# NATIONAL QUALITY FORUM

+ + + + +

## IMPROVING DIAGNOSTIC ACCURACY 2016-2017

+ + + + +

**WEDNESDAY**
**APRIL 12, 2017**

+ + + + +

The Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 9:00 a.m., Mark Graber and Kathy McDonald, Co-Chairs, presiding.

**PRESENT:**

MARK GRABER, MD, FACP, Co-Chair; Society to Improve Diagnosis in Medicine, RTI International Plymouth
KATHRYN MCDONALD, PhD, Co-Chair; Center for Health Policy and Center for Primary Care and Outcomes Research
JENNIFER CAMPISANO, JD, Booby and the Beast Blog Phoenix
MICHAEL DUNNE, PhD, bioMerieux, Inc.
DAVID GRENACHE, PhD, University of Utah
MARILYN HRAVNAK, RN, PhD, ACNP-BC, FCCM, FAAN, University of Pittsburgh
NICHOLAS KUZMA, MD, St. Christopher's Hospital for Children
PRASHANT MAHAJAN, MD, MPH, MBA, University of Michigan
LAVINIA MIDDLETON, MD, The University of Texas MD Anderson Cancer Center
DAVID E. NEWMAN-TOKER, MD, PhD, Johns Hopkins University School of Medicine
MARTHA RADFORD, MD, MA, NYU Langone Medical Center

**DAVID SEIDENWURM, MD, Sutter Health**
**THOMAS SEQUIST, MD, Partners Healthcare System**
**SUSAN SHERIDAN, MIM, MBA, DHL, Patient-Centered Outcomes Research Institute**
**HARDEEP SINGH, MD, MPH, Veterans Affairs Center of Innovation and Baylor College of Medicine**

**NQF STAFF:**

**HELEN BURSTIN, MD, MPH, Chief Scientific Officer**
**MARCIA WILSON, PhD, MBA, Senior Vice President, Quality Measurement**
**JOHN BERNOT, MD, Senior Director**
**TRACY LUSTIG, DPM, MPH, Senior Director**
**ANDREW LYZENGA, Senior Director**
**VANESSA MOY, Project Manager**
**CHRISTY SKIPPER, Project Manager**

**ALSO PRESENT:**

**PAUL EPNER, Executive Vice President, Society to Improve Diagnosis in Medicine**

**KERM HENRIKSEN, Patient Safety Program Officer, Agency for Healthcare Research and Quality (AHRQ)**

**DIVVY UPADHYAY, Senior Research Associate, Social & Scientific Systems**

CONTENTS

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

(9:04 a.m.)

MR. LYZENGA: All right. I think we're going to get started. Thanks, everybody, for coming out for this meeting. We really appreciate your attendance, taking the time.

We've got a busy couple of days ahead of us, I think. We hope to make this a working meeting, sort of dig into the long list of concepts we have. And work through them gradually, and in a sort of iterative process. And come up with a smaller list of prioritized measures. We think we've come up with a pretty good approach for doing that.

I should maybe start out with an apology of sorts. We had, you know, certainly anticipated, and hoped to get you materials to review much earlier, and in advance of this meeting.

Sort of had a plan worked out. And then ended up going back to the drawing board a little bit, late in the game, and rethinking our

approach and the materials we would need.

So, just wanted to acknowledge that we sort of did you a disservice by not giving you more time to review these things in advance. But that said, I think we'll have time to work through these things today and tomorrow. And we look forward to a productive meeting.

Mark, did you have any welcoming remarks, or anything like that you wanted to --

CO-CHAIR GARBER: Good morning, everybody. Thanks so much for being here again. So, I always knew that the NQF did miraculous things. And there's evidence right here today.

We left with I think 40 measures. And we've come back and there are like 200. So, it's like the miracle of the loaves, or the oil, or something.

So, as you all know, I think, the chairs do almost nothing. It's the NQF staff who really do all the heavy lifting behind the scenes. So, huge thanks to Andrew and John, and everybody at the NQF, for somehow massaging all

this, and getting it to the point where we can discuss it today.

Missy Danforth had a minor emergency at Leapfrog, and wasn't able to be with us. So, huge thanks to Kathy for pinch hitting as co-chair of the meeting. And I'll turn it over to her for comments, if she'd like to make some.

CO-CHAIR MCDONALD: Yes. So, put up with me basically. Yes. Looking forward to a good couple of days with everybody, and lots of participation that's, you know, pretty much baked into this.

We're not going to be able to get away without really engaging. And I think it will be very useful for all of us to do that. So, good to be here. And now, I think we do introductions. Is that right?

MR. LYZENGA: Maybe we could just do another quick round of re-introductions for everybody in the room. Go this way? Just a quick introduction.

