hospitals that do have big trauma units, big burn units, big coronary bypass or coronary surgery units, they don't have small other ICUs. They have literally thousands of other ICU patients. So it doesn't actually for them -- they are much happier.

We actually have a separate risk-adjusted measure that we publicly report for the coronary bypass. If we try to push the coronary bypass patients into that thing, the thoracic surgeons would have a fit because --

DR. O'CONNOR: We have seen that.

MR. DUDLEY: Yes.

CHAIR YAWN: They have fits, you know --

MR. DUDLEY: No comment.

DR. NEFF: So you would sort of parallel that with you would have a separate sort of mortality measure for them? There would be a separate length-of-stay measure for them?

MR. DUDLEY: Correct.

DR. NEFF: I mean the same sort of --

MR. DUDLEY: Yes.

DR. NEFF: So it would be kind of parallel in that way?

MR. DUDLEY: Yes.

DR. NEFF: Okay.

CHAIR YAWN: Okay. So that really is not a concern then. Okay. So that gives us a "C" there.

DR. NEFF: Uh-hum.

CHAIR YAWN: All right. Is there a 4 --

DR. NEFF: I think 4e, data collection strategy and implementation, I think actually you have answered. The only thing I had had in there was whether it was really going to be easier or not than some of the other sort of acuity assessment and length of stay, sort of as you described the APACHE and SAPs and other things that actually took longer than this.

So I would probably put this up to a "C" then.

CHAIR YAWN: All right. And we've got, let's just see quickly --

DR. NEFF: Oh, I did have one other question. Why just the first 100 patients in each quarter rather than a random sampling throughout a quarter? Are you worried at all about early-in-the-month bias or anything like that?

MR. DUDLEY: Well, everyone is doing it the same way. So the point is comparability. Also, by doing it that way, you just make it easier for them. So, if you have to do random samples, then they have to send it out to someone who does the randomization.

DR. NEFF: Yes, that's correct.

CHAIR YAWN: The only problem is academic centers in July, when they get new residents, you know, their July quarter may be worse than the other two because the first 100 are probably taken care of.

MR. DUDLEY: Right, but they actually --

CHAIR YAWN: Well, they run into that.

MR. DUDLEY: No, I mean they want it to go this way, too. Everybody wants it. It is a big operation to do this. So they want it to be easy.

CHAIR YAWN: No, I agree.

DR. NEFF: And presumably, you could also, in the same way that we do for a lot of measures, you could also classify different hospitals in different categories, right? So you could not only look at overall comparison hospital to hospital, but if you wanted, then, to look at hospitals less than 200 patients or ICU volume per year, or whatever, or academic? And then you could sort of say, well, this is why we are worse

overall, but look how we are compared to those of our own?

MR. DUDLEY: Right, but we actually, for the public reporting option --

DR. NEFF: Yes.

MR. DUDLEY: So the hospitals get their data quarterly. For the public reporting option, it is a rolling 12 months. So, if you have a July problem, that is one quarter of your total data.

DR. NEFF: Uh-hum.

CHAIR YAWN: Well, and you have a July problem every year probably.

Okay. Are we finished now?

DR. NEFF: Uh-hum.

CHAIR YAWN: Okay. Very good. Well, thank you very much.

We are going to take a quick break for lunch. We are going to have a working lunch because we hope to get through, we need to get through the mortality measure for this, and we have got to get to at least one of your measures, since you are here. Well, I said at least one. And we have about an hour and a half to go.

So lunch, please. Grab it and come back.

(Whereupon, the foregoing matter went off the record at 12:29 p.m. for lunch and went back on the record at 12:40 p.m.)

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

12:40 p.m.

CHAIR YAWN: You're just flip-flopped, but because you are here and he is not, I guess you get picked on.

DR. NEFF: Right. Okay.

MS. FORMAN: Which number?

DR. NEFF: Oh, 24.

MS. FORMAN: Twenty-four.

DR. NEFF: Okay. Same player, same bat channel, so to speak, and now, essentially, with a mortality measure.

So kind of going through, again, using mortality as a surrogate really for quality of care, all of the things that can matter to hospitals, to patients.

CHAIR YAWN: Being alive or dead does matter.

DR. NEFF: It does, yes. I don't know if there's too much to say about the impact.

Lots of patients; unfortunately, still lots of deaths. Big cost. Opportunity for improvement.

CHAIR YAWN: Okay. So 1a --

DR. NEFF: Yes.

CHAIR YAWN: -- is a "C"?

DR. NEFF: 1a is a "C".

CHAIR YAWN: Okay.

DR. NEFF: 1b, I underlined "mortality variability for patients admitted to the ICU persists", which is true. This is

in their dataset and nationally as well. Lots of factors that go into that, but, certainly, always on the list of everything everybody wants to track. So opportunity for standardizing tracking, comparing, publicly reporting.

CHAIR YAWN: So 1b is a "C"?

DR. NEFF: As well.

CHAIR YAWN: Yes. Okay.

DR. NEFF: And then, I am sorry I have those on here, instead of on my remit.

Outcome or evidence to support the measure focus, I think really, basically, I don't think I've got anything here that was -- trying to prevent death. Basically, whether you look at randomized trials, observational studies, all risks, things that you are always having mortality as an outcome measure. I don't think, unless there was something otherwise very specific in this particular subset, I had it as "C" as well.

CHAIR YAWN: Yes, this one seems to be a pretty straightforward "C".

DR. NEFF: Yes.

CHAIR YAWN: Okay. We got that.

DR. NEFF: Okay. It is like trying to find something in a category --

CHAIR YAWN: Do we have strengths, weaknesses? Okay.

DR. NEFF: For the measure, lots of variability across institutions. It is hard to know what variables might impact this

outcome, though. So you never know for sure if you have everything in your model, but nothing unique to this measure.

MS. PACE: Does the model have a socioeconomic status in it?

CHAIR YAWN: Yes. So you could look by this --

MR. DUDLEY: It does not -- sorry. It does not, and there will be disparities.

CHAIR YAWN: Yes, and so we could assess disparities.

MR. DUDLEY: Yes.

CHAIR YAWN: So there is a reason for it not to be in the model.

DR. NEFF: Right. It is not there --

CHAIR YAWN: But you pull that information usually in some way when the patient is admitted to the hospital.

DR. NEFF: You haven't lost it by adjusting for it. So it is there.

MR. DUDLEY: No.

CHAIR YAWN: Right. That is, I think, the important part because he says it doesn't take into account and adjust for it. Well, we didn't want it to adjust for it. So that is okay. So what is the weakness actually is sort of a strength.

DR. NEFF: Is a strength, yes.

MR. DUDLEY: It is a strength, yes.

CHAIR YAWN: Right.

MS. PACE: And that is actually in

our criteria, the guidance in our criteria.

CHAIR YAWN: Yes. But when we send this on --

MS. PACE: Right, right.

CHAIR YAWN: -- we need to move it to a strength instead of a weakness.

Okay, 2?

DR. NEFF: Pretty darn precisely specified.

(Laughter.)

CHAIR YAWN: Okay.

DR. NEFF: You know, basically, they have the same criteria that we just talked about in the last study in terms of ICU for four hours, greater than 18. So this is adults, not the kids. Isolating out the traumas, burns, CABG patients, for all the reasons we discussed the last time.

I didn't see that anything there seemed odd.

You just have to ignore my "P's". I think I was in one of those, like when you have evaluate and you have the five, and you always do four, and you never do five. I have nothing specific on this other than --

CHAIR YAWN: I just had a question about 18.

DR. NEFF: Age of 18?

CHAIR YAWN: Yes. This says greater than 18.

DR. NEFF: Oh, 18 and over.

CHAIR YAWN: What do most things

say about adults? Don't most people say adults are 18 and over?

MR. DUDLEY: It should be 18 and over.

DR. NEFF: Uh-hum.

CHAIR YAWN: Okay. This one just happens to say greater than 18 years of age. So we need to --

MR. DUDLEY: It should be greater than or equal to.

CHAIR YAWN: Okay. So we need to change that?

MR. DUDLEY: Yes, please. Sorry about that.

CHAIR YAWN: Thank you. We will.

DR. NEFF: And actually, in their exclusions, they have less than 18. So that is great. So, if you are 18, you want in --

DR. NEFF: Too bad.

(Laughter.)

CHAIR YAWN: You just have excluded. Well, I think there are some 18-year-olds, yes. Let's move forward.

(Laughter.)

DR. NEFF: Okay. Let me go on down to b.

CHAIR YAWN: Okay.

DR. NEFF: Hold on.

CHAIR YAWN: And this is very similar to the way it was tested with the other one.

DR. NEFF: Yes, exactly.

CHAIR YAWN: So it was a "C" before, and --

DR. NEFF: It was a "C" before. I had a "C" again.

CHAIR YAWN: All right.

DR. NEFF: And I don't think I had any other -- I have no other. Five percent random sampling, auditors, blah, blah, blah.

About validity testing, the same thing, the 40 percent subset, the 60 percent doing the model.

CHAIR YAWN: Okay.

DR. NEFF: All very legitimate.

DR. MILLARD: I have on question.

DR. NEFF: Yes.

DR. MILLARD: It seems to me some of the data probably in our hospital around palliative care is that moving to hospice, this doesn't -- this says we don't care whether or not they were palliative care, if they were made DNR. This is just mortality, correct?

MR. DUDLEY: I they went to the ICU.

DR. MILLARD: If they went to the ICU.

MR. DUDLEY: If they have just come to your hospital for palliative and they never go to the ICU, then it is not an issue.

DR. MILLARD: But if they go to the ICU and then they are made palliative care, that still counts?

MR. DUDLEY: Right.

MS. PACE: Well, is it -- but DNR, you say?

DR. NEFF: No, not in this one.

CHAIR YAWN: No. They are still, even if they are not --

MR. DUDLEY: They are not excluded from the population.

MS. PACE: Oh, okay.

DR. MILLARD: So, unless they are made an DNR within four hours, no?

MR. DUDLEY: No, unless they are taken out of the ICU within four hours.

CHAIR YAWN: Right.

MR. DUDLEY: So the point here is some people are saying, oh, well, I admit people who are palliative care to the ICU.

DR. NEFF: Right.

MR. DUDLEY: The response from the community is: why on earth are you admitting those people? We are not going to give you credit for admitting those people to the ICU.

CHAIR YAWN: Well, and I think the four hours ICU is very reasonable because a lot of hospitals use it for post-op for a certain group of patients --

MR. DUDLEY: Right.

CHAIR YAWN: -- because they don't have recovery open in the middle of the night.

MR. DUDLEY: Right.

CHAIR YAWN: So I think that is one of the reasons for that four-hour, they are

out of there. Okay?

So 2c is "C"?

DR. NEFF: 2c is "C", yes. I had it as a "P", but only before the --

CHAIR YAWN: Because you were just into --

DR. NEFF: Well, I was into "P", but it was also my concern about excluding big chunks of people in these other categories. So I am up to "C" there now.

CHAIR YAWN: Okay. So 2b, justification for exclusion.

DR. NEFF: Oh, sorry. That is where I just was. Hang on.

Oh, yes, because 2c was a "C".

CHAIR YAWN: Uh-hum.

DR. NEFF: 2b actually is a "C" because we have explanation for the traumas, burns, CABG.

CHAIR YAWN: Okay.

DR. NEFF: So it is "C" there.

CHAIR YAWN: So it is fine there. So 2e?

DR. NEFF: And then 2e, this is a similar issue. This is, actually, I think, the same description we had for the others.

MS. PACE: Right, and it was actually the risk model stuff was actually under validity.

DR. NEFF: Yes. Validity, yes. It was just moved into the section, and all appropriately done without any concern. I

mean they have adjusted for what they know. I mean you know --

CHAIR YAWN: So "C"?

DR. NEFF: There were no glaring deficits that I could see.

CHAIR YAWN: All right.

DR. NEFF: So it is "C".

CHAIR YAWN: So 2e is a "C". And 2f?

DR. NEFF: Similarly, we have ranges -- well --

CHAIR YAWN: This one has been publicly reported. So this one we are using publicly-reported data --

DR. NEFF: Yes.

CHAIR YAWN: -- to show that there are --

DR. NEFF: The differences.

CHAIR YAWN: -- opportunities for improvement.

DR. NEFF: That's true. That makes it easier.

CHAIR YAWN: Yes.

DR. NEFF: And then, similarly, well, comparison in 2g, multiple data sources.

CHAIR YAWN: That would be "C"?

DR. NEFF: Yes.

MR. DUDLEY: So can I just clarify?

DR. NEFF: Uh-hum.

CHAIR YAWN: You may.

MR. DUDLEY: Because I afraid of what happens after I leave. So everything for

your review, all the way up through 2f now, is a "C"?

CHAIR YAWN: That is correct.

MR. DUDLEY: And those are from Margaret, Margaret's review through f is a "C"?

DR. NEFF: Uh-hum.

MR. DUDLEY: Okay.

DR. NEFF: By 2f is a "C".

MR. DUDLEY: No, I mean everything, 1a through 2f is a "C"?

DR. NEFF: Yes.

MR. DUDLEY: Okay.

CHAIR YAWN: So far.

MR. DUDLEY: I just want to make sure it gets recorded that way.

DR. O'CONNOR: I get the sense that I am getting this report card.

(Laughter.)

DR. NEFF: I think that is what is happening.

MS. PACE: Can I stop for one second? I just want to ask a question about, in the list of risk model variables, you have "received CPR". So received CPR when? Is that even if it is the final event?

MR. DUDLEY: No, no, no, it's not.

CHAIR YAWN: It is before.

MR. DUDLEY: Yes, this is before.

MS. PACE: Okay. I just wanted to make sure. Okay.

CHAIR YAWN: Well, and it is within

a certain period of time before.

MR. DUDLEY: Yes. So everything in this model goes within one hour either direction of ICU.

MS. PACE: Okay.

CHAIR YAWN: Okay.

DR. NEFF: Okay. So we were down to 2g. Actually, we had a "C" on that.

Then disparities in care, kind of the same issue we discussed. It is available in the model, in fact, because it wasn't adjusted for. So it could be analyzed, should it choose to.

And I think in terms of strengths and weaknesses, this will be data actually that, although sort of some version of it may be in the project impact not readily available for use, it is going to be publicly reported. It is easier to do than APACHE, all those sorts of things. So I didn't have any actually other issues.

CHAIR YAWN: Well, then, our other reviewer has a couple of issues. One is the change in code status. I don't know how you get that.

DR. NEFF: Which point is he? Sorry.

CHAIR YAWN: It is c; although it takes into account full code versus full code, it doesn't take into account somebody's status changes while they are in the unit. I have no idea how you would ever do that, and I would

think that might be biased potentially, too. Because if someone gets really sick, they might change.

DR. NEFF: Right.

MS. PACE: And we would want risk variables to be present at the start there, not something --

MR. DUDLEY: As for Weakness A on the screen there that we don't get details of the severity of each illness, that is correct. So that is true. But we don't really view it as a weakness, but rather as a strength for the following reason: APACHE does, but if you use MPM or APACHE, you get the same assessments of hospital quality, but you spend three times as much effort getting the APACHE data. So it turns out that getting that additional data doesn't have a substantive impact on a hospital's performance ratings. So we intentionally not --

CHAIR YAWN: Okay. Thank you.

All right, and e and the weaknesses, and the APACHE and all of those, I think you have addressed those.

Lead time, we have addressed that previously.

Okay. So we are down to 3.

DR. NEFF: We are down to 3. Yes, it is pretty understandable.

CHAIR YAWN: Yes, it is pretty straightforward.

DR. NEFF: Yes.

CHAIR YAWN: Alive or dead, yes.

MS. PACE: And I think the other issue that I just was rattling in my brain is in a way, what does it add to kind of some of the other models that are out there? Obviously, some of the time-saving and then, also, I don't know if you would say other sort of more specificity or more easily done or this sort of approach to predicting mortality, as opposed to all the other ones that are out there.

MR. DUDLEY: The main thing is it gives you the same assessment. It is not quite as accurate on a patient-by-patient thing, but across hospital populations it gives you the same assessment for much less cost.

CHAIR YAWN: And it is not intended as mortality prediction score.

MR. DUDLEY: For individual patients, no.

CHAIR YAWN: No. So that doesn't matter.

DR. NEFF: Okay. So then, based on that --

MR. DUDLEY: I don't know how this process works, but we just rolled past in 3a a "P" from the reviewer.

CHAIR YAWN: No, I know. I saw it.

MR. DUDLEY: What are the implications of that?

CHAIR YAWN: If you go back up -- I

don't know. See, it says, "was not specifically tested for interpretability, but overall the website was tested and is widely used." I think that is why he gave it a "P", is that he didn't think you had tested it.

MR. DUDLEY: I guess, but is it any different than any other mortality measure?

CHAIR YAWN: Oh, no. No.

MR. DUDLEY: Okay.

CHAIR YAWN: But that is why we just, overall, it is a "C".

DR. NEFF: Yes.

CHAIR YAWN: I just chose to ignore his "P".

MR. DUDLEY: Okay.

DR. NEFF: Yes. Because in the context of the conversation --

MR. DUDLEY: In the process, does his "P" matter anymore?

DR. NEFF: No.

MR. DUDLEY: No?

DR. NEFF: Well, I mean it does to make sure that we are thinking about all of these things, but in the context of the added conversation here, it is more --

MR. DUDLEY: And what goes on to the Steering Committee will be the "C"?

DR. NEFF: Correct.

CHAIR YAWN: That is correct.

MR. DUDLEY: I do care. This is very important.

(Laughter.)

I care.

CHAIR YAWN: You started out as an internist once in your life, didn't you?

MR. DUDLEY: I did.

CHAIR YAWN: I am married to one. He is also obsessive/compulsive.

(Laughter.)

MR. DUDLEY: I just want it proved.

CHAIR YAWN: I know. I know.

DR. O'CONNOR: He prefers the word "focused".

(Laughter.)

CHAIR YAWN: I like the word "obsessive/compulsive".

MR. DUDLEY: I'll deal with OCD. That's all right.

CHAIR YAWN: If you live with one, believe me.

(Laughter.)

Okay.

DR. NEFF: So we are pretty much in the same --

CHAIR YAWN: 3e, "C"?

DR. NEFF: Yes.

CHAIR YAWN: It is the age issue. Now we know it is 18 and over.

DR. NEFF: Right.

CHAIR YAWN: I'm sorry.

DR. NEFF: And there is really nothing -- I mean adults -- children are children.

And distinctive or additive value,

in a way, it is not applicable, but it is also mostly entirely meaningful. So it says "C" and "NA" at the same time.

CHAIR YAWN: Right. Yes, it is good that it is not additive.

DR. NEFF: It is good that it is not additive, right. And I put easier than APACHE, would be data will be widely used.

CHAIR YAWN: All right.

DR. NEFF: So then we are down to feasibility.

CHAIR YAWN: Wait a minute.

DR. NEFF: Oh, no? Oh, sorry.

Patient and family goals, the same thing that we have dealt with --

CHAIR YAWN: A, the weakness, we have said they do take it into account in the DNR, as best we can figure it out.

B is not provide insights as to cause -- oh, okay.

CHAIR YAWN: Of poor performance.

DR. NEFF: Oh, yes.

CHAIR YAWN: Well, yes, okay.

DR. NEFF: But that is why we are tracking it.

MS. WINKLER: Do you have a comment on that? I mean in terms of how hospitals responded and used the data for quality improvement efforts.

MR. DUDLEY: Yes. The care is so heterogeneous, the poor performance can come from a lot of things. It may be your

ventilator strategy, whether or not and how you feed people, timeliness of antibiotics. It can be a ton of things.

We have two collaboratives, one in the southern part of the State and one in the northern part of the State, where they are getting together and people who do well are saying what they do. And some of it will be applicable and different to your hospital if you are doing poorly, and some of it will be the same. So you try to pick out the pieces.

MS. PACE: I will just make a comment, too. We often get this comment about outcome measures. It doesn't tell you exactly what to do. But it can be different for each hospital.

So the idea is you look at where you are not doing so well, and then you have to investigate what is the cause.

MR. DUDLEY: Yes. In my experience, in the absence of measuring the outcome first, no one investigates. So you have to start to measure it, and then you get to the understanding.

CHAIR YAWN: Well, and this is not an outcome measure that says every patient is admitted to the hospital's mortality. This is a more limited group of patients. So I mean I can understand when somebody says, okay, let's give our overall mortality rate per year for God knows what. Well, that is not very helpful to me.

But this is a very specific group of people, especially since they have pulled out the burns and the bypass, and all those other people. You now have a pretty limited group of people who are usually there. So I think that that is how we answer that one.

All right?

DR. NEFF: So, then, we are at 4, which is exactly, I think, the same issues we had before. There is still coding, abstraction that needs to be done by someone other than the person that is doing the clinical data. Then the electronic source of what they have, it is readily available. Do you know what I mean? What is currently available? So I think we are doing "M" and the "C" is how we are --

CHAIR YAWN: I don't know that this one is, though.

DR. NEFF: Okay.

CHAIR YAWN: Is there any hospital now that doesn't electronically have the fact -- well, if they don't know about the unit, whether they are in the ICU. Okay. They all have electronically that they are dead.

MS. WINKLER: But they need the risk factors.

CHAIR YAWN: Yes. Okay. Got you.

MS. WINKLER: It is the same issue.

CHAIR YAWN: Okay.

MS. WINKLER: Not specific to the

study, just electronic --

CHAIR YAWN: I am sorry, you are going to get an "M" whether you like it or not. So there.

(Laughter.)

MR. DUDLEY: As long as it is only one.

MS. WINKLER: One per study.

DR. NEFF: It is a weird "M". It is an across-the-board "M".

Nothing weird about the exclusions that requires additional data. So that is a "C".

And then, really, the same discussion we had about avoiding the high-risk patients. There is good reason for that. So I put a "C" there as well.

CHAIR YAWN: All right.

DR. NEFF: And, really, I just have feasibility, just in terms of the handout.

CHAIR YAWN: Okay. Let's see --

DR. NEFF: And that is no different for anybody else.

CHAIR YAWN: Well, the weaknesses he has, I think you have discussed them already.

DR. NEFF: Yes.

MS. WINKLER: I just wanted to verify -- scroll up just a little bit, Alexis -- in terms of the model is readily available; anybody can adopt it. The cost would really be involved in just data

abstraction, you know, personnel versus electronic, you invest or not.

DR. NEFF: Right.

MS. WINKLER: So that is also a significant feasibility issue, actually.

MR. DUDLEY: Do you care that we give away the -- so you can put electronic software -- we have software that you can put into your computer system. Then you could report the data directly to us, and we would give that away for free. And that is true for this one and the linking one.

MS. WINKLER: It certainly doesn't hurt. We care. It is a nice characteristic.

MR. DUDLEY: And we give away the data collection forms, if you don't want to send us the data.

CHAIR YAWN: It doesn't get you a "C-plus". I'm sorry.

(Laughter.)

MR. DUDLEY: But I tried, though.

MS. WINKLER: It's a good thing.

CHAIR YAWN: Yes. No, it is, and we will mention it as a strength.

MS. WINKLER: Of the feasibility?

CHAIR YAWN: Yes, that you will give them that prepared data and --

DR. NEFF: It's not restricted?

MS. WINKLER: I guess one question. You are building your data model or your risk model off of a portion of your sample from California. If this were to be nationalized,

if you will, strategies for remaking the risk model on a national sample?

MR. DUDLEY: Well, the reality is that you have to re-estimate the risk model continuously anyway.

MS. WINKLER: Right.

MR. DUDLEY: And you should do that for whatever population you want to do. So, if Barbara decides in Minnesota we want to do it, then they shouldn't accept anyone's model, a national model, a Rhode Island model, a California model. They should look at the risk data from Minnesota.

But, even when they have done it, that model, in reality, is probably only good for a day, but we only turn it over every quarter, and we recalculate it every quarter. So that there is no way -- because what is happening across the country, and it is a good thing, but it is happening faster in California, I think because of this public reporting, is that mortality rates are going down for any given level of risk.

So you want to constantly be updating and reflecting where performance is today.

CHAIR YAWN: So, again, that is a strength of very rapid reassessment of the risk model.

DR. NEFF: And kind of the importance of having support for whatever model system you are using. So you are not

just buying a box, the software, and forgetting about it for five years and using the same thing. You have to kind of --

CHAIR YAWN: He likes to do that yadda yadda.

(Laughter.)

MR. DUDLEY: Tweet, tweet, yadda yadda.

(Laughter.)

CHAIR YAWN: We don't yadda yadda; we tweet. But he yadda yaddas.

Okay.

MR. DUDLEY: Give me cookies and I'll do anything.

CHAIR YAWN: Thank you very much. We appreciate it. Feel free to stay and eat more cookies or leave, whichever works for you.

MR. DUDLEY: If I eat more cookies, I won't fit out the door.

MS. WINKLER: Well, thank you very much because I know it was a long trip for you to come and talk with us.

CHAIR YAWN: Yes.

MR. DUDLEY: Thank you.

CHAIR YAWN: We appreciate it.

Okay, we are going to move on now. Which ones do we have?

MS. FORMAN: Maybe we can do COPD.

CHAIR YAWN: All right, let's do COPD.

MS. FORMAN: Okay. That is Neff

and Millard.

CHAIR YAWN: And which number?

MS. FORMAN: I'm sorry, it is 18.

CHAIR YAWN: Eighteen.

DR. RASTOGI: So just a quick overview about it. This measure is kind of not a standard NQF type of thing that has been there in the past.

He is creating an episode-based construct here. Many of the measures that are competing, and we have 21 ECRs that we have been working on. So there are 21 measures that he was having to finally submit.

And they have a common framework. We defined the episode triggers. We defined the length or the time window for that episode, and then it is completely claims-based, you know. At this point, we haven't been able to include what we call Channel 2 or Channel 3 data, where we would have EMR information or patient-specific information.

So the only information for patient purposes is only demographics, age, and gender, and then an enrollment file, whether they enrolled as they had planned, and for what time period.

Then these measures that we are developing, it is not directly meant for provider performance measurement, but it is more for provider self-improvement. So they can look at their own results over time and see how they are working towards improvement.

So we don't intend to have it used as a provider profiling mechanism. However, we do intend to use it for public reporting. We would give population-based information. So, in Minnesota, the percentage of patients who had potentially avoidable complications in a COPD setting are "X", and in California it is "Y". So those kinds of information and detail won't be available.

So the potentially-avoidable complications, which is the crux of the matter here, are what we are measuring, and there are several of them. The most important in these chronic conditions is hospitalizations. So all hospitalizations that happen during this time window are considered potentially avoidable if they are related to the initial index condition. So when we say related, we have to try to make sure that that hospitalization is not for an unrelated condition, for example, hip replacement in COPD.

But the main idea behind it is, can we keep the patients out of the ER? Can we keep them out of the hospitalization, out of the hospital?

The idea is not to say no hospitalizations have to happen, but the admission, the percentage of COPD patients that do get admitted and the percentage of COPD patients that do end up in the ER, and then try to get a trend over time, can provide

us work towards decreasing that rate.

We also make available more actionable data. So, besides hospitalizations, if they have other conditions which have been flagged or labeled as potentially avoidable complications, things which may be across the board because this is a patient-centered approach, so if they have a urinary tract infection, they have defib thrombosis, they have pneumonias, they have other things that are going on, you know, what is the incidence of that particular PAC, as we call it, and over time can that be reduced?

So the incentive in the cost-sharing model, it is a shared savings model that ultimately is being proposed as the PROMETHEUS payment architecture. That is kind of the full picture of it, but the limited piece is the patient outcome measures for using the potential avoidance of complications.

CHAIR YAWN: Good. Thank you.

All right. And are you the first reviewer?

DR. MILLARD: I am, and I have to confess that I have spent more time on this, I mean on this one thing, than everything else, review of the other. I do not understand the model. I do not understand the use of the different -- the inclusion of what seemed to be singularly unrelated diagnoses.

So, as a clinician just trying to

make sense, and not as a statistician, I understand kind of where you are going. And certainly, when you said this is part of 21 different -- so that is the reason why all these other illnesses are that, and I have a problem with, therefore, saying, we are saying the percent -- what we are really saying is not the proportion of COPD exacerbations that have potentially avoidable complications from COPD, but we are saying the patients with COPD who have potentially avoidable complications in general unrelated to COPD. Because how is urinary tract infection related to COPD?

DR. RASTOGI: So you are absolutely right that some of it is confusing, but we are thinking in terms of a patient-centered approach. So if you think in terms of this medical home model, or whatever, the physician or the treating provider is supposed to look at the patient as a whole. If, during his watch, while he is caring for COPD, the patient develops a urinary tract infection, then that could have been something that could have been managed proactively. That is all we are saying, right?

So many of these PACs, I agree, are across the board, across all episodes that we are doing, and they are pretty much common. There are episode-specific PACs, but then there are some which are common across the board.

CHAIR YAWN: Could you help me by

telling me, when you give us data -- say Minnesota is far too large a group -- we will say for southeast Minnesota, are you going to tell me what the PACs are? Because it really isn't very helpful to me to say 87 percent of all my people with COPD had one of these. So what?

DR. RASTOGI: And that is exactly what I said. We make the data actionable. In one of the files that we have provided to you, you can see if the hospitalizations are there, you can see what were the causes of hospitalizations. So we say, what is the percentage of patients who had hospitalizations and what were the chief drivers of hospitalization, the frequency and the cost associated with it.

CHAIR YAWN: Okay. So, if I have 70 patients submitted with COPD or who have COPD and were admitted, you are saying that you believe it is actionable if you tell me 30 percent of those were due to a hip fracture, 10 percent were due to thrombophlebitis, and 50 percent were due to pneumonia?

DR. RASTOGI: Right, and then we have drill-down capabilities. So we have produced like the standardized SAS programs, which we are making available to all the health plans. So that right now 11 different databases it has been tried on. They use the same standardized SAS package.

So the data that comes out is

exactly the same. So we can compare it from one client to the other or one employer base to the other.

Once you look at the percentage -- so, say, you say, oh, hip fractures are very common in my population. You could drill down and see exactly how many patients and what patients, the names of the patients who had this problem. Then, if you wanted to do a retrospective reconciliation, you could go down and do a chart review, or whatever.

So that is how it has been implemented for pilot sites across the country. So the physicians who are -- and from Minnesota, Medicare is participating, Health Partners is participating, and there are other groups that are also looking into it.

So, when they run the data through this, they look at exactly what patients, they define the COPD patients, what were the complications that we are calling as PACs, and on the professional side as well as on the state side, and what are the costs associated with that, and then trying to see why is a certain potentially avoidable complication more prevalent in a given population.

CHAIR YAWN: So it really only works with a health plan?

DR. RASTOGI: Well, we have the employee coalition --

CHAIR YAWN: But that doesn't help

me much, as an individual physician, because, yes, this patient is employed by, let's say, IBM, but I only take care of their left leg because they have six other specialists who take care of all the other pieces. How does this work for me to get actionable information out of this?

DR. RASTOGI: Yes, and it is a very good point you raise because we have been so ingrained in the fee-for-services these days, everything is encounter-based, and a patient body limb- or part-based. You know, we are not really looking at the whole patient.

So moving to an episode-based approach is kind of changing the paradigm a little bit, trying to get the providers to focus on everything that is going on with the patient.

But how most of the people are using it, like, for example, Medicare, and the population, the whole medical population, through the database, then they passed it out into Fairview Clinic, this kind of thing.

CHAIR YAWN: Oh, right. They have people that are responsible for that patient's care?

DR. RASTOGI: Yes.

CHAIR YAWN: In fee-for-service, I don't have that at all.

DR. RASTOGI: Yes.

CHAIR YAWN: So, again, an HMO, maybe managed care, Medicare and managed care

Medicaid, what percentage of the population does that represent in the United States?

DR. RASTOGI: Right. And when you go down from the --

CHAIR YAWN: But do you know? I am asking you a question.

DR. RASTOGI: Okay.

CHAIR YAWN: Do you know what percentage of the U.S. population is covered by managed care, Medicaid managed care, or Medicare managed care?

DR. RASTOGI: Well, I can look up the numbers.

CHAIR YAWN: I don't know, but that seems who this applies to, but it would be very hard to apply it to other groups. I am just trying to get a sense of that.

DR. RASTOGI: Why would it be hard to apply to other groups?

CHAIR YAWN: Well, again, I do fee-for-service.

DR. RASTOGI: Yes, this is also fee-for-service. This is completely in a commercial database. It is not in the HMO population.

CHAIR YAWN: Oh, the ones you mentioned are all HMO populations. Medicare is; Health Partners is.

DR. RASTOGI: No, I am giving an example of who is using it. Now, in Rockford, Illinois, that is the Employee Coalition on Health, they have a very different system,

right? And they have three big hospital systems that are participating in it, right? Partners in Massachusetts, they are completely different, right? Crozer-Keystone in Pennsylvania, it is different. Geisinger has taken this thing, documented it, is running through the CMS system, the data piece.

So people are running these programs onto different databases. The way the payment system will work is very different than with outcome measures that we are presenting today, right?

The payment reform that is coming through -- we are not talking about accountable care entities or something. We are talking about a system we can work out how the payment reform would be and how the providers would be made responsible in a shared-savings model.

CHAIR YAWN: Yes, I am not worried about the payment so much as I am worried about the actionable items.

DR. RASTOGI: Sure.

DR. NEFF: Can I ask you one question, I think semi-related, but just to get a feel for this?

DR. RASTOGI: Yes.

DR. NEFF: So the measure is proportion of COPD patients that had potentially avoidable complications, and there is a whole list of potentially avoidable complications. Is the assumption that those

complications are somehow related to their care of their COPD? Or is it totally that this just happens to be a COPD population, and you are looking at all these bad things that can happen?

DR. RASTOGI: Okay.

DR. NEFF: I mean, is it anything specific about it being COPD?

DR. RASTOGI: Yes. So the sequence by which it goes is you start the trigger. So the patient has to have a COPD trigger. Then you look for one year's worth of claims, starting from the trigger date. During that time window, all the things that come for that patient are looked in, and there's a filter logic that works.

So, if a claim has a COPD-related diagnosis code on it, then the claim gets filtered then as relevant to COPD care. If it doesn't have one of those filter codes, then it is considered as irrelevant to COPD care and it is taken out.

Now, of all the claims that do get considered as relevant, then they are sorted out into whether they are typical or PACs, based on these definitions. So, if they have a single PAC code, then that claim gets put into the PAC bucket or the others get put into the typical --

DR. O'CONNOR: There are more than 50 of these associated PACs?

DR. RASTOGI: Right, there could

be. I haven't counted, but it varies from --

DR. O'CONNOR: Yes, well, there are. If you think they are struggling on the adult side with the COPD, can you imagine me struggling on the pediatric side, looking at these 50-plus --

CHAIR YAWN: For asthma, yes.

DR. O'CONNOR: -- for children with asthma.

CHAIR YAWN: Yes. Well, but they seem to be the same for children for asthma and adults with COPD.

DR. NEFF: They are just like any bad thing that could happen to you.

CHAIR YAWN: And if I had a patient with -- oh, I don't know -- hypertension, would they be the same?

DR. RASTOGI: Similar, you know, and like hypertensive emergency would be specific for hypertension, but, yes, urinary tract infection would be common. Pneumonia would also be there for hypertension.

CHAIR YAWN: And diabetic emergency --

DR. RASTOGI: Yes.

CHAIR YAWN: -- with hyper- or hypoglycemia?

DR. RASTOGI: Yes.

CHAIR YAWN: Okay. I think that is where we are having our problem because --

DR. RASTOGI: Because we don't think of patient-centered these days, right?

CHAIR YAWN: Well, some of us do. I mean I am a family physician. Yes, I do, but my specialty colleagues here don't.

DR. RASTOGI: I am a specialist, too.