MEMBER SHERIDAN: Good morning. I'm

Sue Sheriden. I'm the Patient Family Advisor at CMS.

MEMBER SINGH: Hi. Hardeep Singh, Patient Safety Researcher at Baylor College of Medicine and Houston VA.

MEMBER KUZMA: Nick Kuzma. I'm a pediatrician in Philadelphia.

MEMBER HRAVNAK: Marilyn Hravnak. I'm a nursing faculty at the University of Pittsburgh. I lead our PhD program. But I'm also a nurse practitioner in the ICU.

MEMBER RADFORD: Martha Radford. I'm Chief Quality Officer at NYU Langone Medical Center in New York City.

MEMBER NEWMAN-TOKER: David Newman-Toker, Johns Hopkins Health Services Researcher on Diagnostic Safety and Quality.

MEMBER MIDDLETON: Lavinia Middleton, MD Anderson Cancer Center. Director of Quality Operations and Deputy Chief Medical Officer.

MEMBER SEQUIST: Morning. I'm Tom Sequist. I'm the Chief Quality and Safety

Officer at Partners Healthcare.

MR. HENRIKSEN: Kerm Henriksen, Patient Safety Program Officer at AHRQ.

MEMBER DUNNE: Mike Dunne, Senior Fellow, bioMerieux Clinical Microbiologies.

MEMBER GRENACHE: I'm David Grenache, Professor of Pathology at the University of Utah, and Medical Director of a clinical chemistry lab at ARUP Laboratories.

MEMBER MAHAJAN: Prashant Mahajan, Vice Chair of Emergency Medicine, and a pediatric emergency physician at University of Michigan Ann Arbor.

MEMBER HUNT: I'm David Hunt. I'm the third David so far. I'm a medical officer, and Medical Director for Patient Safety at the Office of the National Coordinator.

MEMBER SEIDENWURM: David Seidenwurm, fourth and final David. And I'm the chair of the Quality Committee at Sutter Medical Group, and a neuroradiologist.

MEMBER CAMPISANO: I'm Jen Campisano.

I'm a patient advocate.

MS. LUSTIG: I'm Tracy Lustig. I'm part of the NQF staff.

DR. BURSTIN: Good morning, everybody. Welcome back. Helen Burstin, Chief Scientific Officer. And especially pleased to work through the CMS issues, and able to join us today.

DR. BERNOT: John Bernot, also part of the NQF staff here.

MR. LYZENGA: Andrew Lyzenga, part of the NQF staff.

MS. SKIPPER: Good morning. Welcome back everyone. Christy Skipper, project manager.

MS. MOY: Good morning, everyone. My name's Vanessa Moy. I'm also an NQF staff.

CO-CHAIR MCDONALD: Do we have any committee members on the phone?

(Off microphone comment.)

CO-CHAIR MCDONALD: Oh. Any committee members on the phone? No? No committee members on the phone? Okay. And then, guests? Paul. Yes.

MR. EPNER: Paul Epner, Executive Vice President, Society to Improve Diagnosis in Medicine.

CO-CHAIR MCDONALD: Divvy? Yes.

(Off microphone comment.)

MR. UPADHYAY: Morning. I am Divvy Upadhyay. I'm a Senior Research Associate at Social and Scientific Systems. Thank you.

MS. SKIPPER: And I'm sorry. And this is Marcia Wilson, also part of the NQF staff.

CO-CHAIR MCDONALD: Okay. Welcome, everybody.

MR. LYZENGA: So, I think maybe we can just jump in. John, you want to talk a little bit about some, sort of thoughts and potential revisions we've, maybe proposing to our measurement framework here?

DR. BERNOT: Yes. Yes. Thank you very much. And again, welcome, everyone. Thank you so much. As Mark said, it's absolutely amazing the involvement that we've had, both from public comments, to participation in between

sessions. I don't think any of us expected that we would crest 200 measure concepts here in a short period of time.

So, we did have the pleasure of going through all of these concepts. And trying to put this into something that made sense, in a way that we can now analyze it, and hopefully come out of today with some priorities. And maybe even some ratings in different categories as to where the measures stand. And we'll talk about that. Can you go to the next slide? Maybe the next one after that then.

Okay. So, this is how we left it at the meeting last time. We said we had, we came up with three domains, structure, process and outcome domains. And within those a number of sub-domains that we'll talk about in just a few moments here.

And as we went through these different concepts we came up, and they really clustered into three pretty big concept areas. Not surprisingly, not totally different. But around

the organization, around the diagnostic process, and around the patients and the caregiver.

And the reason I point this out is for a few reasons. One, within those though there was a lot of crosscutting of measure types. So, within the organization we're seeing structure and process measures. Within the patient and caregiver we're seeing structure, or process and outcome measures.