CHAIR YAWN: I understand. Yes, you take care of left or right coronary arteries. I understand that.

DR. NEFF: And I guess it doesn't seem that this is necessarily unique to COPD patients. Maybe that is what I am struggling with. This is just is you could almost stick anything in the slot and then run this list of complications and then print out a report?

DR. RASTOGI: Yes. There are some which are specific, but you are right, most of them are not, right. Most of them, you are right, are anything bad that can happen that could be avoidable, but when the actual data comes out and you could see that -- I don't know if you can bring up that Excel worksheet that we had attached, but it is called "Risk Adjustment".

And you could look at the -- the outputs are different, you know. So, when you look at the COPD population, the top drivers are pneumonia, lung complications, et cetera. You do see some DBTs and all that, but they are low down in frequency, right?

When you look at an asthma population and adjust for the pediatric, then the top drivers are different.

So each provider could look at their specification, say, "What for my patients?" Now, you're right, we don't have a provider database which goes across health plans at the time being. We are trying to reach out to others who are advocate for provider specificator, and then see if we could run it on that.

But, you know, we are using whatever these administrative claims databases is aggregated, and over there in these large databases we can see what is actionable, based on what percentage of patients have a certain complication, and they drill down to the provider.

CHAIR YAWN: So are you trying to suggest that these could be very useful because you spend all the time and energy getting all these PACs, and then you can put pretty much any patient population at the front end?

Are you trying to make efficiency? Is that why you are doing it this way?

DR. RASTOGI: The reason is, if you think in terms of a particular patient, right, and say you are a primary care provider, you have a patient who has COPD and they end up in the hospital with acute myocardial infarction, right? Would you completely wash off your hands and say, "That has nothing to do with me."? No.

CHAIR YAWN: Most people with COPD

die of heart disease.

DR. RASTOGI: Yes.

CHAIR YAWN: But I don't take care of the patient in the hospital anyway.

DR. RASTOGI: Exact.

CHAIR YAWN: A hospitalist does.

DR. MILLARD: I think that sort of the systems issue or the primary issue that I think I went into this review of this with is trying to figure out specific avoidable complications directly related to COPD. Maybe that is the subspecialist trap, recognizing that, as my fellowship director said, yes, I did an internal medicine residency, too. So I mean I need to look at the whole patient.

But I went with the assumption that these -- and the problem that I had logistically with this process was that I kept trying to figure out what does urinary tract infection have to do with COPD, and that isn't well-described upfront in the model. You have described it better. I understand it better.

I think this is really more chronic disease, potentially avoidable complications, because I suspect you could put in CHF; you could put in chronic renal failure; you could put in a lot of things, diabetes, and you would have the same PACs. Then you could go forward with this very nicely.

But I think if you use the single -- there is just an issue of perhaps syntax when you say this is COPD, patients with PACs.

It makes you feel like you are looking through the lens of a COPD patient, as opposed to patient. They may have COPD, but they are really what they are. And that changes the entire evaluation process.

DR. RASTOGI: Yes, and to some extent, it is specific to COPD patients, based on how I described about the claims, you know. So, then, we look at, say the same patient also has congestive heart failure, right? So they would be, for the same patient, there may be another set of claims that get put, right? So, for the first patient, for that patient, when you are looking from COPD lens, there are a handful of claims that get put in. But when we look at it from the CHF point of view, there may be another handful of claims that get put in.

Now I agree with you that some of it is based on coding practices. If the coding is not very good or complete, you may pull a different set of claims versus another. But to the extent that we believe that is all that is available to us, and in administrative claims data, then we are evaluating a particular patient based on the condition that is the trigger.

CHAIR YAWN: I think that one of the things that it is a very different model to think about patient-centered for an individual patient and then a quality-of-care measure for something that would go across

specialty and primary care and who is responsible for what, and how do we figure out what is actionable? I think that is part of our issue.

DR. RASTOGI: Yes.

CHAIR YAWN: So let's try to see if we can go on and see what happens as we go on through some of the very specific --

MS. WINKLER: Just one question or clarification. The level of analysis for this measure would be plan, system, large group, and it would not -- so you are talking about those large entities that would compile all of them?

CHAIR YAWN: But it would have to be a very large system.

DR. O'CONNOR: It would have to also, if you fast-forward and say, from this data, you get this information, and you had variations across the country, then we are left with, well, there are 53 conditions in there. And even though each one might be actionable, you say to yourself, what's the next step? I mean, how does this lead to quality improvement?

MS. WINKLER: Do you have any experience with what the people who are using it do?

DR. RASTOGI: Right. And you know, the database has been run, like we said, and this here is, 2010 is our implementation year, where people are starting to work toward

practice re-engineering. So they identify two or three big drivers of cost, and they try to focus into that and say, how can we reduce it?

Geisinger Health System has adopted a similar improvement care model where they have the elements of practice re-engineering, right? So, of course, they don't have the chronic care model. They just had the coronary bypass, cardiac surgery, so more the procedural ones.

So they have taken all chronic care, you know, program, a SAS program, normally identify the issues and now they are working on practice re-engineering on specific ways of improving, you know, to decrease the PAC rates.

CHAIR YAWN: So this model hasn't really been tested either -- I mean you are in the process of?

DR. RASTOGI: Yes, it has been tested to the extent that we can measure the PAC rates or the proportion of patients who have PACs. Now maybe their PAC rates will decrease over time. That part has not been tested.

CHAIR YAWN: Okay. And if I have patients from five insurance companies and you have collected from an insurance company, how are you going to aggregate my patients?

DR. RASTOGI: Yes. So it is not provider-specific right now. Okay?

CHAIR YAWN: Okay.

DR. RASTOGI: To the extent -- if all five health plans participate in this project, then we would be able to get information from five different. So, for example, Crozer-Keystone is a provider system, and they have Aetna and IBC data. So both of those plans ran this model. Okay? And they put patients specific to Crozer-Keystone, and then have identified -- and they have very similar overlapping results.

They were able to say, okay, these are the cost drivers. Now our aim is to decrease these PACs.

CHAIR YAWN: Okay.

DR. MILLARD: Does NQF want, were you looking for a COPD-specific outcome measurement or were you looking for a general outcome? Were you looking for PACs or were you looking for COPD-specific PACs?

MS. WINKLER: All of the above are possible. We didn't focus it specifically.

DR. MILLARD: Okay.

MS. WINKLER: I mean the call for measures looked for both cross-cutting, condition-specific outcome measures. So slice it any way you want to.

CHAIR YAWN: I think let's go ahead.

DR. O'CONNOR: To call this a COPD measure is a misnomer then.

CHAIR YAWN: Well, these are measures of PACs in patients with COPD, as it

is specified right here.

DR. O'CONNOR: And the PACs are not related to their having the COPD.

DR. NEFF: Some are. Some get this link, right?

DR. RASTOGI: That is exactly right.

DR. MILLARD: But urinary tract infection I think is a PAC.

DR. NEFF: But it wouldn't necessarily be associated with COPD.

DR. O'CONNOR: No, she said it was.

CHAIR YAWN: But we would still get that information. What he is saying is it probably wouldn't show up very often. So it wouldn't be one of the drivers and probably wouldn't be what you would choose.

DR. RASTOGI: But if the claim had both COPD and UTI in the diagnosis codes, then it will get pulled in, right? If they didn't have any COPD diagnosis concerns, so if they just put an office visit to a urologist and he just wrote "for urinary tract infection", didn't put the COPD, then it would fall out.

DR. NEFF: So it is not trying to link the complication to somehow being related in some sort of physiologic way. It is just whether they are both showing up on the note or on the claim at the same time?

CHAIR YAWN: And that is why she is saying that the patient --

DR. MILLARD: But the patient has

COPD.

CHAIR YAWN: Yes.

DR. RASTOGI: The patient has COPD. That is sort of a trigger.

DR. MILLARD: Okay. Yes. So my patient with COPD that I see in my general medical practice, unbeknownst to me ends up in an ER with urinary incontinence, with urinary retention. There's no code for COPD on the ER visit for urinary retention, but it is still tracks as a PAC to that patient?

DR. RASTOGI: Yes. If there is no procedure done for him, say it was -- then it gets excluded, okay? Right.

DR. MILLARD: Okay, but the point is that, if they go to the urologist, and the urologist doesn't have to code the COPD, but I've already coded that patient as COPD earlier, it will still track to that diagnosis, correct?

DR. RASTOGI: No, no. On that particular visit. So say they went to a urologist.

DR. MILLARD: What is the difference between a PAC over a year's period of time -- a PAC over a year's period of time will take all the visits in aggregate, won't it?

DR. RASTOGI: It looks at it claim by claim, and it looks at one year's worth of claims. And for every claim, it makes a determination if it has a COPD-related printer

code or not.

So, if it was completely unrelated to COPD, and the doctor doesn't code for COPD, then that claim gets thrown out.

DR. MILLARD: Even if he has been seen in the year previously for --

DR. RASTOGI: If the COPD is a chronic condition --

MS. PACE: So say a patient went to the pulmonologist, and that pulmonologist identified that they had a urinary tract infection.

DR. RASTOGI: Yes.

MS. PACE: So, on that claim, it would have, and so it would show up. But if the patient, through another vehicle, went directly to a --

DR. RASTOGI: A urologist.

MS. PACE: -- urologist, and they didn't put "COPD" on, then the UTI wouldn't show up?

DR. RASTOGI: Right, and that is the filter process, right?

CHAIR YAWN: But that is actually a huge problem because people who have fragmented care are going to look like they have many fewer PACs than my patient-centered care, where, yes, every time the patient walks through the door, I am probably going to put down they have COPD, if it is reasonably severe, and everything else they have.

But my colleague down the street

who is a cardiologist doesn't put COPD and puts all of those other things, and so they don't ever turn out to be PACs for that patient.

DR. RASTOGI: Right, and it is a good point you raise because that is exactly what I was saying. It all depends on the coding practices. You know, to the extent that the coding --

CHAIR YAWN: Well, it depends on where the patient is taken care of, too.

DR. RASTOGI: Yes.

CHAIR YAWN: I mean it is much more likely in a patient who has fragmented care that they are going to look better. This is one of the things that we, as generalists, worry a lot about. If the patient has totally fragmented care, they are going to come out looking better than a patient whose care is all put together, and somebody knows they have a urinary tract infection, they have COPD, they have congestive heart failure.

DR. RASTOGI: Yes.

DR. O'CONNOR: Yes, and PCP.

CHAIR YAWN: Yes, I know, because I am taking care of a whole patient, and you guys are giving him medicines that conflict.

DR. RASTOGI: Yes. And we look at it more from the provider attribution point of view. So, if a urologist is taking care of the urinary tract infection, should he be attributed COPD or not, right? So is the

urologist responsible for COPD?

CHAIR YAWN: Yes.

DR. RASTOGI: And to the extent he's coding for COPD, he is taking care of COPD, he is recorded. Then he will be considered responsible, right?

CHAIR YAWN: Well, but he's as responsible as I am for the fact that, when he gives the patient the drug that causes whatever, and has a side effect that is a pulmonary side effect, but he doesn't know it, so he doesn't code COPD, then he is not responsible. But I am because I have to take care of the mess he made or she made.

DR. RASTOGI: Right. You know, we can win both ways because we present it to our design team members both ways. You know, the initial model was we were including all the claims that were coming through for one year's worth, and if they didn't have an exclusion code, like if they didn't have a major procedure or something like that, and they didn't get excluded, then all of the claims were included.

But then the question from the design team and some of the pulmonologists on the design team was that this is being taken care of by a urologist, and it is not really a COPD-related problem.

So the printer logic kind of decreases the -- you know, to the extent a coding happens, and they code for everything,

whether they are taking care of it or not, so that during that office visit you did nothing for COPD and you code for COPD, it is kind of, you know --

CHAIR YAWN: But I just think it sounds like your design team -- I don't know if you had any primary care people on it, but I would bet that would be the distinction, and one group would say let's do it this way, and the other group is saying let's do it this way. Because you sound like you have a measure that works for one group, but it doesn't work for the other.

DR. O'CONNOR: The other thing -- pardon me -- is the choice of the term "avoidable complication". For example, some of these things in here, skin and lung care, they may not be avoidable complications. Use of a splint, how does that qualify as an avoidable complication?

CHAIR YAWN: Well, if you hadn't fallen down, you wouldn't have needed the splint.

DR. O'CONNOR: Well, that is my point. I mean some of these aren't complications. These are the way we treat patients. I mean I am struggling with the use of the term "a potentially avoidable complication".

DR. RASTOGI: Right, and some of these are aggregates of codes. So the names on them may just be a representation of the

coding that is behind it, and you could drill down and see, you know, the codes that aggregate up to that.

CHAIR YAWN: Did you take into account -- I mean reviewers, people who have worked on ambulatory-sensitive conditions -- did you look at those before you chose your PACs?

DR. RASTOGI: That is exactly how it started.

CHAIR YAWN: And is a fracture an ambulatory-sensitive condition?

DR. RASTOGI: The PACs that we have used are the ones which AHRQ has defined, you know, for these chronic conditions. All right? So those are the ones that --

CHAIR YAWN: As avoidable?

DR. RASTOGI: As avoidable. That is this hospitalization for ESEs becomes avoidable complications.

CHAIR YAWN: Okay. All right.

DR. RASTOGI: Then the HACs, you know, the hospital-acquired conditions, are also part of --

CHAIR YAWN: Oh, yes. Now hospital-acquired, those are certainly potentially avoidable if you never get in a hospital in the first place. So I don't have a problem with that.

But let's, please, just go ahead and try to go through and see what happens. Okay?

DR. RASTOGI: Okay.

DR. MILLARD: Margaret, if you could help me on this?

DR. NEFF: Oh, yes, I've got your back.

(Laughter.)

DR. MILLARD: Oh, okay. Okay.

DR. NEFF: We're all together.

CHAIR YAWN: You know, I did the two asthma which looked exactly like this, except it says the word "asthma", and they were different age groups. So I think all four of us have been through this, and we can sort of help each other.

Okay. So let's go to 1a.

DR. O'CONNOR: This could be the killer right here.

DR. MILLARD: Importance of measure to report.

DR. NEFF: So it is aimed at prevention. I mean that is clearly -- which is generally always a good thing. I think, then, there is the issue of how it translates into use, but, clearly, the intent is a preventative one.

CHAIR YAWN: Okay. So 1c? Are you comfortable --

DR. O'CONNOR: Basically, it says here, the first question is, "Extent to which the measure focuses...is important to make significant gains in the healthcare quality and improving health outcomes."

DR. NEFF: But then we go down to the subcriteria.

CHAIR YAWN: So demonstrated high-impact aspect of healthcare. Summary of evidence citation. This is 1a.

DR. NEFF: But isn't there also, if you meet one of the MPM key goals, you sort of already are there, even before you get into all the other --

MS. PACE: For 1a.

DR. NEFF: For 1a.

MS. PACE: Only for 1a.

CHAIR YAWN: So I mean I think it is sort of the given "C" before you even start.

DR. NEFF: Just by virtue of the things it hits, even if you don't get into the concept of whether it is having an high impact.

CHAIR YAWN: Yes. I don't think this is one where we should spend a lot of discussion.

DR. NEFF: Right.

CHAIR YAWN: I think we all have better discriminating discussion later.

Okay, 1b.

DR. MILLARD: 1b, opportunity for improvement. Demonstrating performance gaps across providers. Well, by definition, therefore, this is would be an "N", therefore, because you have said this is not provider-specific.

DR. RASTOGI: Right, it is population-based.

DR. MILLARD: This is population-specific.

CHAIR YAWN: Right. So, when it is population-specific, it makes it tough for me to --

MS. PACE: If you look at the actual statement in the criteria, it is across providers and/or population groups, which relates to -- we may have made that too small in terms of providers. I mean we do allow for measures as a health plan level or population level. So that is a good point out to us that we may have to fix that language.

DR. MILLARD: Okay. So long as we can, because I mean it is --

MS. PACE: Exactly. I understand what you are saying. But I think, because it is at the health system level --

CHAIR YAWN: But it is at an insurer level frequently, and that is a huge problem for a specific group trying to improve. Because if I have seven --

MS. PACE: Right.

CHAIR YAWN: -- insurers and only one agrees to give me the data, maybe they have very few of my patients, and it may not work for me. So there is that problem.

MS. PACE: Right.

CHAIR YAWN: This needs to be a very widespread --

MS. PACE: Right. Okay.

DR. MILLARD: Now, in terms of citations, the Hogan article just merely says COPD patients who get hospitalized have lots of chronic illnesses.

CHAIR YAWN: Right.

DR. MILLARD: How is that sort of relevant to this issue of performance gap? I mean I am not sure -- I wasn't able to access the sustainable medical home report. I didn't have time to get to that. But the issue is, I mean, is there literature that suggests -- I mean I guess there is plenty of literature when you just look at the PAC, when you look at the general trend. Some of them, 90 percent of some base populations and 64 percent of other patient populations have it. So, obviously, there is performance gap measurings.

So do you want to say "C" on this and move on? Because this is --

CHAIR YAWN: Yes, go ahead.

DR. MILLARD: Okay. Outcome or evidence support measure focus. Here we have, I think, a real issue of, when there is a summary of evidence and you use the ability to reduce hospitalizations and ER visits by 31 to 50 percent by aggressive pharmacologic management of COPD, that is pharma data from very selected populations that are incredibly -- I think most of that has to do with pharma-related phase 3 trials and not necessarily

generalizable to wild-type --

CHAIR YAWN: Wild-type patients.

(Laughter.)

CHAIR YAWN: Also, it depends on what you choose as your drivers, whether pharmacological therapy is going to make any difference or not.

DR. MILLARD: See, the evidence cited here is essentially related to COPD, not disease-specific, not this generalized.

DR. O'CONNOR: Relevant to the target population.

DR. NEFF: Right. So then that gets at, what are we actually trying --

DR. MILLARD: Yes.

DR. NEFF: -- so are we trying to look here -- if you are trying to support this, we have evidence to support the measure. Are we trying to support interventions for COPD that prevent complications?

DR. RASTOGI: Yes, and what we saw in the example that she was showing, that the chief drivers or the main PACs are the COPD-related ones. They are the pneumonia. They are the respiratory insufficiency. The last two PACs are the ones that you need to look at.

CHAIR YAWN: So, then, why do we bother to do all the others? It just seems like a tremendous amount of data collection if we already know what the main drivers are.

DR. RASTOGI: Yes, the detail is

already there. There's nothing new that you have to do for collection because it is all already available. It is just that the SAS program has churned through it.

Now, if you exclude the other PACs and just keep them -- we didn't know what would show up when we initially started up with this whole analysis. This is after the fact, if you look at it now. If it weren't, you know, you could go and refine them.

But most people found this information very helpful. You know, we ran it through quite a few medical directors, the different health plans, the actual providers in the community who had agreed to take payment through the PROMETHEUS system. They didn't have a problem, but delete the other PACs, you know.

DR. NEFF: So do we think it is odd that that PAC 30, the acute exacerbation of COPD and asthma is so low?

DR. RASTOGI: We could look at the exact coding behind it, and I could tell you exactly what the --

DR. NEFF: Okay. Well, why are there so few --

DR. O'CONNOR: It doesn't pass the smell test.

CHAIR YAWN: The what?

DR. O'CONNOR: It doesn't pass the smell test.

CHAIR YAWN: No, it doesn't.

DR. NEFF: So few exacerbations and fewer than there are phlebitis.

DR. O'CONNOR: Episodes of septicemia.

DR. NEFF: Yes.

CHAIR YAWN: In the COPD population.

DR. NEFF: Right.

DR. RASTOGI: But the acute exacerbation of COPD announcement was the full-length of people in 541.2. So that is the status, asthmatic, exercise-induced bronchial spasm, acute bronchitis and bronchiolitis, COPD with acute exacerbation, 491.2, 541.2.

So all that is in the spreadsheet which I just opened up. But there are about six or seven types of codes that can be put together in that.

CHAIR YAWN: Most people would code a hospitalization which is specifically for COPD as an exacerbation by definition or an ED visit that just has a COPD code as the first code, as by definition an exacerbation.

So it sounds like the way that it was specified may be a problem. Because if I look at this and I take the top two or three, okay, I am going to need to deal with mental health issues as a big driver for quality improvement in my COPD patients.

DR. RASTOGI: Yes, and you are right, this is the principal diagnosis that we

have reported on. So, if patients with COPD have been admitted and they have been reported as principal diagnosis of this, then that is what it is. You know, I can't change the data.

CHAIR YAWN: No, I know, but the problem is, when you use just this as the way to identify them, it is the way that CMS pays and DRGs, it really complicates the way you do it. And the reason that most of us do not use this when we are doing any kind of research -- I mean I would never use the first diagnosis as deciding why somebody was in the hospital.

DR. RASTOGI: Okay. And, yes, it is a good point you raise because we don't believe in the DRGs, either, because DRGs give bad incentives. You know, you get paid more when you have a complication. So we have not done any grouping by DRGs.

CHAIR YAWN: No, I understand that, but the point is that in every hospital that are a whole room full of coders. Now they take what I write about what the patient came in for and they rearrange it to make it with the highest payment. So it is not why the patient really came in. It is what gives us the highest payment.

When you translate that, then, to me doing a quality improvement program based on what the coder decided would get paid the most, it doesn't translate.

DR. RASTOGI: Exactly right. I

agree with you that there's a lot of gaps between the administrative detail and clinical chart review. And this is not clinical at all. It is all based on administrative --

CHAIR YAWN: But can I justify doing a quality measure that gives me the wrong answer --

DR. RASTOGI: No.

CHAIR YAWN: -- to what I should improve?

DR. RASTOGI: This is the starting point, and make everything -- Crozer-Keystone is working just on knee replacement. Their physicians, they have hired a consulting physician who is just looking into: what are the PACs? What are the drivers? They go back into the charts and they say, was it really hemorrhage when it is coded as hemorrhage. And it is very interesting the results that they find because sometimes it is not even hemorrhage.

DR. O'CONNOR: But in that situation, it is a potentially avoidable complication of hip surgery.

DR. RASTOGI: Yes. And then when they go back to clinical charts, it was that there was no hemorrhage. It was just coded as hemorrhage to get more money from CMS.

CHAIR YAWN: But, I mean, that is something we all know about the difference between coding and what is in the medical record. I mean that was one of the things I

was going to tell you about the asthma. You are going to way over get asthma if you only use one diagnosis. You are going to have 30 percent of the people who don't even have asthma, in children anyway.

But what I am saying is the whole basis for this, you said, unless I misunderstood, what I am going to do is I am going to find this out, and I am going to look at this and say, "Oh, look at that," you know, "19 percent of all the hospitalizations were due to mental health." So, if I want improve COPD care, I had better go work on mental health issues.

DR. RASTOGI: But what you would do is you would look at those 19 percent of your patients, the 907, and you would say work with the mental health behind it, right? Because it is grouped up, it is bunched up at several codes.

Like I told you, this data is drillable. So you can drill down to the patient level, and you can see, what are the codes, what are the drivers, why was this patient admitted, and what was the principal diagnosis?

Now if you said that the principal diagnosis is not what I want to go with, and some other diagnosis, then you can pull their charts. That would be more manual extraction.

Now I agree that, you know, we, as physicians and clinicians, have a big problem

with administrative data being used for quality improvement purposes. What Francois, with his Brilliance to Excellence efforts, he is lining up, as you may know, for performance and pay-for-performance programs. He has got the EMR staff coming in. So he has got so many health plans lined up. So he is getting the Channel 2/Channel 3 data where we can get the EMR, we can get the patient-specific stuff.

But, at this point, we are --

CHAIR YAWN: But that isn't any of this. We have to talk about this the way it is.

DR. RASTOGI: Right. So, at this point, we are not there. We just have administrative data.

CHAIR YAWN: Okay.

DR. RASTOGI: We just have claims, and if claims information is not good enough for quality improvement, then, yes, I understand.

CHAIR YAWN: This way, it might not be.

MS. PACE: I just also want to make a clarification that, you know, there is a distinction between what the measure is that is being considered for endorsement and the data analysis that an individual might be able to have access to to drill down.

So what is being presented is that overall percent of COPD patients that have one

of these complications, and that is what you are presenting as being suitable for public reporting, whether it is at a health plan or system level. And this other information would be available for people for quality improvement, but this would not show up anywhere based on what we are reviewing.

So I think we just need to keep that in mind. I mean all this detail is good in terms of how it was developed and then how a provider might use it. But we also have to keep in mind that the ultimate score that would be reported would just be that general percentage of patients that --

CHAIR YAWN: Right. So what we have, let's go ahead and see. So the outcome or evidence to support the measure focus, where are we with that?

I personally don't think there is yet any evidence. I think you are testing it right now. I think there is no evidence to support using this measure yet.

DR. RASTOGI: Right. That's right.

CHAIR YAWN: And in three years, there may be, but right now I am going to say that this is either minimal or none.

DR. RASTOGI: Okay. At this point, we are just identifying the PACs.

CHAIR YAWN: Right.

DR. RASTOGI: And we think people will become aware that these things are happening. These are the reasons why your

patients are getting to the hospital or they are getting, you know --

CHAIR YAWN: Well, all you are doing, it is right now a percent. So, right now, what you are giving us, is there evidence support that telling me 87 percent of my patients with COPD have one of these 50 events in a year, is that supportable? Is that useful? Is that whatever? Can we say -- what do we say about it here?

MS. PACE: Well, I think the evidence is related to -- well, first of all, is it an outcome? Of course, these are outcomes. But, then, is there evidence that there are care processes that affect these outcomes?

So your issue about whether it is useful or whether the measure is constructed properly, I think are good questions. I am not sure --

DR. O'CONNOR: The data doesn't exist yet.

MS. PACE: That is what I am just saying.

CHAIR YAWN: Is there data to say, and I am saying I think there is minimal to no data.

DR. NEFF: As it is right here.

CHAIR YAWN: Yes, as it is.

DR. NEFF: It is describing a lot of other processes in place that could improve upon this, but, as is, it is a little --

DR. RASTOGI: Yes, and the literature is showing, you know, some studies, like they pointed out, Dr. Millard pointed out that some studies are showing, just related to pharmacy aggressive management, there are some studies which are talking about hospitalization and careful management, how you can prevent hospitalizations related to COPD.

DR. NEFF: Right.

CHAIR YAWN: But, again, those are all drilling down. Those are not studies that say, when I know 87 percent of my patients have one or more of those PACs, then I know what in the world to do with that 87 percent, except go to your next step. We have no evidence that anybody knows what to do with that 87 percent at the moment.

DR. RASTOGI: That's right. And you know, talking to various people who are in this quality improvement world, they say it starts with transparency, right? Once you know what is going on and what are the problems that are happening to these patients, then they can act on it.

CHAIR YAWN: Oh, yes, you know, we are not arguing about that. We are just saying the measure, as constructed, what would you like to give it? We need to move on to the next step.

DR. MILLARD: "M".

CHAIR YAWN: "M". All right.

DR. MILLARD: Then the strengths, we can say it includes a panoply of complications that are related to patients with chronic disease, including COPD. And in the weaknesses, we can say there is nothing unique, specific, or directed that would necessarily give us insight into improving the care of COPD patients.

CHAIR YAWN: As constructed without drilling down.

DR. MILLARD: Right.

CHAIR YAWN: Okay. All right. Let's go to the next one, then.

DR. MILLARD: Two.

CHAIR YAWN: Two, measure specifications.

DR. MILLARD: Numerator -- and this has to do with who is included and who is not, and the worksheet -- and it is anybody who has one of those targeted trigger claims that is a PAC-related --

DR. NEFF: Linked to COPD.

DR. MILLARD: Yes, links to COPD.

CHAIR YAWN: Okay. So that is the numerator. Is there an exclusion? Can you go down, please? In one of the others, there was an exclusion for anyone who had that as their first episode, had one of these PACs as their first episode that year they were thrown out.

DR. RASTOGI: Yes, there was some exclusions from the trigger, and it is not that they are thrown out, but then for the

same patient, we go forward until we find a non-acute trigger claim. So the same patient, if you start off in a hospital or you start off with an acute exacerbation, then the provider, so to say, has already inherited a train wreck.

So we wait for the most stable claim. That is when we trigger the COPD, and then we go forward one year.

CHAIR YAWN: Yes, I guess that is okay. I mean people who are train wrecks tend to have more train wrecks. I mean that is just, you know -- and the fact that you have thrown everything and the kitchen sink in there, I guess I can understand. But people who have one exacerbation of COPD are the ones at highest risk for having the next exacerbation.

But this measure is that they have one or more during the year. So that is why you threw them out.

DR. RASTOGI: We haven't excluded the patient --

CHAIR YAWN: Well, but for that episode, you excluded that episode. If that was the first one, you didn't count it. You looked for the next episode.

DR. NEFF: You waited for the next one.

DR. MILLARD: You went 30 days --

DR. RASTOGI: The next encounter, right?

CHAIR YAWN: Yes, the next encounter.

DR. RASTOGI: Okay.

CHAIR YAWN: So I just think we have to make people clear that the reason you are doing that is because you are only counting one per year. If you were counting the number per year, it would make more sense to not exclude those people. But because you are just saying they are either in or they are out, over a whole year having at least one episode, I guess that is an acceptable exclusion criteria. It would not be, in my opinion, if you were counting the number of episodes they have each year.

DR. MILLARD: No, doesn't it count the number? I mean it is just you can't get included as a numerator if you are not already in the denominator.

DR. NEFF: Right.

CHAIR YAWN: I understand that, but I mean just I wouldn't take them out of the denominator if I were saying you could be in the numerator 10 times.

DR. MILLARD: Okay. Right.

DR. RASTOGI: Right.

CHAIR YAWN: So, okay? All right. So what do you want to do with that?

DR. MILLARD: So shall we give that a "P" for partial, because we have questions as to the numerators and denominators?

CHAIR YAWN: Well, is risk

adjustment under this same thing, too? No.

MS. PACE: Well, this is just the specification of the risk adjustment. The actual testing and validation --

CHAIR YAWN: Yes, I know. I am just saying all of the risk adjustment --

MS. PACE: Right.

CHAIR YAWN: -- is here, too.

DR. RASTOGI: Okay. So why would you call it a "P", and what could make it a "C"? Like the exclusion piece?

CHAIR YAWN: No. I am just trying to see if there is something in risk adjustment that makes it a "P". Why is it a "P"?

DR. MILLARD: As opposed to a --

CHAIR YAWN: "C".

DR. MILLARD: Well, I have questions about not including the denominator as the first exacerbation of COPD. In other words, if you are a new patient, never been seen, the first time you are seen is because of an exacerbation of COPD, you are not counted until you have something else. Patients with COPD do have, often they have clusters --

CHAIR YAWN: Yes.

DR. MILLARD: -- but does that underreport? If one of the potentially avoidable complications is an exacerbation, I can't understand why you wouldn't include that as a --

DR. RASTOGI: So, if you are thinking in terms of a provider taking care of COPD patients, and you know that COPD is a chronic condition, you are just taking a snapshot in time, right? A one-year, whether you cut it here or you cut it here or you cut it here, it doesn't really matter, right; it is a COPD patient?

However, it matters if the patient switches providers, right? They go to one provider. This is another and another. And certain providers get labeled as having more PACs versus another, when they start off with a patient.

DR. MILLARD: But you said it wasn't going to be provider-specific.

CHAIR YAWN: We can't tell about providers anyway.

DR. RASTOGI: Okay.

CHAIR YAWN: It would be more if they changed insurance companies.

DR. RASTOGI: Yes. Sure.

CHAIR YAWN: So the first time you see them from your insurance company.

DR. MILLARD: That is why I have questions.

CHAIR YAWN: Okay.

DR. NEFF: It is just it is a little confusing.

CHAIR YAWN: Oka.

DR. NEFF: I mean, as much as we are sort of catching on, it is still a little

hard for us to get our brain around the categories.

CHAIR YAWN: Well, yes, and I think the fact that, you know --

DR. RASTOGI: It is a slightly different approach. You know, it is a completely --

CHAIR YAWN: But you are also going from, if they change providers, but this has nothing to do with -- it depends on what your definition of a provider is.

DR. RASTOGI: Exactly.

CHAIR YAWN: If your provider is an insurance company, yes, you change insurance companies and it might make a difference. But if you change from one physician to the next, since this isn't reported at the physician level anyway, it doesn't matter.

So, okay, I think we will go with the "P" because of that confusion that we have. You may not have the confusion, but we do.

DR. RASTOGI: Okay. Well, I could clarify it, but we are running out of time.

CHAIR YAWN: We need to move on.

DR. RASTOGI: Yes, yes.

CHAIR YAWN: Okay. Go ahead. Let's go on.

DR. MILLARD: 2b, reliability testing. I don't have any --

MS. PACE: When you say you tested the data on two datasets, but, okay, I guess

maybe you did not do any formal reliability testing. So this is issue that was brought up about claims versus --

CHAIR YAWN: Medical records.

MS. PACE: You haven't done any kind of testing to see how close or --

DR. RASTOGI: Right.

CHAIR YAWN: Right. So that I think the reliability, because that is the only way that I know you can test the reliability. I mean we all know there is an issue. We just don't know big an issue, and it is probably different for different conditions, and you have 50 thrown in there. So we don't know.

DR. RASTOGI: And for different health plans, right.

CHAIR YAWN: Certainly.

DR. RASTOGI: Yes.

CHAIR YAWN: So we are going to say "N" for that, I think, because there isn't a reliability test.

Okay, can you go on? Validity.

DR. MILLARD: Validity? Well, I mean, is that --

CHAIR YAWN: That is face validity. Those physicians are different face validity than we saw. But any other kind of validity? No, I guess test/retest. I can't imagine you are going to get a whole lot different number if you go to the same dataset and use the same thing two weeks later, but any other kind of

validity testing that should be included that wasn't?

You're thinking, Karen; I can see.

MS. PACE: Well, I mean, you know, in terms of the score, you know, do the kind of scores that result from this measure, correspond to something else that we know about the quality at that level? So, at the health plan level, it is a little harder to know, but --

CHAIR YAWN: Well, also, when you have, you know, 87 percent of everybody has one of these or 74 percent, that was one of the other problems I really had, is when everybody's got them, yes, what does that mean?

DR. RASTOGI: It means that we are not really paying attention to quality, and we are closing our eyes to what is happening in the real world. Because we just focus on one limb or one aspect of the patient rather than the full patient.

CHAIR YAWN: Well, and that is the U.S. healthcare system.

DR. RASTOGI: Exactly.

CHAIR YAWN: It is called having some specialty care --

DR. RASTOGI: Yes.

CHAIR YAWN: -- instead of primary care.

DR. RASTOGI: It is interesting that the responses we saw, as we are going

into the community presenting this stuff, that physicians like the AGA approached me, the American Gastroenterological Association. They wanted to do the good in colonoscopy measures. They sent two of the physicians to work with me to just line out how to line up the thing.

The same thing with pregnancy delivery, you know, Dr. Elliott Main and Debra Bingham from CMQCC, they approached me, and they are like, "We want to help you develop this for pregnancy delivery, so it has meaning." So they sat down and said, "What's that? What's typical? What's excluded?" They went down the list and they said, "This is what we want included."

CHAIR YAWN: Well, and I think you are getting to exactly what the concern is of the people here. We don't hear exactly that you have done that for COPD.

DR. RASTOGI: Right. For COPD, we presented --

CHAIR YAWN: We still have a very big gunshot.

DR. RASTOGI: COPD was presented to our design team, and they were -- I can get you the names. They were three family practitioners, one pharmacologist, two cardiologists, Dr. Allan Kahn from Blue Cross/Blue Shield, Mary Beth Rosenthal, et cetera.

So there were people on the design

team who have worked with measures and who have worked with these kinds of issues, and it went back and forth. This is a process that has been going on now since 2006.