So, it got us to thinking, do we really have things laid out the right way? Or are we maybe looking at it from the other side? And the other things is, as we were talking internally, when we come up with a framework it's our feeling that we should have the domains as something important.

We're saying, this is a project. Is it really important that we have a structure measure? Maybe, maybe not. But is it really important we have an organizational measure? Absolutely. Is it important that we have a patient measure? Absolutely.

So, when we saw this large volume of data it clustered a little bit differently. And, Vanessa, can you go to the next slide?

So, what we were thinking of doing is actually moving the domains over, the structure process outcome, and actually making it more of a measure type. And you can go to the next slide on here.

And really moving those over and saying, rather than those being the domains, they became a type of measure within the sub-domains. And actually then moving the concepts over to, the concept areas over, and let those become the domain. So, flip flopping these two things on the side of the sub-domain. You can go to the next slide.

And then, this is what it would look like. And this is by no means us trying to change the committee's wishes. I think this is how the data clustered. And we definitely want to give this as a proposal.

By no means is this a directive that

this is the way it should be. But just as we had a lot of time, many, many hours going through this information, it actually seemed to cluster a little better this way, and maybe make a little more sense.

In the process of this also, the sub-domains would need to be tweaked. But largely, the sub-domains still held up. And I'm going to show you how that would look if we did go through with this proposal.

So, just a couple of more slides, and then we'll open it wide open for questions. So, if we did this, we would have three domains, organization, the diagnostic process, the patient and the caregiver, with the sub-domains.

And I will take the time to read them, just because I think it's important. So, in the organization, the external environment, organizations diagnostic, QI activities, patient access, and workforce.

Also, I'm going to show you next now these match back to the old sub-domains. So, you

don't have to try to figure that out in your head.

Diagnostic process, as you can see, very much going in line with the National Academy report, the info gathering, the integration, interpretation, and some other things on efficiency, errors or accuracy, and follow-up.

And lastly, the patient and caregiver, patient engagement and patient experience. So, okay, you can go to the next slide. And this is how they compare. And this will be the last slide. I know I'm throwing a lot of information quickly at you.

But no change in the external environment. Where we saw the organizational features we broke that up into a couple of different components that had clear, clear delineations in the amount of concepts that came in.

Technology and tools was one of these ones that got us thinking about this whole thing. We were seeing technology and tools more so than

just in the structure. And that's where we originally had it.

Technology and tools was being found in different areas, how it interacts with a patient, how it interacted with the process, and things. Does this really fit the way it is? Work, people and workforce, essentially the same thing.

So, no real change in the diagnostic process, except for the expansion. Really, we took that one, and really opened it up into a lot of different ones. And it was truly because of the volume, and the clear separation of the types of concepts that were coming in. In our mind there's a clear separation.

No change to patient engagement. No change to patient experience. And again, the outcomes, sub-domains would be moved in as measure types, into the domains.

So, that is a lot. It is lot. And, can you go to the next slide? Actually, let's go back to, just so people can see. Keep going

back. We can just stop here, so this is available for everyone to see.

And I want to turn it over to the committee, just to get your gut reaction, if this makes sense, or if we want to revert back. Again, everything's fluid, and nothing is set in stone for us whatsoever. I'll turn it over to the chairs.

CO-CHAIR GARBER: Comments from anybody? Please.

MEMBER NEWMAN-TOKER: Yes, I'd like to. The one thing that I found myself tripping over a little bit was the patient and caregiver, and then the sub-domains. Both being about patient engagement and patient experience.

I don't know if there's a provider experience piece of this story that exists out there. But if, I would either, if we don't care about other provider fields, might make that a patient domain, the patient and the caregiver domain.

Oh, did you mean caregiver? You

didn't mean the providers? You meant the patient's family? That was confusing to me. So, I would see, think if there's some way patient and family, or something slightly different. I don't know.

DR. BERNOT: Absolutely.

MEMBER NEWMAN-TOKER: Patient family and caregivers. Or something that would make it more, clearer to me anyway. It wasn't --

DR. BERNOT: That makes sense.

MEMBER NEWMAN-TOKER: -- You didn't mean physicians and nurses?

DR. BERNOT: I did want to comment on the first, about the clinician provider experience in the, we put that in the workforce part. And there actually was some concepts on burnout that people had proposed here.

MEMBER HRAVNAK: I like it. I only had one question, which is that when we came up with those original domains, we must have had a reason for why those seem to be the one to go with.

And is that reason okay to give up? I mean, was that based on how NQF does other things? Or was it based on something nationally? Or is there anything to lose by giving up the original, I guess was my only question.

DR. BERNOT: I'll briefly, but I think I'd rather have the committee make that decision. But there was nothing that we felt that we lost in the process.

We did come up with those as we were looking at the report from the National Academy, and just the way that they had laid out some of their things. So, it, I don't want to say it was arbitrary that we picked those to start. But it was the starting point based on that.