So, like I had mentioned in one of my emails to Alexis, that version 1 was initially developed, then 2, and now this is the version 3 codes that have come out.

DR. NEFF: And I think the other thing, as this evolves, because I mean I think what you are hearing, too, is that there is a lot of enthusiasm for the work you are doing to try to figure out a good way to use administrative data, because we know that in a lot of cases that is what we are stuck with, and nobody has really figured out a good way to use it.

So, despite all of this, you know, we encourage the effort. I think what you are hearing from us, though, is just that it feels weird when the codes that are on there are just maybe the first-pass codes. So it may not really reflect what drove them to the visit, and then just the lack of maybe specificity to COPD -- I would think if I were in my gut summing up some of the persistent themes kind of with this particular measure. So it is fine, I mean --

CHAIR YAWN: Yes, and I apologize, I really do have to go. But I think that we have heard a lot of the --

MS. WINKLER: I would like to try

to finish out this one, especially with all the measures being very similar, I think, going through it.

DR. RASTOGI: Yes.

MS. WINKLER: The transcript will be available. We will be able to share it with Barbara. But if we can just kind of address each of these subcriteria --

DR. MILLARD: Barbara, what is your vote on validity testing? What is your sense of it?

CHAIR YAWN: Oh, I don't think they have more than face validity.

DR. NEFF: So maybe an "M".

DR. MILLARD: "M"? Okay.

DR. NEFF: I mean because there are groups that came up with it. We are just not seeing it, but somebody did.

CHAIR YAWN: Yes. No.

DR. NEFF: I mean so there's --

CHAIR YAWN: Oh, it is not whether their testing was right or wrong.

DR. NEFF: No.

CHAIR YAWN: It was that face validity seems to be the only testing that has been done for validity, which would make it "M", I agree.

DR. MILLARD: Exclusions justified. The inclusions are limited to very few criteria. And most of these are pretty straightforward.

I am not sure how you could get

excluded if you had lung volume reduction surgery if you didn't already have a diagnosis of COPD. That's why. You have to have COPD in order to have lung volume reduction surgery.

DR. RASTOGI: Yes, but once you have COPD and then you got lung volume reduction surgery, then that patient gets excluded. So it is a retrospective analysis. You know, when we create the models, which patients are finally selected into the models, so this one they are removed.

DR. MILLARD: Why would you exclude retrospectively patients for lung volume reduction? They have bad COPD. They have lots of complications.

DR. RASTOGI: That's right, and then it would -- you know, the cost, because, like I told you, this is more like a cost model. The cost associated becomes a dog versus a tail, you know, what's wagging what? We didn't go into the costs associated with that and the complications that may happen after the volume to overwhelm that whole episode.

DR. NEFF: So then thinking that their visits subsequently are more a consequence of their surgery and their LVRS even though they still had COPD as the original driver?

DR. RASTOGI: Yes.

DR. MILLARD: Can I ask an analysis

-- one lung volume reduction surgery, they had to suggest that you actually do better.

DR. RASTOGI: And I was thinking more from terms of cure because I am a surgeon and I have done some of those lung volumes. So I said, okay, these patients just do so much better. You know, maybe it would be like a heart transplant.

DR. MILLARD: But they have general --

DR. RASTOGI: Yes.

DR. MILLARD: But so many of the PACs are non-pulmonary.

DR. RASTOGI: Uh-huh.

DR. MILLARD: I don't know. We can say it is a "C".

MS. WINKLER: Any conclusion on reading? For exclusions?

DR. MILLARD: On reading?

DR. NEFF: For whether exclusions were justified.

DR. O'CONNOR: 2d.

DR. MILLARD: 2d, yes, I would say "C".

DR. NEFF: Yes, okay. You exclude patients that did not -- this is what I underlined and was trying to remember -- did not complete enrollment for the entire episode. And that I thought was creating some bias. Am I misreading that?

You excluded patients that did not complete enrollment for the entire --

DR. RASTOGI: So the enrollment thing is, if they are enrolled with that health plan, then some detail may go to some other health plan. So we are thinking you are not capturing that entire episode.

DR. NEFF: Okay.

DR. RASTOGI: That is how we were looking at it.

DR. NEFF: Okay.

DR. O'CONNOR: But if the patient stayed within the same health plan, I thought I understood you in the beginning to say that a patient who, let's say, was seen by urologist who treated a urinary tract infection, but didn't code for COPD, even though the primary care physician had already coded as COPD, you knew the patient had COPD, that that UTI event was filtered out?

DR. RASTOGI: Yes, that particular claim, yes.

DR. O'CONNOR: But if the UTI had been seen by the primary care physician who also coded for COPD and UTI, it would have been filtered in?

DR. RASTOGI: Filtered in.

DR. O'CONNOR: I have a problem with that because it is the same event, and it is just being treated by the entry source.

DR. MILLARD: And that would be program exclusions. Well, I would agree, yes.

DR. O'CONNOR: So I couldn't go with a "C" on this.

DR. MILLARD: You could say "M"?

DR. NEFF: I mean the bulk of this may be pretty well-supported. I think there are --

MS. PACE: And that is actually related to the numerator because that patient will still be in the denominator of COPD, but how that event is counted --

DR. O'CONNOR: But the PAC --

MS. PACE: Right, exactly.

DR. O'CONNOR: -- won't be counted the same way.

MS. PACE: Right, exactly.

DR. O'CONNOR: So I was trying to support Margaret's use of a "P" rather than a "C".

DR. NEFF: Yes.

DR. MILLARD: Okay. Fine, "P".

DR. RASTOGI: But the important thing to remember is, to the extent that PAC is a printer code, then those would be entered in. Okay? So if they didn't put COPD, but they put some of the codes that are there, as a required PAC, then they would be filtered in. Okay? And that is all in the expanded trigger section. Right? So all the acute exacerbations, some of the diverticular disease, the tracheostomy, everything that is on that page 2 of the triggers, which we call expanded triggers, and those are included.

DR. NEFF: So those are included regardless of whether they had COPD triggered

with them? No?

DR. RASTOGI: Right. So, if they didn't have a COPD code, but they had a pneumonia code, right --

DR. NEFF: Uh-huh.

DR. RASTOGI: -- or they had a --

DR. NEFF: But not a UTI?

DR. RASTOGI: Yes, urinary tract infection would be brought in.

DR. NEFF: Yes. Okay.

DR. RASTOGI: They are sort of remotely linked to COPD.

DR. NEFF: Right.

DR. RASTOGI: But we didn't think of them as printer codes, you know. But you are right, like to the extent if somebody overcodes and puts COPD and UTI, then they would have the ITI code.

DR. NEFF: Yes.

DR. RASTOGI: But just with the majority of the -- otherwise, there is so much junk in the data; we will get rid of the majority of the junk.

DR. NEFF: Okay.

DR. MILLARD: 2e, risk adjustment for outcomes, resource use measures.

DR. RASTOGI: You know, a code 15 has stayed there for over 11 years now. So you kind of know what gets put in when you put some of these things in it.

DR. NEFF: Yes. Some of it is primary more for function than maybe for --

DR. RASTOGI: Yes. Yes, not so much clinically, you know, Telesymmetry and Engenics, and all. You see all kinds of things there. When I go to the health market episode-based group, you know, system, then you will see, you know -- so working with different databases, you realize that sometimes to get the optimum results, you have to create these kinds of --

DR. NEFF: Well, the trick, then, is knowing what optimum there is. You know what I mean? If you are sort of molding the data, you've got to be pretty sure you're --

DR. RASTOGI: Yes, and that is why that consistency is important, that if you have the standardized SAS programs, and then all the databases are run through the same program, and that is available as a freeware on our website.

DR. NEFF: Right.

DR. RASTOGI: So anybody can down --

DR. NEFF: It is all right there?

DR. RASTOGI: Yes. Anybody can download it and run it through the database.

DR. NEFF: Yes.

DR. RASTOGI: Then it could be comparable across different populations.

DR. NEFF: So 2e, I think I was lost in this, just because there was so much.

DR. O'CONNOR: Do you have a biostatistician on staff?

MS. PACE: No.

DR. O'CONNOR: No?

MS. PACE: No, uh-uh.

DR. O'CONNOR: You might consider it.

MS. WINKLER: It is definitely something we have talked about. We have definitely talked about it.

DR. MILLARD: I would have to recuse myself from the 2e because I don't -- I mean, a bootstrap, I don't think I have ever done a bootstrap in the old system in my life, other than trying to figure out to put on shoes.

DR. NEFF: Yes. Other than knowing it, I couldn't actually speak to --

DR. RASTOGI: Yes, this particular modeling was done by Mass Crew. They have biostatisticians on staff.

So, basically, the bootstrap technique, what it does is it takes those -- it is a pretty standard, you know, validation technique. So they take the whole database, and they take, say, 200 different samples within that same database, and then they define which way to present the important or significant ones. If they stay significant in more than 80 percent of the sessions or runs, then those variables are selected.

MS. PACE: So do you have any -- so, in the results, you just reported the adjusted R-square, but did you do any of

calibration plots or --

DR. RASTOGI: Yes. All that, we can supply to you, whatever you need. But those details, you know, the analysis is there. There is inflation factors that are calculated and all the coefficients, you know. Dr. Arlene Ash from Boston University has reviewed many of these models.

DR. NEFF: It is pretty complex modeling, to be sure.

DR. RASTOGI: Right.

DR. NEFF: First of all, SAS gives me PTSD, but other than that -- and I think this piece of it could be validated with a biostatistician and talking with --

DR. MILLARD: Can we put in an asterisk and say --

DR. NEFF: This could just pass for now and could sort out later.

MS. WINKLER: Actually, we just had insight over who we might get to do that. This is a multi-advanced model.

DR. MILLARD: 2f. 2f, "identification of meaningful differences in performance. Accountability for and measurement of PAC occurs at the practice medical group and provider system or purchaser or payer level, not from the individual physician performance. Calculates absolutes, not relative values."

The objective of the measure is to encourage the unit being measured to

progressively reduce the amount, not to discriminate performance between two units of measure."

MS. PACE: So I am not exactly sure what you are getting at. The 80 percent, you would be -- that would be a risk-adjusted rate? Or are you saying, when you are saying absolute rate, are you referring to a non -- I am not sure how that fits with --

DR. RASTOGI: So the idea here was this, like I mentioned earlier, it is not comparing one provider with the other, right? What we are really trying to get at is, can the same provider improve the PAC rates, right? So, if they have 80 percent today, maybe a year or two years down, can they make it 76 percent or 77 percent?

MS. PACE: Right.

DR. RASTOGI: That is the whole idea behind it.

The risk adjustment, like I was telling Reva in a separate meeting, and all that, is more comparing, say, one population versus the other, if the severity of the patient is higher in this particular population versus in the other, right?

And the risk adjustment model is not done on PACs. It does only on typical care. So, then when you look at patients who have a typical episode of COPD, you understand what are the other co-morbid conditions that are present in that particular patient. Those

are the risk factors or risk variables that go into the model to determine the severity of the patient.

MS. PACE: Okay. So correct me, then, if I'm wrong. So what you are saying is you are not really using that risk model for this score.

DR. RASTOGI: Yes.

MS. PACE: So I am not sure why you presented it.

DR. RASTOGI: Okay. So, if you are saying the score is 80 percent PAC rate and all, that is right. If you look at the 80 percent PAC rate in one population and you look at another one, that can be adjusted by the severity index. That is why we showed you how this severity index is calculated, right?

So all you are doing is, say the severity index for this population is 1, and it is 1.2, here is it 80 percent and here it is 75. Then you can adjust the 75 based on the severity index and say what it is. That is only to that extent it is used.

MS. PACE: Okay.

DR. MILLARD: But we don't know what meaningful -- has it ever been validated what a meaningful difference in performance is?

DR. RASTOGI: A meaningful, yes.

MS. PACE: Well, what we are getting at there is, and this is where it gets a little unclear because you keep talking

about quality improvement, but if you do public reporting, there is going to be comparison.

DR. RASTOGI: Yes.

MS. PACE: So you would expect the risk adjustment to --

DR. RASTOGI: Exactly, and that is why that severity index is calculated, and that is how --

MS. PACE: So what we are getting at here, so say you are publicly reporting the score on two health systems.

DR. RASTOGI: Yes.

MS. PACE: How do you determine whether, you know, 80 percent in one and a 78 percent in another is a difference or if that is just due to measurement error? Have you done any work on that yet?

DR. RASTOGI: Right, and the only thing we can say is, to the extent the severity is almost the same, then we would say, you know, this is 80 percent and this is 76 percent. Now how different is 76 percent from 80 percent, what's the P-value and all that, no, we haven't calculated those.

DR. MILLARD: So, in a sense, it is not really validated in terms of the differences?

DR. NEFF: I think you get stuck, too, with just the reliability kind of affects this, even though it is in its own little category.

DR. RASTOGI: Uh-hum.

DR. NEFF: You know, if you are still at the stage, where you are not quite at the stage yet where you have done that sort of to the chart, reliable, you know, where you kind of can know that these numbers that you are seeing up here are really reliable, it, unfortunately, filters into all this and makes it harder to trust the differences you are seeing at this stage.

Now it sounds like you are moving forward and will have some of this in the near future.

DR. RASTOGI: Yes, and like Barbara pointed out, working with claims data, it is very different.

DR. NEFF: Yes.

DR. RASTOGI: When you match it up with chart review, you know, when I go to United, we did so much chart review, and it doesn't match up sometimes.

DR. NEFF: Right.

DR. RASTOGI: But this is what we have. So, if you are going with the administrative claims data, this is what you are stuck with. Now how do you match up and how do you compare? And you say, so should we not do the administrative data? That is a completely different question, you know.

DR. NEFF: I know.

DR. RASTOGI: And, currently, it is correct that we are using the fee-for-service

system. So there is a lot of different incentives for coding, which Barbara was pointing out. That is absolutely right. But whatever exists, that is what we can point out.

DR. NEFF: Yes.

DR. RASTOGI: And when the system changes to, say, episode-based payment, the coding practices may change.

DR. NEFF: Right.

DR. RASTOGI: And then we may have different kinds of drivers of costs that come up.

MS. PACE: So probably for now, for 2f, "M", yes.

DR. MILLARD: 2g, are we going to say the same thing as we said earlier; we pass on it because that is sort of --

MS. PACE: 2g is probably not applicable.

DR. MILLARD: Yes.

MS. PACE: It is only the administrative.

DR. MILLARD: Disparities in care, 2h --

DR. RASTOGI: We didn't do any disparities.

DR. MILLARD: And there's no disparities. So do we say "NA" or not?

MS. PACE: If there's no disparities identified, then it is "NA".

DR. NEFF: I mean it,

theoretically, could be extracted, if you wanted to.

MS. PACE: Right.

DR. NEFF: I mean, so in the same way it is not lost; it is just not being sought out.

DR. O'CONNOR: We could create a new category, "WNL", "we never looked".

(Laughter.)

DR. MILLARD: And I think we have discussed the strengths and weaknesses fairly well.

Okay, usability, No. 3. Meaningful, understandable, and useful information.

I think it would be somewhere in between "P" and "M", I think, on that, aren't we?

MS. PACE: It sounds like from your comments about how you interpret this overall score. What do you guys want to do?

DR. MILLARD: Margaret, what do you --

DR. O'CONNOR: Well, they say right in here, I mean that it is not applicable today.

DR. NEFF: Right. Yes. When I was reading this sort of in the other context, like the PACs make sense in the sense that you are trying to prevent complications, I mean that concept. It is just it is not quite linked yet to the COPD, in particular. So, I

mean, I think, as described here, it is probably not quite there now.

But the structure of it --

MS. PACE: I mean they presented that they have experience in presenting, but you are talking about presenting the whole analysis, not just the score, right? This discussion you put about presented the analysis to medical directors, CEOS, that have found it useful, that's --

DR. RASTOGI: Yes.

MS. PACE: That is using the whole system?

DR. RASTOGI: Right. And showing the actionable part, like we showed them which are the drivers. When they look at the top drivers, then they know that this is where they need to focus their efforts. So they found that very useful.

And, yes, the entire list of facts may overburden you, but, like one of you guys pointed out, that the ones that are not so relevant to COPD fall down on the list as low points, and then the ones that are very relevant for that particular episode rise to the top.

So you can see, to the extent people want to make it actionable, they have all the information that is there.

And in some databases, it flips, too. You see other things that are popping up.

DR. NEFF: I guess I can't get over the one spreadsheet we have where the acute exacerbation of COPD is like second from the bottom, and you have other stuff that is way at the top. You know, somehow, that doesn't --

DR. RASTOGI: That is the coding, yes. That is how they coded it.

DR. NEFF: That is the backbone of this whole thing, right, is the coding?

DR. RASTOGI: You could call it anything. You could call it PAC 21 and not worry about the name, right? But the important thing is you have to see what are the drivers, you know. So, to the extent that you call every hospitalization as acute exacerbation, then that is fine. You know, you could label it that way, too.

So it is not so much semantics. It is more about going into realizing what are the drivers and what are the codes behind it.

DR. NEFF: I think I am at "M". I mean as it stands.

DR. MILLARD: 3b, 3c, relation to other NQF-endorsed measures and harmonization.

MS. PACE: Well, this AHRQ PQI, I guess you didn't put that in there, but they identified it, but I don't remember exactly what that --

DR. NEFF: The similar related measures --

MS. PACE: Right.

DR. NEFF: -- the AHRQ PQI 15 or something or QI --

MS. PACE: Yes. Well, they identified it. The question is, and it wasn't relevant?

MS. FORMAN: No. We had a discussion and then we took it out.

MS. WINKLER: I think that the methodologies and the targets for each of those are somewhat different.

MS. PACE: Well, the methodologies are different, but is it a measure of complications?

MS. WINKLER: No, it is a measure of avoidable hospitalization.

MS. PACE: Oh, okay. All right. Okay.

Well, but it includes hospitalization.

MS. WINKLER: Yes, it is teeny portion of it, yes.

MS. PACE: Right, right. Okay.

MS. WINKLER: And it is condition-specific.

DR. RASTOGI: In COPD, 57 percent in that thing was due to avoidable hospitalizations, you know, when you look in terms of stays and percentages and all that.

DR. MILLARD: So we would say "C" on that 3b? 3b.

3c, distinctive or additive value. "Describe distinctive, improved, or additive

value that this method provides to existing NQF-endorsed measures."

Potentially avoidable complications encompass the majority of these effects.

DR. NEFF: And I guess what you are saying, it might just be an "NA"? Because we don't have anything on the same topic.

MS. WINKLER: Not to this degree. So the only thing is going to be the avoidable hospitalizations from the PQIs, but not the full breadth and extent of this.

MS. PACE: Right, but you are talking about harmonization?

MS. WINKLER: No.

DR. NEFF: So there's not really a way to answer that for this one. That question doesn't really apply.

DR. MILLARD: Okay.

DR. NEFF: I think -- no?

MS. PACE: Well, I think what Reva is saying is, if we don't have anything that addresses this --

DR. NEFF: Right.

MS. PACE: -- then it would be a good thing --

DR. NEFF: Oh, right.

MS. PACE: -- as the distinctive and additive value. But I think the other piece of this is it is only distinctive and additive value if you think that there is a valid way to measure the kind of issue.

So it is somewhat influenced by

some of your other discussion.

DR. NEFF: Right.

MS. PACE: So it is a little bit tainted or affected, or could be.

DR. NEFF: Right. It could be affected positively or negatively by the other --

DR. MILLARD: So do you want to say just an "M" then?

MS. PACE: And probably the more important thing is in the strengths and weaknesses, to say that, you know, this is not addressed by other measures.

DR. NEFF: Right.

MS. PACE: However, value-added depends on our discussion about the scientific acceptability or something.

DR. MILLARD: Yes. I mean the concern I have is that the PACs are defined so broadly that the effect of COPD management, good COPD management, may not really be influenced, the influencer.

DR. NEFF: So not a "C", but not an "NA". Because there's potential, again, for a big addition --

MS. PACE: Right.

DR. NEFF: -- if the target can get something. Get rid of sort of the noise of all the sort of the billing codes that aren't relating to the charts.

DR. MILLARD: The other thing is I have a note here about age, 18 versus 40. Are

the PACs, was this 18 or 40? Is there an age limit on these or not?

DR. O'CONNOR: It's above the age of 18, I think. They segregated two populations, pediatrics -- oh, I'm sorry.

DR. RASTOGI: Yes, for asthma, it is 2 and 17 and then 18 and above; for COPD, it is 18 and above.

MS. PACE: Oh, so that is a harmonization?

DR. O'CONNOR: Yes, that is a harmonization.

MS. WINKLER: Do you have -- I mean you must have -- the data you can stratify by age to know really how many under the age of 40 and what impact that has on the whole --

DR. RASTOGI: Yes. Yes, it is very easy, you know, because once you have the whole data, you can just cut wherever.

MS. WINKLER: And the reason we care about it is, when you put together NQF's portfolio of measures around COPD, you like to look at them as a package. And for those really focused on COPD, you would want them to be able to implement them all. But if it takes different algorithms and different implementations and has different rules, they don't do it.

DR. RASTOGI: Yes.

MS. WINKLER: So the harmonization will facilitate implementation. So, to that degree --

DR. O'CONNOR: But as it is written.

MS. WINKLER: It is 18.

DR. O'CONNOR: It is 18.

MS. WINKLER: Right.

DR. NEFF: So the recommendation would be for to harmonize.

MS. WINKLER: The harmonization would be useful.

DR. NEFF: For the COPD population.

DR. O'CONNOR: The way it is currently constructed, if the patient alpha 1 antitrypsin deficiency, they would have been included in the COPD population 18 and over probably.

DR. RASTOGI: That's right. We don't exclude.

DR. MILLARD: They may never split.

DR. RASTOGI: Yes, we don't exclude specifically.

DR. MILLARD: Okay. So feasibility. Data generated by a byproduct of the care process. The answer is yes. So it is "C".

Electronic sources. Are all the data elements available electronically? That is how you get --

MS. PACE: Could I go back to this? I realize this is an area that we have to do some better descriptions. But this is based off of codes that are generated by someone other than the people doing -- so we would not

consider ICD-9 codes on claims as data generated during the care process.

DR. NEFF: But that is for the billing, isn't it?

MS. PACE: Yes, but that is not for care.

DR. NEFF: Oh, no, but it is --

MS. PACE: It is for billing.

DR. NEFF: I guess when I look at that, I am like you are having to do something more than what is already happening as a consequence of their clinical stay, which they are going to get billed.

MS. PACE: Well, that is what I am saying. We need to define that.

DR. NEFF: Yes.

MS. PACE: But, obviously, we need to define that better. But the real intent of that is, you know, a blood pressure that is taken by the clinical person and used in the treatment of their care versus coding goes through another person.

MS. WINKLER: Kind of like what Barbara was talking about.

MS. PACE: Yes.

MS. WINKLER: You know, you do the chart and take care of the patient, but someone else abstracts that, interprets it, assigns codes to become part of the billing.

MS. PACE: Right. So we are not saying that is bad. We are just saying it is not data that is generated by the people --

DR. O'CONNOR: That is probably true in the hospital, but not necessarily in the ambulatory setting. Because when I see a patient in the office, I am the only one that codes.

MS. PACE: Okay. That is good.

DR. O'CONNOR: And in fact, if I have overcoded, I will get a tap on the shoulder.

MS. PACE: Okay.

DR. O'CONNOR: But I'm never tapped if I have undercoded. They don't care because they are not going to arrested for undercoding.

MS. PACE: No, that is a good distinction.

DR. NEFF: Yes, you're right.

DR. NEFF: Well, and I guess, you know, as you guys tweak these more and more, figuring out what really the goal of that is, is it extra work, which wouldn't be the case with billing because that is going to happen anyhow, or if you really wanted to focus just on the clinician activity, I mean not --

MS. PACE: Right. No, no, that is good. I think that is the --

DR. NEFF: Because the other flip side of that were when someone had to actually go and abstract data specifically for --

MS. PACE: Right.

DR. NEFF: That is even a third.

DR. MILLARD: It is not like the

ICU stuff.

DR. NEFF: Yes, right, but you've got to have a body that would never go to get that unless for this measure.

MS. PACE: So we need to definitely --

DR. NEFF: There could be value in knowing all of that.

MS. PACE: No, that's good. That's good.

DR. MILLARD: And electronically available.

DR. NEFF: Yes. Yes. No "M" on it, the only one. It is actually all electronic. Yay.

DR. MILLARD: Exclusions due to specific -- require additional data sources beyond what is required. But it is all done electronically. Clear.

Subject to inaccuracies, errors, and unintended consequences.

DR. NEFF: I had, is it validated? How good is the code? I mean it is really what we have been talking about. Oh, in fact, that is what you guys said. "PNC analysis is as good as the coding." Indeed.

DR. MILLARD: So are we going to say that that is "C" or "P"? I am not sure about the -- I mean the coding is the coding. It is "C".

DR. O'CONNOR: I don't think that they are at any increased risk of anything

else we have discussed today.

DR. MILLARD: Yes.

DR. O'CONNOR: There is nothing uniquely specific to this particular measure that would probably downgrade it. We have given every other one a "C" on that.

DR. MILLARD: So, yes.

Data collection strategy --

DR. NEFF: Wait, wait, wait. What's 4d?

DR. O'CONNOR: 4d.

DR. MILLARD: Yes.

DR. O'CONNOR: Is this data susceptible to inaccuracies? That is probably true of any dataset.

MS. WINKLER: Is it more so with the coding than, say, the abstraction of the data elements for the mortality model?

DR. NEFF: Oh, so you guys are reading this as, once the data is done and then presented, is it at risk -- not you guys. But that is sort of, is it at higher risk for being misinterpreted, not so much whether the data are accurate or inaccurate.

DR. O'CONNOR: Oh, I see the distinction.

DR. NEFF: I think it is all how you -- I think you are right, though, the way we have interpreted this previously was, once you have the dataset and you are presenting it, is there risk of it being inaccurate, misinterpreted, more so than others? In the

past, we have said no more than anybody else. That may be an issue unrelated to the accuracy of the coding.

DR. MILLARD: Because in this case the coding is very important, because if the urologist doesn't mention COPD, it gets lost.

DR. O'CONNOR: Because we --

DR. RASTOGI: But that is not so much, and I don't know why we are so hung up on it. The main complications will be captured here. So they are in here looking at the one complication that doesn't --

DR. O'CONNOR: Is that a reflection of the data or the interpretation of the data? No, this is for the raw.

DR. NEFF: Yes, right.

MS. WINKLER: I think probably it could be any and all.

DR. NEFF: Right.

MS. WINKLER: If we think about it, it is collection of the data. So this would be coding errors. This would be interpretation, as opposed to, if the data element is in an EHR, you only click it once and it is what it is; it doesn't get translated.

But then you also potentially have inaccuracies in how you combine data. Methodologically, there is a potential. So I don't think it is all of them or any or all of them, if you can kind of envision the kinds of problems you see.

DR. MILLARD: And unintended consequence, the question is whether or not that is the -- could the data ever be used to look at an individual physician?

DR. NEFF: Could people use the data in a way that you are not planning on them using it?

DR. RASTOGI: Sure. Yes.

DR. NEFF: I mean which is sort of what we have said about anything could happen.

DR. RASTOGI: Sure. Yes.

DR. NEFF: So I think we have been doing, you know, I think it has got a little bit more risk because of the coding interpretation issue, which you highlight yourself, more so than our other ones, which maybe didn't have the coding pieces. And then all of them are at risk for just being misused, which you can't do anything about.

DR. MILLARD: Would you say "P"? Just say "P"?

DR. NEFF: Yes, I think so, just to touch different things, yes.

DR. MILLARD: That will work.

DR. NEFF: Data collection strategy.

MS. PACE: Well, basically, there's no issues. I mean it is all administrative claims.

DR. NEFF: Well, that's true. How they are getting it is, you know, it is fed in in a direct way. Yes.

DR. RASTOGI: Yes, the data formatting is very important, you know, and stuff like that. We have seen if they don't have procedure codes in the claims data, then it causes problems. If they don't have multiple diagnosis codes in the data, then it causes problems in the risk adjustment.

DR. MILLARD: So I am saying that is "C".

MS. PACE: Yes.

DR. MILLARD: That is good. Okay.

Strengths and weaknesses in relation to subcriteria feasibility. We have sort have been all over the map on feasibility, haven't we?

DR. NEFF: Yes.

MS. PACE: Yes.

DR. NEFF: I mean I think the gist of this is just the challenges of the administrative data, I mean really, than just whether it is, as is, ready as a measure for itself. There may be other work that you are already planning to do before it gets to that point, but, you know, it is good to figure out where the holes are, based on the criteria.

DR. RASTOGI: You know, just I would like to try one more thing, you know. Like I know we were kind of starting off trying to understand the whole thing.

The symptom logic I don't think is intended to -- the intention there is to get rid of the noise, you know. So, if the

majority of the complications which are COPD-related, all those complications themselves serve as filter codes, you know. So, to the extent they have something like that that is going on which physicians and pulmonologists have identified as COPD-related, all those are critical. So not only if the COPD diagnosis is present, but any of those other diagnoses are present, it is --

MS. PACE: When you say, "filter code", you mean codes that identify that the patient would be in the denominator?

DR. RASTOGI: No, the claim. So the patient is the trigger, right?

MS. PACE: Right.

DR. RASTOGI: So, then, if the trigger code comes in, then the patient gets counted. Then we would count all the things that happen for that patient for one year, you know.

MS. PACE: Okay.

DR. RASTOGI: So there are some codes that get excluded because it is a major surgical procedure.

MS. PACE: Okay. All right.

DR. RASTOGI: So, even though it is happening in a COPD patient, it is not related to COPD, so the claim gets thrown out.

Then other things get thrown out because of the splinter thing, which somehow we kind of kept going round and round on the urinary tract infection thing. But to the

extent if it is a very popular potentially avoidable complication in COPD, it would have featured into the filter, you know.

The same thing for asthma. If it is not an important complication related to that condition, it won't be a filter, you know. So that is the point I want to make, is when the claims are pulled in, we want to make it as relevant to COPD as possible. That is why the outward results that you see are more relevant to COPD.

DR. NEFF: You know, I wonder, and maybe you have this in here, although I don't think I saw it, as you are sort of evolving this over time, if there might be some sort of schematic.

DR. RASTOGI: Yes.

DR. NEFF: You know, as to what data is moving through and filtering and --

DR. MILLARD: See a flow.

DR. NEFF: Did I? Oh, maybe I did see it.

DR. RASTOGI: And then we have a website which has -- I don't know if you had a chance to look at that, but that has an entire playbook on COPD. In the playbook, you have all the demographic information. You have the flowcharts. You have all kinds of, you know, information. There is a slide deck there which shows the entire process on how it goes, you know.

DR. NEFF: Yes.

DR. RASTOGI: And this was all part of the RWG grant. So the grant money will run out the end of this year. So our development is done, you know. So now we are in the implementation phase, which starts next year.

DR. MILLARD: Do we have time or not to go over those COPD ECR playbook decision tree?

MS. WINKLER: Well, sure. There is no reason not to.

DR. MILLARD: Because I would like you to -- because I spent too much time on this trying to figure it out.

MS. WINKLER: Where is it?

DR. MILLARD: It is under --

DR. RASTOGI: Yes, it is the first tab on that one. Thank you for sharing.

MS. WINKLER: It is in the risk adjustment.

DR. MILLARD: Yes, risk adjustment.

DR. RASTOGI: That same worksheet, yes.

MS. WINKLER: Is that what you are talking about?

MS. WINKLER: Yes.

DR. MILLARD: Go to the top.

DR. RASTOGI: So, if you go to the top --

DR. NEFF: Oh, yes, okay.

DR. RASTOGI: So we start with the development of the database had 4.7 million covered lives and $95 million. Then it goes

through the various steps in which we do the enrollment. You know, some patients get excluded because they don't have continuous enrollment. We are allowing a 30-day gap. So that is that next piece.

DR. MILLARD: Okay. So, first, does the beneficiary have a trigger code --

DR. RASTOGI: Yes.

DR. MILLARD: -- and a physician on a professional claim? And the trigger code is --

DR. RASTOGI: On the old codes worksheet in the first tab. It says, "COPD trigger".

DR. NEFF: And then there is the extended.

DR. MILLARD: And that is bronchitis, emphysema, but that also includes end-stage renal disease? No, that is a terminator. Or is that --

DR. RASTOGI: So that tab before that, you know, as you are looking at expanded trigger --

DR. MILLARD: Oh, okay. I thought you were on --

MR. AUSTIN: -- the one before that.

DR. MILLARD: Okay. Okay.

MR. AUSTIN: Yes.

DR. MILLARD: Essentially, a COPD code?

DR. RASTOGI: Yes, uh-huh.

DR. MILLARD: Okay. Then patient episode. Now patient episode, is that the number of patients or is that the number of times the diagnosis is read.

DR. RASTOGI: No, number of patients.

DR. MILLARD: Okay. Okay, so that is not patient episodes.

DR. RASTOGI: So they start an episode.

DR. MILLARD: That is patients?

DR. RASTOGI: The episode is one patient for one year. That is the episode. Right?

DR. MILLARD: Okay.

DR. RASTOGI: And if they didn't have the one year's worth of claims, then the episode doesn't form.

DR. MILLARD: Is the patient over 18?

Okay, keep going.

DR. RASTOGI: Yes.

DR. MILLARD: There's the answers, that they had less than 10 percent below -- no. They have one year continuous -- okay.

Now what are reasonable episode costs?

DR. RASTOGI: So we are removing the outliers. So, if the entire episode costs more than $2 million, the medical part has to be more than $1 million and then the pharmacy $1 million. So the entire thing, you know.

So, then, if it is more than that, then the episode gets removed.

And this moved from our linked purposes, you know, the entire piece, so that the cost doesn't get skewed by one or two patients.

DR. MILLARD: Okay. Then it says, "Is the episode free of acute diagnosis codes and termination codes?"

So if they didn't get diagnosed --

DR. O'CONNOR: If the episode is free of an acute episode, you go straight down.

DR. NEFF: Oh, and you exclude the acute because you are not doing that first hospitalization? Is that right?

DR. RASTOGI: That is right.

DR. NEFF: Yes.

DR. RASTOGI: The initial trigger piece.

DR. NEFF: Right. So that is a freebie, basically.

DR. RASTOGI: Right.

DR. NEFF: And then they get triggered on the next one.

DR. RASTOGI: In the next one, right.

DR. NEFF: Yes.

DR. O'CONNOR: Wait. But she just defined as episode as a year.

MS. WINKLER: Yes. So a year starting from the trigger.

DR. RASTOGI: Right.

MS. WINKLER: Assuming the trigger isn't one of those acute exacerbations.

DR. RASTOGI: Right.

DR. O'CONNOR: And if they have exacerbations during that one year --

DR. RASTOGI: Yes, and maybe it could be written better, but here what they are doing is we are removing -- and, you know, there are several steps that happen, and we have just shown some major steps here. But here we are removing all of the exclusion criteria.

So, in the all codes tab, you may have noticed we have a medical tab and a procedure tab. The medical tab links up to the CCS classification AHRQ in the dataset. So all the 10,000 codes which have not been put into the expanded triggers are now being grouped using the CCS classification.

You know, if they have any of these HIV conditions, cancer, if they have some of the other conditions, pregnancy delivery, et cetera, those are exclusions. So the patient, if they have those conditions in the presence of COPD, they get removed. So there is termination; you know, ESRB patients, et cetera, get removed.

We also exclude claims, you know, if they had a major surgical procedure, then those are excluded. So that is identified in the procedure tab.

So all that cleanup happens at that stage.

DR. MILLARD: Of course, you know, when you were talking about outliers were removed --

DR. RASTOGI: Yes.

DR. MILLARD: -- one-third of that patient population was removed as an outlier in this.