I don't think there was anything more to it, in my recollection. But I really would like to make sure that I'm not missing something.

MR. LYZENGA: I should note that some of those ideas were pulled from our colleague, Hardeep's diagnostic, what's the framework? The diagnostic safety, Safer Dx Framework, which had

sort of broken things down into structure, process, and outcome categories. And that kind of made sense to us, and aligned very well with many of the categories that were in the National Academy's report.

But again, as we reviewed the measure, concepts that would see, it seemed like we were having many of these categories, things like information gathering and documentation, that were cutting across structure, process, and outcome, you had measures reflecting each of those things.

And it kind of started to feel like a bit of an artificial distinction to say that information gathering and documentation was only related to process measures.

But that, you know, you could also have structure measures representing those elements within that, and outcomes related to them. So, that was kind of the thinking.

CO-CHAIR MCDONALD: Hardeep?

MEMBER SINGH: Yes. So, I think

conceptually speaking this is fine. I was thinking, are we missing anything? So, I tried to sort of in my mind do the, you know, when I think of diagnostic errors and how to study the topic.

In general three things come up, system issues, provider/team issues, and patient issues. And I think you've got those covered. You just have titled them differently, which is, you know, works.

You know, a couple of things I'll say. Okay. So, if you -- you mentioned provider burnout a second ago. Where would that go? I think you can still put that in the organizational workforce issues, because a lot of those problems are organizational. And so, I think it is covered.

The other way I was trying to break it down to see have you missed anything. If you sort of go away from the structure process outcome, you know, model that we were using, I think we're still okay.

Because a lot of the structural things are under the organization. A lot of the process things are already covered. And a lot of the outcomes that we're interested in are patient outcomes.

And are we missing anything, like provider outcomes? Maybe not. Because then, as I said, they may be under workforce. So, I think this is okay. No matter what titles we use, I think we've covered it broadly. So, I'm fine with it.

CO-CHAIR MCDONALD: Thank you. Martha?

MEMBER RADFORD: Again, I like it. There was a fair amount of misassignment around structure process outcome in the original list. So, we probably need to kind of tighten that up a little bit.

CO-CHAIR MCDONALD: And the thought was that you'd still do structure, process, and outcome. But sub-domains, so that will still --

MEMBER RADFORD: Matter, right.

(Simultaneous speaking.)

CO-CHAIR MCDONALD: David?

MEMBER HUNT: I feel like a very old dog, just recounting the history of the Donabedian structure, process, outcome. I don't think, you know, and I can share. I always keep a copy of his original, that article on structure, process and outcome, from the Millstone Quarterly.

(Off microphone comment.)

MEMBER HUNT: Huh?

DR. BURSTIN: It's on my desk.

MEMBER HUNT: Oh, okay. But just to speak to it. If we read that and look at this, I think this is wholly concordant with the basic tenets of what he was trying to forward, as far as this is concerned. And I feel like this still has a tremendous amount of rigor and elegance. And I don't think we're losing anything with this.

CO-CHAIR MCDONALD: Thank you. And I actually have to add too here that I think it

nicely reflects the National Academy's model. Because that's got the patients sort of throughout. It's got the process, and the kind of main part within the work system, which is sort of the organization. So, if you had to parse it into three subcategories, this would be aligned with that. Prashant?

MEMBER MAHAJAN: So, to me honestly it didn't make much of a difference, in the sense it was what I thought that the structure, process, outcomes would have captured all these. I think it just reassigned them. But eventually we are doing the same thing.

I mean, what I am anticipating, at least coming out under organization, under the sub-domains, we will eventually have measures related to structure, process, outcome. Maybe not everyone will have everything.

But the only thing I would add is probably try to balance it in such a way that we just shouldn't end up having too many of the process measures or the outcome measures. But

check off each of them a little bit.

CO-CHAIR GARBER: Yes. So, one thing is missing that is in the IoM report, is the impact of the external world. So, these are all measures that are perfect I think for healthcare organizations. But don't take into account the external factors. So, I guess, is it an assumption that we're --

MEMBER SINGH: This is the --

CO-CHAIR GARBER: -- just developing?

MEMBER SINGH: It's the external environment. I checked off for it, and it's there. So it will, it's under organization, which is sort of a little artificial. But it affects the organization.

So, if there's a change in Medicare policy or reimbursement structure, you are mediating this with the organization. So, I think it's okay. I don't know if it deserves a separate category. I thought about that also before I --

(Off microphone comment.)

MEMBER SINGH: The organization as a mediator.

CO-CHAIR GARBER: Is that, John and Andrew, is that where you're thinking external means?

(Off microphone comment.)

CO-CHAIR GARBER: Models?

MR. LYZENGA: Yes, yes.

CO-CHAIR GARBER: Liability reform? Those things? Okay.

MEMBER SINGH: Well, I mean, if you want to be precise you could just, you know, retitle the organization as something else, you know, organizational and policy issues, or external environment issues, or something of the like. If you want to like really be crystal clear, you could do --

MR. LYZENGA: I don't think we're necessarily tied to these names, by any means. If you have a better way of describing things --

MEMBER SINGH: So, the other thing we could do is, as we create the measures, or

measure concepts, see what new things we have, and change the --

MR. LYZENGA: Yes.

MEMBER SINGH: Yes. And change the names accordingly.

MR. LYZENGA: Yes. That's absolutely the idea. Both the domains, and then the sub-domains as we're going through them, consider as we're working through them, are these the right sort of, you know, names for the sub-domains? Or should they be tweaked or modified at all? Are we getting the appropriate buckets, again, to capture the measures and the conceptual areas they're trying to cover?

CO-CHAIR MCDONALD: Is anybody else wanting to comment? David?

MEMBER NEWMAN-TOKER: Just the, so the question of whether the patient access piece, it is something that the organization has obviously a direct impact on. But so does the patient experience. It's directly impacted by the organization.

I wonder whether it belongs in the patient bucket, rather than the organization bucket, the sort of, you know, access, engagement, and experience? Almost feel like they're, that they belong together.

CO-CHAIR MCDONALD: So, I was looking at that too, wondering about that. And then I made an argument for why it should be in organization. But I want to hear what the argument was that was yours.

DR. BERNOT: Yes. So, first of all, this one is one that Andrew and I at the last minute were going back and forth. So, I definitely, I could definitely see that going either way.

The one thing I did with, at the risk of turning everyone's attention away from the discussion, I did want to point out, we put a pretty thorough definition list of exactly what we think fits into each of these.

For example, just based on Mark's question. Not, then again, all this can be

changed. But what we thought external environment, we said policy costs, legal issues around diagnostic quality.

So, you at least, if you look at this, it's the proposed new domains and sub-domains sheet. It really at least lets you have an insight into our thinking. Again, all of that can be changed. But just, so for clarification purposes.

We also tried to include, based on some of those really rich discussions on the last webinar, about where the different communications occur, provider/provider, provider/patient, system/patient. We had a lot of discussion.

And we wanted to make sure that we at least addressed that in our definitions. So, I think I've said enough. But I just wanted to make sure that you had that in front of you.

MR. LYZENGA: I should add, we, just again, much of this emerged from our going through the list many time, of the concepts. And those patient access measures seemed to us

largely to be related to organizational activities, and sort of characteristic, things like ease of getting an appointment, geographic access, you know, that patients receiving sufficient long appointments.

Most of them seemed to kind of be those things that were in the organization's control. So, that was again, and we sort of, you know, speaks to the process by which we did this.

It was a sort of inductively reasoning, I think that's the right word, from the concept list to our sub-domains in some sense, with some input or, you know, considering other things as well. But that was the reason behind it.

MEMBER SINGH: And we're not looking for perfect, you know, mutually exclusive categories, correct? I mean, there could be a little bit of, there's never going to be a perfectly mutually exclusive --

The other thing I was thinking is, for some organizations patient access is huge. I

mean, for VA, probably even Kaiser, and some of the other organizations, it's a huge issue. And it's an organizational issue.

The other thing I was thinking was, which I think maybe David's also referring to is, access that means, I, you know, didn't go and see the doctor for a month, or a year, or six months, and my diagnosis got delayed. And some of those issues could be patient engagement related. So, I think maybe that's okay as a balance.

MEMBER NEWMAN-TOKER: I think, you know, just now reading your definitions here for patient access to care, includes timely availability of human and diagnostic resources.

I almost think you might want to think about changing it from patient access to care, to some access to diagnostic services, or something like that.

Just because the implication here is that, you know, you've got on call radiology for 24/7, you know, like you've got, organizationally and institutionally you have the capacity to

provide diagnostic services.

Some of those things aren't really patient access, per se. They're more like internal workings of access to care, you know, at off hours, and this, and that, and the next thing. So, I think probably the better resolve is to move access to patients, but to remove the word patient from access.

CO-CHAIR MCDONALD: Are there any comments on the patient and caregiver block there? I know from the last time that we had discussions about patient engagement, sort of patient partnership, patient activation.

There's, you know, kind of all these ways of thinking about kind of what the role of the patient is, and where the vulnerabilities might be. Is having sort of two buckets appropriate? And are the those the right buckets? Put that out there.

MEMBER CAMPISANO: I think that makes sense. And, you know, I think that the way that these domains and sub-domains are laid out makes,

in general makes a lot more sense than how it was previously, just from a lay person's perspective.