DR. RASTOGI: No, I don't think it is one --

DR. MILLARD: Yes, you go from 419,000 to 272,000.

DR. RASTOGI: That is the enrollment.

DR. NEFF: Yes, the enrollment got rid of a bunch.

DR. RASTOGI: Yes, enrollment cut into halves.

DR. MILLARD: Yes, and reasonableness of cost. Okay. So you grouped the two.

So, then, we go down. So they are free of the termination codes. Then do they carry a COPD-related diagnosis code? Yes. So free of medical exclusion criteria and other -- so you get to 97,000 patients.

DR. RASTOGI: Right, and, you know, I double-checked because I was also worried about why the number dropped so much from 272,000 to 97,000. And what I realize is in this particular output we are showing only the

commercial population because we were creating the model on the commercial to compare it with the other health plans. So it is from 18 to 64 years of age. That is what is on the website right now.

Now we have version 2.3, which is -- so version 1 is what is on the website. Version 2.3, which we did the entire population, 18 all the way to 120 years, or whatever.

(Laughter.)

So then the drop wasn't there, you know. So all this is, you know, and like you were saying, if you want to cut it at 14, you know, then it is very straightforward at that.

DR. MILLARD: And the difference between PAC and typical?

DR. RASTOGI: Right. So it is patients that have overlap. Right? So the same patient could have some claims which are typical and some claims which are PACs. There would be very few patients who would have only typical claims or only PAC claims.

DR. MILLARD: What is a typical claim?

DR. RASTOGI: So anything is not PAC is typical. Okay? So, basically, during that whole one-year episode time window, we have removed the exclusions of the irrelevant claims. Whatever is left is the relevant. Then those are, then, sorted out. Do they have a PAC code on it or not? If they don't,

then they are typical.

So the dollars add up, but then the patients are overlapping. So you wouldn't be able to add the two patients and make it equal to relevant because the same patient may have typical as well as PAC claims.

DR. O'CONNOR: Do you think there is a dataset or a study that has shown that what has been identified as a PAC is truly an avoidable complication? I mean they were defined that way for the purposes of the study. But is there any other data to suggest these are truly potentially avoidable?

DR. NEFF: It sounds like, at least from you were saying, it sounds like these came from AHRQ PACs, how they defined them? No?

DR. RASTOGI: No. AHRQ --

DR. NEFF: None that I know of, but --

DR. RASTOGI: Yes, AHRQ hasn't defined PACs. They have only said that hospitalizations which are there, right?

DR. NEFF: Ah, okay.

DR. RASTOGI: So, then, most of the definitions for PACs are clinically-based and based out of the design group suggestions and all this.

DR. NEFF: Okay.

DR. RASTOGI: So you are right that some of it could be controversial. People could be questioning them. But, you know, to

the extent they were part of the CMS definitions, so you know all the DBTs and, you know, fracture of femur, et cetera, those have been put in, and you have seen --

MS. PACE: I just wanted to mention the prior project I was working on. Hospital Outcomes reviewed some measures that were somewhat similar to this. I just thought I would tell you, you know, some of their comments are similar to some of the things that you have raised.

One was about, you know, the reliability of the data items, and they were especially concerned with reliability when the claims data are used to measure the outcome. If it were just being used for like risk adjustment, they thought that had maybe a little bit different -- that since it is the outcome, that it carries more weight of being concerned about reliability.

And they had a couple of measures. One was where they just kind of identified all complications and risk-adjusted, and so you didn't have this idea of what was preventable or not preventable. There's advantages to that, in that, you know, you just measure everything, you risk-adjust, and you look at differences.

They also had some measures where the developer had tried to identify preventable re-admissions, preventable complications. Then the issue came up of, who

made those decisions? How repeatable is it?

I mean, in some of these systems, they were looking at like 10,000 diagnosis pairs. When you start multiplying -- so the questions started coming up, well, you know, would another group who looked at these 10,000 pairs come up with the same list?

So I think the bottom line that I think came out is -- and it is probably something that we just don't have a good -- and the way you kept presenting it is the system, and so did they.

So the question is, you know, whether NQF, do we need to think about some other things when we look at these kinds of systems versus a discrete measure? Because when you submit a measure to us, that is the measure that ultimately we think should, you know, if we endorse it, should be publicly reported.

But the real value in these big systems is for quality improvement and being able to drill down into that data, and it doesn't exactly fit with our traditional NQF-endorsed measures.

So I don't know if any of that resonates with you, and some of that kind of overlaps with some of the comments that you all were making on this measure. So I don't know if you have anything else to say in those regards.

DR. NEFF: Well, it reminds me a

lot when I sat on the IRB and you would get these really complicated, very specialized type studies, whether it is genetics or emergency consent or something. And they are in enough different fields that there almost is a different way to deal with them. I mean it is a little bit what you are saying, that it is not so much that it is COPD or CHF or hospital outcomes. It is that it is almost its own little category.

DR. RASTOGI: Yes, and it is kind of --

DR. NEFF: I don't know --

DR. RASTOGI: Yes, it is kind of similar to your first definition. Like we are calling them as PACs, but we don't know how many of them are avoidable, right? So we are identifying these complications and we are saying that, really, for a patient, they shouldn't have these, right?

Now to what extent can they be preventable? You know, time will tell, like we have discussed, right?

And it is something like you were saying, you know, HACs, PSIs, you know, all the EHR to define, you know, these are standard definitions are good across the board. Then there are some additional which are more specific.

But if you think in terms of a whole patient, and a patient-centered approach, then you want to make sure that they

don't have any of those complications. So that was the premise behind it.

MS. PACE: And I think, you know, it is one of those things, when you have this claims data and you have all of this data, it is hard to even think about how you would do a reliability study, if you are including everything versus if you could construct a measure around those things that are most frequently the complications. You know, then you can actually manage.

So I don't know what the answer is. I just know these are the issues.

DR. MILLARD: In designed clinical studies, we are supposed to narrow down to -- I mean the best studies where you have such a simple, straightforward outcome, small, narrow population, you know exactly the questions you are going to ask, so there's no variables.

Now we are looking at an entire patient population.

DR. RASTOGI: Yes.

DR. MILLARD: So it is an entire -- you have to sort of throw your assumptions out the window and come in with an entire --

MS. PACE: I mean it is also kind of, you know, because of the large databases, it is, in a sense, a little bit like data mining versus kind of constructing a measure conceptually first.

So it presents a lot of different questions and challenges to us. I don't know

if you all have any suggestions for us, you know, because I think we are going to see this more and more. We are seeing it more and more, and I am not sure that we know exactly how to handle it or address it.

DR. RASTOGI: Yes, and when we began our analysis, we didn't know what to expect. We didn't know what the percentage would be. We had no clue what the big drivers would be.

And you're right, like right now, you know, almost everything is game; you know, it is all right. And then you look at the risk-adjustment model also, and most of the variables that you fed in are classic co-morbid conditions that you go in. But the output that you get is very specific for every patient population, you know.

So, for COPD, we are seeing different risk drivers. For CHF population, you see different things. You know, so it becomes very, very condition-specific when you start looking at the outputs.

DR. O'CONNOR: But I think before it rises to the level of an approved NQF measure, that there ought to be some aspect to it that provides value to the people who are going to employ it. And I am not convinced that that exists.

I mean you have shown these differences when you have run these models between various states, Arkansas and Alabama.

I can't remember which ones you have mentioned.

But how that is going to drive a different approach is unclear to me.

DR. MILLARD: I mean, if I know my patients, if I know that 25 percent of my patients with COPD have exacerbations that are potentially avoidable, then that is a number that I can go back and say, okay, how do I change practice to improve my outcome.

DR. RASTOGI: Thank you.

DR. MILLARD: But, right now, with this, I don't know. I just know --

DR. RASTOGI: But if you want to look at specific ones, you can see, right? So, if you only interested in acute exacerbations, then permission is there. You can ignore the other rows, and you say, okay, what percentage of my patients have --

MS. PACE: But that is the system.

DR. O'CONNOR: That is the drillable-down data, yes.

MS. PACE: And NQF right now doesn't have a category of endorsing that kind of system analysis, and maybe that is something we need to think about. But what we would be endorsing is this measure that say, what percent of my COPD patients have any of 50 --

DR. RASTOGI: Yes.

MS. PACE: -- PACs?

DR. O'CONNOR: Because the next

question is --

DR. NEFF: Yes, what is the meaning of that?

MS. PACE: Not even of my patients, but of this patient population.

DR. RASTOGI: That is right.

MS. WINKLER: Is there value in that information being measured, say, for a health plan, a health system, a large medical group, in terms of information that can be used to help look at that system in terms of the things we can hypothesize could improve some of these potentially avoidable conditions, such as better care coordination, less fragmented care, that sort of thing, even though you are working at a high level?

DR. O'CONNOR: I am not sure that the drillable data would be of any value. Because in my health system, and we've got 400 patients, I mean 400 physicians, and we take care of, roughly, 180,000 patients. We are looking at 8,000 or 9,000 COPD patients. So, once you do the cuts, we are probably down to 2,000 patients, 2500. Then you are looking at PAC versus typical. I don't know how we could go find the data.

MS. WINKLER: Would it be valuable information for you and your whole organization to know that the most common things that come up on this list are whatever they are for you? You know, one, two, three.

DR. O'CONNOR: Sure. If you had

COPD-associated "X", rather than COPD-associated A through Z. Because the current, the way it is constructed, you are going to come up with a number, 87 percent.

MS. WINKLER: Right.

DR. O'CONNOR: I don't know how you do that. But if you tell me that in this measure when we apply this data, a system with good outcomes has only 19 percent of its patients with complication "X", whereas, one that is poorly managed has 57 percent of its patients with complication "X", that is something you can work with.

But to have a number not associated with something you can do anything about makes it difficult to understand how it is going to be used. Do you follow what I am saying?

MS. WINKLER: Yes, I do. So you are talking about having the aggregate complication versus having a discrete level --

DR. O'CONNOR: I think the work that has been done is phenomenal. It is incredible. I mean it is a dataset that is just absolutely golden. But rather than hide it with a total percentage, what you would like to know is, when they did the COPD analysis and then looked at 11 different plans, and they found that the most common associated complication was, and it varied by 3X between plan 1 and plan 11, that would be incredibly important information to have.

But to say that it varied between

74 and 90 percent, and we are talking a 16 percent difference between the high and low, everybody has --

DR. MILLARD: Although to be the devil's advocate for our colleague, he said all you needed to do was say mortality and everybody starts saying, "Oh, wow, this is important. Why?"

And you could say all you have to say is PACs are high; that's bad. Then it is up to the healthcare plan to say, why? What is the difference in the analogy between the ICU model and the mortality and the primary care? Because in both cases you have to drill down and get the data. Is there a difference, really, when you think --

DR. O'CONNOR: Sure, there is because in this complication model, if you are just looking at numbers, they are between 76 percent and 81 percent. I mean nobody is going to care.

MS. WINKLER: Let me go back to your statement about, is that a characteristic, or your question, is that a characteristic of outcome measures? Because, frankly, when you look at an outcome measure, it is the endpoint of a whole bunch of things that did or didn't happen along the way.

In order to really act on it and move it and change it, you really have to do some background analysis to figure out, what are all the contributions to get you to that

outcome? So is that really a characteristic of outcome measures compared to process measures, where, you know, only half your patients got beta blockers? It is pretty easy to figure out what to do.

Whereas, an outcome measure, is that just inherently part of the deal, is it is a conclusion? And you look at it and say, "Hmmm, what does this tell me?" And you will need to do some back analysis to really understand all the factors that are contributing to that. But is that something that is so much specific to this measure or the ICU mortality measure? Or is that a characteristic of outcome measures in general?

DR. MILLARD: I think that is -- I would agree. The difference may be that this data, I don't know as much what it means. Mortality, we know that that means.

DR. NEFF: And in some ways, mortality, you didn't even totally drill down to find what you think was your problem at your hospital. If you, then, just had some new intervention that maybe wasn't even something that you figured out was the cause, you could then see its effect, even without necessarily knowing that that was what you were trying to change. Do you know what I mean? It is something you can track, if nothing else. But, yes, it gives you an answer, but it doesn't tell you the why.

DR. O'CONNOR: To answer your

question, I think you are absolutely right.

DR. NEFF: Yes.

DR. O'CONNOR: It is inherent within an outcome measure --

MS. PACE: Right.

DR. O'CONNOR: -- that has wrapped up into a conclusion.

MS. PACE: Right.

DR. O'CONNOR: If there is something wrong with one particular health system compared to every -- it doesn't tell you what is wrong.

MS. PACE: Right.

DR. O'CONNOR: It just tells you that there -- it is sort of like a sed rate. You know, a sed rate of 95. I don't know what the hell is causing it. I just know there's something wrong.

MS. PACE: Right.

MS. WINKLER: Well, that is what I was wondering when you started talking about this actionability aspect of it. I think the actionability around using outcome measures is a little more complicated, a little different, because it requires a localized analysis of what you think, or the literature suggests, are the likely contributors and the factors leading to it.

DR. O'CONNOR: There is a credibility factor because let's suppose you put this up on your website for public commentary. I can't imagine the public

comments you get about listing 53 PACs for pediatric asthma. I mean every pediatric asthma doc in the country is going to come out of his or her shoes saying, "What are you talking about? This is a 3-year-old child. You've got DBTs."

So, to make it credible, you would have to make it relevant and use the items that were mentioned as being relevant for that particular condition.

MS. PACE: That was one of the discussions in the other TAP, is that, once you have done this analysis and identified what are the most important drivers of complications, then make the measure around that, which becomes more understandable, more verifiable.

I mean that was just a comment. I don't know whether is the way to go, but that parallels what you are just saying there.

DR. O'CONNOR: Because getting buy-in from pediatric asthma docs is going to be difficult.

MS. WINKLER: We are kind of at the end of this conversation, but, essentially, you have all looked at the asthma and the pediatric asthma --

DR. O'CONNOR: Yes, everything we just said applies to --

MS. WINKLER: Well, that is exactly what I was going to say. Is that the case? Was there anything, in addition, that was

specific to asthma or the pediatric asthma you want to say over and above what we have already said about this kind of measure in general?

DR. O'CONNOR: Just one or two minor comments really.

Let's see, what section? This would be 2b, the reliability testing section. You have done it on 11 datasets; 10,500 children with asthma were included in this analysis. And the PAC complication rates range of 47 to 79 percent.

I don't think there is a pediatric asthma doctor that is going to believe that 47 percent of healthy children with asthma have a complication like this. It just doesn't rise to the level of believability. I don't understand the data.

I would have to see specific information to understand why between 50 and 80 percent of the kids with asthma along to the potential avoidable complications.

DR. RASTOGI: And then you can see that same example sheet that we submitted and the last two tabs, you know, the percentage of PACs and --

DR. O'CONNOR: My point is that the only things that I read in the PAC that seemed relevant which were curious potential avoidable complications of hospital or ER visits related to asthma or acute exacerbation of asthma, even those two wouldn't account for

50 to 80 percent of the kids with asthma. Our hospital, the ER rates are dramatically lower than that.

And looking at everything else in here, they don't all apply to children. So I have a hard time understanding what --

MS. PACE: Which measure? What number is that one?

DR. O'CONNOR: The pediatric asthma one.

MS. PACE: Twenty-one. Do you want to open that Excel file that she was referring to?

MS. FORMAN: It's up.

MS. WINKLER: Oh, okay. Sorry.

Donald, the screen. I am just wondering, are we getting glare?

DONALD: Yes. See, there's some right in front of the camera?

MS. WINKLER: Yes, that first one, if we could just --

So which tabs should we look at?

DR. O'CONNOR: It says 23 percent of the kids were felt to have a mental or behavioral illness.

DR. RASTOGI: This is the admissions.

DR. O'CONNOR: Is this the entire list?

DR. RASTOGI: Right. So, if you scroll up -- oh, maybe it is at the top. Is it? Right, yes.

Okay. So we can see that, you know, this is the stay part of it, you know, what percentage of stays --

MS. PACE: Hospital stay you mean or --

DR. RASTOGI: Hospital stays, right, for various conditions.

DR. O'CONNOR: So these are hospital-associated complications, not necessarily ambulatory?

DR. RASTOGI: So patients with asthma who were admitted to the hospital, the principal diagnosis on the hospital stay has been identified here.

DR. O'CONNOR: That not really a dataset defined by the measure, though.

MS. WINKLER: So I guess that is the question. The peds asthma is keyed off of patients that were hospitalized. Yes?

DR. RASTOGI: The peds asthma is the whole episode. Right?

MS. WINKLER: Okay.

DR. RASTOGI: And if they had, if a pediatric patient had a treatment for asthma, it is a pediatric asthma as this one starts.

MS. WINKLER: Right.

DR. RASTOGI: Then during that one year, if they get hospitalized, then all the hospitalizations are aggregated here, and we are looking at the principal diagnosis for those hospitalizations.

DR. O'CONNOR: So, if a child had

an episode of asthma and was never admitted to the hospital, they would never appear on your dataset?

DR. RASTOGI: They won't be in this tab, but in the next tab, which is the professional tab, you would see the potentially avoidable complications.

DR. O'CONNOR: If a child was admitted during this year for what appeared to be a dehydration episode because you have electrolyte disturbances in 5 percent of the kids, that would appear in this?

DR. RASTOGI: Yes. So, if they did an exclusion, and they were --

DR. O'CONNOR: But point is that a child, you know, a 4-year-old who gets a virus, is vomiting with diarrhea, has to be admitted to the hospital for dehydration. How is that a potentially avoidable complication in an asthma population? That is a childhood illness, and that is a routine childhood illness. We see it all the time.

This is for the professional charges for ambulatory?

DR. RASTOGI: Right. So these are the professional claims. Then, here you can see what were the top drivers for --

DR. O'CONNOR: Adverse effects of drugs, 30, 29 percent?

And wound care, splints, and ostomy, 20 percent?