CO-CHAIR MCDONALD: I think it being pretty accessible to the patient community is important. And I think this does a better job at that. It doesn't seem as sort of technical.

MEMBER SHERIDAN: I'm nodding my head, because although I'm with, I'm representing CMS, I'm also a patient advocate. So, I'm agreeing with Jen, that I think this is understandable, and easy to digest as a lay person.

CO-CHAIR MCDONALD: Well, it sounds like everybody's liking this. Anyone having sort of remaining concerns, or questions?

MR. LYZENGA: And again, we'll, as we're working through --

CO-CHAIR MCDONALD: Yes.

MR. LYZENGA: -- these activities, just kind of keep it in the back of your head. If you think something's kind of off, you know, raise that during our discussions, and we'll see if we can rework it.

**CO-CHAIR MCDONALD: Thank you.**

**DR. BERNOT: Just to summarize the two things I have so far, that we'll work towards. And that is, if there's something adding patient family caregiver, to make sure we clarify that.**

**And then the other one is getting rid of the patient part, and maybe something along the lines of access to diagnostic services or resources. Something in those, to clarify those two sub-domains.**

**And if anything else comes up, just let us know and we'll get something out to the rest of the Committee, either tomorrow or as follow-up.**

**DR. BURSTIN: Just as an aside to that, just having done way too many of these over the last ten years. The other thing is, once you actually go through the concepts, it's very often that you then go backwards. And then it becomes more clear what the domains and sub-domains are.**

**So, I wouldn't get too hung up on structure, not to go back to Donabedian. But I**

think you've probably done enough for structure. And it's probably fine to just kind of let it flow, and come back to it as needed.

CO-CHAIR MCDONALD: Okay. I think, no, Martha, you're not trying to -- Okay. Okay. That's fine. You get to continuously go for it. No problem. Okay. So, should we move on to --

MR. LYZENGA: Yes.

CO-CHAIR MCDONALD: -- the next item?

MR. LYZENGA: Yes. So, yes, I guess we can kind of jump right into our work here. We're running slightly early, which is great. So, just to give an overview of kind of what, the process we're following here.

We're going to sort of do this iteratively. And we're going to do this a few times. First, what we're going to ask you to do is sort of sit down by yourselves, take your laptop, and work through the measure list fairly quickly, kind of gut instinct.

What are your top five to ten measures? We've actually asked for you to give

us a specific number in each sub-domain. Which are your top measures in each sub-domain?

We've got a spreadsheet we'll ask you to work in, and just put an X next to the ones that you feel are the top measures, the most important the highest priority, whatever sort of criteria you would like to apply at this stage.

Tomorrow we'll be sort of applying our more standard criteria. But in terms of this first cut, we just want you to give again sort of more of a gut feel. Which ones of these are the most important? Which do you think should be pulled out for further discussion and prioritization?

Those will then go to, you'll break out in small groups. They'll essentially do the same exercise, sort of taking fresh eyes, discussing amongst themselves. Do we have the right sub-group of measures pulled out of each sub-domain to then review and rate.

Those groups will report back to the full Committee, which again will kind of go

through the same exercise, asking the question, do we have the right concepts here pulled out of each of these sub-domains? Getting different perspectives who weren't in the sub-groups.

And after that we'll have sort of our subset of measures that will then be more fully reviewed and rated. We'll do the same kind of exercise in terms of gaps. We'll take out the individual part.

But we'll go into small group discussions, and assess gaps within each of these sub-domains. Those small groups will then bring back to the Committee, who will kind of go through the same exercise, look at gaps, review the small group work, bring their other perspectives.

And then finally, and we'll get into this tomorrow, we'll do the small group again. And they'll do those ratings against our criteria for each of those, that sub-group of concepts that has emerged from today's work.

And then we'll bring back to the full

Committee for a check, and to finalize the ratings against the criteria. So, sort of at a high level, that's our approach here.

We'll kind of be doing the same thing a few times, but in different groups as sort of an iterative process. Each kind of doing a check on the next, and providing different perspectives at each stage.

So again, the first part of this is that we're going to ask you to sit down and work through the concept list yourself, and pick out your top concepts from each sub-domain.

Hopefully all of you do have a laptop. I know, Martha, you didn't. But you kind of worked it out on paper, which is great. And we'll incorporate that into the spreadsheet. If anybody else has any issues accessing the spreadsheet, let us know, and we can work out something else.

I think we've got a good bit of time here, until I think 10:30 a.m. is when we had scheduled. So, we can sort of make this into a

break/working session for you guys to just run through the list.

If you get through it quickly, feel free to take a break, look through some of the materials. And sort of think about what we're going to be doing subsequently. But we'll give you, try to give you enough time to get through that list.

We had done it ourselves, a few of us and some other staff, and generally found it to take about 30 to 40 minutes to work through the list and give out top measures. So, hopefully that will be consistent with your experience.

DR. BERNOT: And, Andrew, I don't want to put Vanessa on the spot. But is it possible, since we have a couple of minutes, to actually bring up the, to do a demonstration of the Google doc for the group? So everybody knows what we're looking to do. Sorry to put you on the spot.

MEMBER RADFORD: While you're doing that, can I just ask a question?

CO-CHAIR MCDONALD: Sure.

MEMBER RADFORD: So, have you resorted these measures according to the new -- Okay.

MR. LYZENGA: We have.

MEMBER RADFORD: So, and you're going to send that to us? Okay.

MR. LYZENGA: Yes. We sent it out. And these are in our new sub-domains. And so, I'm glad you guys like them.

CO-CHAIR MCDONALD: Marilyn, did you have -- Yes.

MEMBER HRAVNAK: On the asterisked things. So, are we not to rate those? Are they already accepted? Or --

MR. LYZENGA: No. You should consider those in the same way. Rank them if you think they're important. We just thought it might be worth marking those.

Those are actually fully developed and specified measures that we identified through our environmental scan. That may be something you want to take into consideration as you're doing your rating.

The others are just sort of these concepts that we came up with in the last meeting, and since then with Committee members entering things into that working document. So, those measure concepts are pretty vague in general.

The fully specified measures are a little bit more fleshed out, and have, in some instances, been implemented in programs. So, something you may want to consider.

Many of those, I should note, are very narrowly focused on specific sort of conditions or topic areas. So, that's another thing to take into consideration.

I don't think you should feel like you have to prioritize those over the concepts, if you think there's a concept that better reflects, you know, the sub-domain, or some important aspect of diagnostic quality. Just something we thought you should know.

CO-CHAIR MCDONALD: Yes. It does seem like that's useful for when we get to the rating

stage. So, at this stage, where we're just trying to think through from our own vantage points, which concepts seem more, something that matters to us, that we haven't kind of created a rating around them.

Then, it's about, okay, just, you'll each have your own reason for why you think a set of concepts are more important than another set of concepts in any domain. And that's perfectly valid. Because that's the point of doing things independently to start out with.

CO-CHAIR GARBER: Any other questions?

CO-CHAIR MCDONALD: Vanessa, is this a good time for you to show us?

MS. MOY: So, yes. So in this Google doc there are 12 different tabs. And in each tab there's your name on it for your column. And on the very left hand corner it tells you how many concepts to take out of those total concepts.

Let's say there's 25 out of this tab for diagnostic efficiency, for instance. So, you'll select your top five measure concepts that

you think should belong for this sub-domain.

And then you can mark it by, there's a drop down list. And if you think that's a, one of your top ones, you just mark it. And if it's not your top, you can just leave it blank.

And that's what you do for the rest of them. And for the rest of the tabs, on the upper left corner it will say, like how many measure concepts to select.

And the ones with the asterisks are the ones that were existing measures, as we mentioned previously before. So, yes. Just let me know if you have any questions, and how we can help you.

(Off microphone comment.)

MS. MOY: On the Google doc link.

(Off microphone comment.)

CO-CHAIR MCDONALD: The asterisk means that it's not just a concept --

(Off microphone comment.)

CO-CHAIR MCDONALD: It's a measure. Oh, where would you see it?

(Off microphone comments.)

MS. MOY: So, yes. If you scroll down to --

MR. LYZENGA: It's the first character in the cell.

MS. MOY: Right there.

CO-CHAIR MCDONALD: Yes. So --

MS. MOY: Yes, sorry. So, if you scroll down there's like a little asterisk on one of the measure concepts. So, that would mean that it's like an existing measure before.

MEMBER NEWMAN-TOKER: So, you want us to each fill in our named column for ourselves? The question is, do you care whether we hide or delete the other columns? Do you, how are you going to be reintegrating them.

MS. SKIPPER: Please don't hide or delete any --

MEMBER NEWMAN-TOKER: Okay.

MS. SKIPPER: -- other columns.

MEMBER NEWMAN-TOKER: Okay.

MS. SKIPPER: We did freeze the screen

so that your name should appear in that top row regardless of how far you scroll down.

MR. LYZENGA: Or, if you scroll across it should keep those domains on the left, I believe. So, you can kind of move it over.

MS. MOY: Yes. So --

MR. LYZENGA: Well, almost.

MS. MOY: So, you can locate your name by like scrolling to the right or left.

MR. LYZENGA: It depends on whether they're the last people in the list.

MS. MOY: And you can scroll --

(Off microphone comment.)

MS. MOY: Yes. So, the rows are frozen.

MR. LYZENGA: Already frozen. So, you can just scroll just scroll over to where your name is, and work down.