I don't know how this is 30 to 40

of asthma patients --

DR. RASTOGI: So, when you think in terms of the asthma patient, and you see can these conditions be avoidable, and when you are thinking of the entire patient in a medical whole kind of idea, now you want to avoid many of these problems.

DR. O'CONNOR: I guess I am more of an English language purist in the sense that, if you tell me this is a potentially avoidable complication, I get the sense that I, as a physician, have done something incorrect that I should correct to prevent this complication from occurring again. I don't see that this data leads me in that direction. I think it has been mislabeled, basically.

DR. NEFF: And also, just from a purist perspective, respiratory failure, I mean we are getting into these issues with coding as well. If we call it acute respiratory failure, it bills one way. And if we call it pulmonary insufficiency, it bills the other. But, from an AHRQ perspective, it is a big difference. One is a bad thing post-op; the other isn't. And I don't know what those mean.

DR. O'CONNOR: And it is even worse here because these are ambulatory charges.

DR. NEFF: Yes.

DR. O'CONNOR: Respiratory failure in an office?

DR. NEFF: Maybe that has got some

whacky code.

DR. O'CONNOR: Yes, it's got to be a quirk.

DR. RASTOGI: Now these are professional visits.

DR. NEFF: Yes.

DR. RASTOGI: So it could be inpatient or outpatient professional. Yes, these are professional services.

DR. O'CONNOR: But go back to the previous slide where you had inpatient.

DR. RASTOGI: That is this one.

DR. O'CONNOR: The previous tab.

You have respiratory failure here as well.

DR. RASTOGI: Yes.

DR. O'CONNOR: So 20 percent of the charges on the next page come from the hospital. That means 20 percent of the charges come from an ambulatory setting for respiratory insufficiency or respiratory failure?

DR. RASTOGI: Look at the end.

DR. O'CONNOR: Yes, I know.

DR. RASTOGI: Okay. Over here in this 149 and in the next tab in the professional, it was in the 6,000.

DR. O'CONNOR: So that makes it even -- 8,800 --

DR. RASTOGI: Yes.

DR. O'CONNOR: That makes it even worse. How did 8,700 children get coded as

respiratory failure?

DR. RASTOGI: Now this is number of occurrences. Okay? It is not number of patients.

DR. O'CONNOR: Number of occurrences.

MS. PACE: Of the PAC?

DR. O'CONNOR: Have you ever coded respiratory failure in your office?

MS. PACE: Well, this is also professional visits to the hospital.

DR. O'CONNOR: Yes, but that is all they have in this --

MS. PACE: Oh, 140 --

DR. MILLARD: Although if they were multiple, is this daily charges?

DR. O'CONNOR: No, that would be on the hospital side of the charges.

DR. RASTOGI: Yes. The costs are associated with the professional bills.

DR. O'CONNOR: This probably just illustrates the point I am trying to make, is that the credibility factor needs to be addressed.

DR. RASTOGI: It is amazing that the data is this way because it is. And people usually, you know, when physicians see it for the first time, they jump out of their skins, too. They go back and they look, and they say, why is it happening so much?

DR. O'CONNOR: I think what you need to do is to plow the ground and plant

some seeds in a sense that this data needs to go through a peer-review process to become credible before --

DR. RASTOGI: Yes, you know, it has been vetted by several physicians. But, yes, on the NQF side, too, it would be nice if you could, yes.

MS. PACE: Well, I think I understand your development process involved expert panels and things. I guess the question that comes up, or at least in these prior measures, is, you know, how reproducible is that one small group? And I don't know what the numbers come to in this. In those prior measures, you know, they even reported, "We had this group of physicians look at 10,000 ICD-9 code pairs." Then, just from a logistical and people start saying, how did six physicians look at 10,000 pairs of diagnoses and arrive at this?

I mean it just starts raising lots of questions. I don't know what the answer is or if there is an answer.

DR. RASTOGI: In these kinds of outputs, some of the tables are available for all the 11 different databases we ran the data on. So we can provide those to you, too, if you wanted to look at that.

And like I mentioned, this is version 1.0. Now, based on version 3, that we have the latest coding, the latest outputs are just coming out.

But the validity testing that they have done --

MS. PACE: Right, because this PAC is actually, you know, a string of diagnoses codes --

DR. RASTOGI: Exactly.

MS. PACE: -- that go into forming that PAC.

DR. RASTOGI: Exactly.

MS. PACE: And it just this becomes this kind of exponential number of things that --

DR. RASTOGI: Yes, it is --

MS. PACE: -- a group has done in a computer algorithm. It is something that, as people start looking at the measure, have difficulty actually kind of comprehending --

DR. RASTOGI: Yes.

MS. PACE: -- when you start talking about that. I mean I think --

DR. RASTOGI: That is exactly --

MS. PACE: -- that is part of the disconnect of dealing with people instead of computers.

DR. RASTOGI: Yes, exactly.

DR. O'CONNOR: I guess part of it, too, under 1c, part of the justification here is that, if properly managed, these avoidable complications -- well, I am not sure that these are avoidable complications. Some of these are just merely listing some routine childhood illnesses.

There seems to be a message that is being sent that probably ought to be retrieved --

MS. PACE: And you know, that is kind of two different philosophies of how to view these kinds of measures. One is just identify complications and measure it and risk-adjust, and then look at, are there some people, some providers that actually have fewer numbers, and what are they doing? And not try to do the value judgment of what is avoidable.

Or the other kind of philosophy is, no, we only want to measure it if it is absolutely avoidable.

But I think this is kind of a mixture of what you are --

DR. O'CONNOR: But from NQF's perspective, I would think if you are going to approve a measure, that people who are going to come to you to use the measure want some assurance that, in fact, there is a benchmark that they can compare themselves to and change some stuff and improve. And I am not sure that that is going to occur here, the way they've got it all wrapped up into one grand number.

But that is an application issue. The data is fascinating. It is incredible.

All right?

MS. WINKLER: So is there anything more to say? Or have we said it?

DR. O'CONNOR: No, I think we've said it.

In the beginning this morning, you had mentioned whether or not there might be other ideas for other measures at some point in the future.

MS. WINKLER: Uh-hum.

DR. O'CONNOR: Let me just put one on the table. It is the elephant in the room that nobody ever talks about. That is compliance.

MS. WINKLER: Uh-hum.

DR. O'CONNOR: How often patients, why patients with asthma are so very different than patients with, say, diabetes. Under the best of circumstances, the national data would suggest that the refill rate for preventive medications runs between three and four units per patient per year. That is about a 75 percent non-compliance rate.

MS. WINKLER: Yes.

DR. O'CONNOR: If you do frequency distribution analysis on the users, you will find out that less than a third of your patients are actually using more than six units per patient per year. And they tend to drive and bring up the ones that are only using it once or twice a year.

That is the major issue facing us in asthma care in the United States. Everybody seems to know about it, but there really isn't any movement that I can detect --

MS. WINKLER: Yes, we sure have. Yes, management came through with several measures. NCQA particularly has several adherence/compliance, compliance/adherence, pick your name, measures around use of medications, suboptimal use of medication, appropriate use of rescue medication, things like that.

So those measures exist. To the degree that they get implemented, get used to make any changes, I think we are still in the early stages of that. So there are some.

But I guess some of the questions around these conditions here being asthma on types of outcomes --

DR. O'CONNOR: I don't know what the refill rate is for COPD. I just know for asthma.

MS. PACE: Yes, it is pretty bad.

MS. WINKLER: Yes. Well, hopefully, we are going to be getting some more data as some of these measures get implemented more and more.

But I think in terms of outcome measures for asthma, I mean this is more sort of a negative side. What are positive sides? What is a good outcome for a patient with asthma, you know, for a kid or an adult, as a result of appropriate, effective treatment? You know, is it functional, and they do what they need to do? Impact on your life, all these --

DR. O'CONNOR: Well, according to Dr. Millard, we can stop worrying about school attendance.

DR. MILLARD: If you live in Dallas, Texas.

(Laughter.)

MS. WINKLER: If you live in Dallas, Texas. What? Does nobody go to school?

(Laughter.)

DR. O'CONNOR: Nobody goes to school. So it doesn't make any difference.

(Laughter.)

DR. MILLARD: No, my read is that there was such a push for attendance data, that somewhere in the school administration somebody cooks the data on school attendance because they are paid on attendance. When 97 percent of kids in an urban school district are counted as present on a daily basis[1] --

[1] Please note the following clarifying statement about this comment from Dr. Millard:

"In the process of responding to a colleague's comment about our recently published study in CHEST that notes children with asthma do not appear to miss more school than non-asthmatic classmates, I began the discussion by raising the question as to whether the reported data was "cooked" by school administrators, referring to the economic incentive to report high levels of school attendance for state funding purposes. At the point that I was going to refute that charge with arguments to the contrary, the conversation changed to a different path and I was unable to finish my entire thought. As such, therefore, what is recorded, is not an accurate reflection of my sentiments, and represents, indeed, the opposite opinion from what I was intending to state and what I believe

DR. O'CONNOR: It makes you wonder.

DR. MILLARD: I'm sorry, you know, kids in Texas are no different than kids in California or --

DR. O'CONNOR: Kids are kids.

MS. WINKLER: Well, I think the one thing that I think would be helpful, because we are going to need to have at least one, if not two, follow-up conference calls to kind of give yourself a chance to think through some of this, what we have talked today.

We are going capture all this and put into a single form and let you all look at it, to be sure we have reflected what you have said.

But, also, I think a lot of this is kind of tough, complicated stuff, and there is an opportunity to reflect. And we do have the time to do that.

In some of the materials we sent you upfront was, on this topic, the very end of sort of the briefing memo, was what the Steering Committee as sort of a bit of framework of types of outcome measures. One of them was functional status. One was symptom control. One was -- shoot, I can't even remember now. My brain is fried.

But the idea of those are potential types of outcome measures, and not all will

---

to be true: that children with asthma, when properly cared for, do not have to miss more school than their non-asthmatic peers."

apply to all different conditions. But perhaps these become the ideas of good outcome measures, information that actually will be usable to a wide variety of stakeholders, something that gives us something more to work with, to say, you know, how is it going? How good is it? Can we do better? And what are the things that are particularly important?

Yes, I was going to say there they are. Patient functions, symptoms, quality of life. We saw something about it today. Intermediate clinical outcomes. You know, you see that much more readily in something like diabetes or blood pressure control, or something like that. It may not be as applicable here.

But experience of care or caregivers. But knowledge, understanding, behaviors. There's where your adherence/compliance comes in as an outcome measure.

Healthcare service utilization. This is the ER visit for asthma or the hospitalization or the re-admission, or something like that.

The clinical morbidity, aside from mortality, related to disease control and treatment. And the classic example is amputations in diabetics. You know, you get something really dreadful because you just weren't taking care of business.

Then safe and healthy living

environments. Adverse events, clearly, you know, we have talked about that at length today, as well as mortality.

If there are any others that you can think of? But in terms of focusing on these, a variety of ways of describing the outcome for patients with COPD and patients with asthma --

DR. MILLARD: Well, the big asthma metric that is supplanted rules, too, in terms of -- because you gave a number, the asthma control test, because at least that has a number. It has a number associated with it.

DR. O'CONNOR: And I don't know how it is in other parts of the country, but we do an ACT on every patient on every visit.

DR. MILLARD: The ACT in the primary care world has been dead on arrival. Nobody does it.

MS. PACE: And what does that entail?

DR. O'CONNOR: It is five questions.

DR. MILLARD: Like a scale and it is added up and --

DR. O'CONNOR: A score of 19 or less suggests there are issues of control.

MS. PACE: These are all --

DR. O'CONNOR: "Over the last four weeks, I walked in and he had a nighttime disturbance" --

MS. WINKLER: But the patient tells

you what is going on. You are asking the patient.

DR. O'CONNOR: It is a simple number.

DR. NEFF: They are saying that you wouldn't otherwise extract just by a chat.

MS. WINKLER: That is a good intermediate control. Really, that sounds like sort of your intermediate outcomish -- although it may change not in a linear fashion, but may go up and down.

DR. MILLARD: Well, depending on the exacerbation. That is now, when you look at the asthma guidelines, that is some metric of -- that really does combine a lot of the metrics of asthma control. Really, the only thing it lacks is some sort of objective measurement.

MS. WINKLER: One of the interesting things that we talked about in the Steering Committee was in terms of particularly things like function, is where the data comes from.

You can get data from the patient, either through a structured questionnaire or they report, the patient or family reports and tells you about it, as opposed to, you know, clinician observation of. I mean you could tell, if you have to take care of them in the ER, you know, every week, you are going to get a similar sort of answer, but it is of inadequate control, but it is going to be of a

different type of data.

And there was a large support for using these patient-reported data as a really good source of these kinds of questionnaires, particularly when they can be embedded into your medical record. It could be embedded into your EHR. You know, you just ask the questions, check the box, and calculate your number, and that is recorded onward.

DR. O'CONNOR: That is a challenge right now because we have an EHR. Currently, we are scanning ACTs in, but we have talked to the people about doing -- who talked about a flowchart?

MS. PACE: I was asking about that.

DR. O'CONNOR: The ACT is perfect for a flowchart, and that would be just ideal because we could do the ACT when they come in, and the nurse could just simply put the number in a box on the flowchart, and it is part of the medical record.

MS. WINKLER: Right.

DR. O'CONNOR: It is retrievable and it is there for --

MS. PACE: Right, exactly. It would be in a field.

MS. WINKLER: See, I like that because people keep asking about, oh, well, we can't get patient-reported data; it is too expensive, too burdensome to do questionnaires.

Wait a minute. Wait a minute.

That is what history is, you know.

MS. PACE: It is in a structured format.

DR. MILLARD: But it key is to direct it. I mean it is very simple because asthma lends itself to very focused --

MS. WINKLER: But it is standardized. It is structured.

DR. MILLARD: And it has been validated.

MS. WINKLER: And it lends itself to individual goals.

DR. O'CONNOR: And there's a pediatric version as well.

MS. PACE: But, I mean, it would be similar to, you know, the measures you looked at first today, the health-related quality of life and the six-minute walk.

DR. O'CONNOR: Exercise tolerance, yes.

MS. WINKLER: Yes. So, I mean, it would be in that vein, and you could look at changes or you could just look at levels. I don't know how it would be best to construct the measure, but would be a great -- but those are the kinds of things that there actually is to think about in posing, what would be good outcome measures for asthma?

I mean there was a big emphasis on COPD today, which, of course, HHS will like very much. But asthma is still a huge issue for younger populations, and we just don't

have quite as much in that arena, it would seem. So trying to get a handle on that of where we would like to go, learn more about the outcomes around asthma treatment for both children and adults.

DR. MILLARD: You would get a lot of money from pharma if you put the ACT out there because that is what drives, I mean that is what drives prescriptions, is bad asthma control.

DR. O'CONNOR: And it turns out that, I mean quite truthfully, I will see a patient in the office and their pulmonary function could be perfectly normal, but their control is just absolutely dreadful. PFTs are a poor positive predictive value; for normal PFT, it is really very low. I mean, yes, if the PFTs are bad, you've got to pin your ears back because there is something really wrong. But you can have an out-of-control asthma patient who has a normal blood test. That is the issue.

That is where something like the ACT comes in because it absolutely does give you a different objective point. PFTs, there is no reason why soon they shouldn't be all retrievable in an EHR as well.

MS. WINKLER: Good.

Margaret?

DR. NEFF: Theoretically.

MS. PACE: You're saying that that wouldn't really give you much indication of

quality of control, by looking at pulmonary function?

DR. O'CONNOR: No. I will do PFTs on patients when they come in the first time, especially if their persability needs to be demonstrated. But, after that, I would rather rely on the clinical history and an ACT because I have too often been burned by a patient what appears to be normal pulmonary function who is actually doing quite poorly.

DR. MILLARD: Although the critique of the ACT and clinical assessment is you can't guess lung function. In the adult population, it may be the reverse thing. Because we are doing a study right now with a control of breathing, a non-medical intervention to try to downregulate asthma symptoms. And I get to see all these people that I have never seen before who have asthma diagnosis and guess their lung function before they have their methacholine challenge. I am a terrible clinician. I can't guess their lung function to safe my life.

I mean because adults at least get used to having low lung function, and you can say this person doesn't have any symptoms at all. Are they going to qualify? And you look at their PFT --

DR. O'CONNOR: It scares you. It scares you.

DR. MILLARD: It scares me to death. Now it is smokers.

MS. WINKLER: Yes.

DR. MILLARD: And that has been sort of the reason why --

DR. O'CONNOR: That is amply stated, and it is actually quite solid information, too.

DR. MILLARD: Maybe a lot of us aren't as effective anymore.

DR. O'CONNOR: Okay. The other thing is that oftentimes you have to do -- it is a gestalt, a clinical history, a physical exam, and ACT, and PFT, and exhaled nitric oxide, and you put everything together then, and come up with some --

MS. WINKLER: Margaret, in terms of the intensive care unit, I mean it looks like we are moving towards something for mortality, which of course is the big one there, but are there are some other things that, thinking about outcomes in a broader perspective?

DR. NEFF: Well, I mean I think you brought up a little bit sort of healthcare utilization, sort of recidivism kind of comes to mind a lot in the concept of, whether it is ICU bouncebacks, which is a big deal. People coming out of the ICU are coming back within 24 hours. And the same thing could be said for the ED, and we know the ones that come back two or three times are the ones that are the worst off. It is all kind of part of the same process.

There is also sort of that

flowthrough. So it is kind of a triaging flow of patient utilization of resources in a way that actually gets them where they need to go and not letting go of them too soon or too late.

So it kind of works in a little bit with that ICU length of stay, but the bouncebacks, I would say, are probably pretty high on our radar right now because they are probably -- there is something modifiable in there. I think you could expand that to other venues within the hospital pretty easily as well.

I don't know how that would morph into an outcome, but it is definitely sort of an issue that we are trying to find sort of process that improves outcomes. So it is kind of they sort of have to all sort of link together.

MS. WINKLER: All right. Is everybody pretty much tired out?

Thank you, guys.

(Whereupon, at 3:30 p.m., the proceedings in the above-entitled matter were adjourned.)