CO-CHAIR MCDONALD: Basically it's kind of an honor system. It's beneficial to do it on your own. So, go ahead and just look at the items and the concepts, and --

MR. LYZENGA: And again, this is --

CO-CHAIR MCDONALD: -- consider them. And then mark your X.

MR. LYZENGA: Yes.

CO-CHAIR MCDONALD: Because my understanding is that it's not that you're going to kind of use these initial rankings in any way. It's --

DR. BERNOT: Right. This is an informal --

CO-CHAIR MCDONALD: This is for our own process to be able to record and remember what we were thinking. And then we'll use that when we get into the small groups.

MR. LYZENGA: First cut, group exercise will do the same sort of thing, and review what comes out of this. And then the full Committee will kind of review that.

(Off microphone comment.)

MEMBER SINGH: On the second, there's a second column next to, choose a number of top measure concepts, on the instruction sheet.

What's that?

CO-CHAIR MCDONALD: So, Hardeep and everybody, see up at the top, it shows -- No, go back to one of the tabs though. Yes. So, see how at the top it says, please select five concepts from this sub-domain?

MEMBER SINGH: Yes.

CO-CHAIR MCDONALD: So, always look at that. Because sometimes it will be five, sometimes it will be a different number.

MEMBER SINGH: Right, right. But what's the other --

MR. LYZENGA: He was just talking about Column C --

MEMBER SINGH: I'm talking about, go back to the instructions.

MR. LYZENGA: -- on the instructions page.

MEMBER SINGH: What is the 4123 on the right side?

CO-CHAIR MCDONALD: Oh.

MR. LYZENGA: Digital detritus? Or is

that not anything?

MS. SKIPPER: So you all can ignore that last column. Sorry about that.

MEMBER SINGH: Just want to make sure we're following instruction, that's all.

CO-CHAIR MCDONALD: Okay. So, do people feel ready to, I saw a couple of people maybe had already done their homework.

MEMBER NEWMAN-TOKER: Just one --

CO-CHAIR MCDONALD: But most haven't had a chance.

MEMBER NEWMAN-TOKER: One particular question.

CO-CHAIR MCDONALD: Okay.

MEMBER NEWMAN-TOKER: So, the total number of, just give us an approximate for how much time we have for each of these things. We've got about 45 minutes total, and there are how many tabs?

MR. LYZENGA: Twelve tabs. So, you know --

MEMBER NEWMAN-TOKER: So, we got two

or three minutes for --

MR. LYZENGA: Yes. And five to eight minutes I would guess for each one.

MEMBER NEWMAN-TOKER: All right. Okay.

CO-CHAIR MCDONALD: We'll have to move.

MR. LYZENGA: Yes, got it.

CO-CHAIR MCDONALD: But they tell us we can do it in this amount of time. So, we can. David.

(Off microphone comment.)

CO-CHAIR MCDONALD: Okay. Yes. Raise your hand if you're having any challenges, and the --

(Off microphone comment.)

CO-CHAIR MCDONALD: -- staff will help.

(Off microphone comments.)

CO-CHAIR MCDONALD: Okay. Have at it. Yes. And if you have any questions, we'll just have folks helping.

MS. SKIPPER: And I'll resend the link with the Google docs in just a moment.

CO-CHAIR MCDONALD: Oh, yes.

MS. SKIPPER: For those of you who don't have it.

MS. MOY: Yes. I would like to add also, if you need to have a laptop, we can also loan out some laptops for you if you need one.

MEMBER SINGH: And it saves automatically, right?

(Whereupon, the above-entitled matter went off the record at 9:43 a.m. and resumed at 10:51 a.m.)

CO-CHAIR GARBER: This is complicated, and hard to tackle. But thank you all for going through that exercise. I'm going to turn it over to Andrew, who will give us instructions on where we're going from here.

MR. LYZENGA: Yes. So, yes, I'll reiterate. Thank you for sticking through that, and toughing it out. That, we did anticipate would be sort of the rockiest part of this. And

lose a little bit of momentum maybe in the beginning. But we think we'll gain it back in the group stage.

So, again, we just went through an individual exercise to pull out, to select what we think are the top measures or measure concepts in each sub-domain.

We'll now break into some groups. Each group will be assigned a few sub-domains, and will essentially go through that same kind of exercise.

We'll review what the individual results were, which measures or measure concepts rose to the top. And we'll ask you to sort of consider these questions on the slide there.

There's a little bit more detail and guidance in the discussion guide, if you have that. But basically, have the correct concepts risen to the top and selected from each sub-domain. Discuss that amongst your group.

Among those concepts, can any of them be further specified or better described? Some

of them, again, are very vague, and just sort of, you know, a general idea of measurement. Can we get a little bit more flesh around those sort of bones of those concepts? We'll try to do that in other sessions as well.

Should any of the concepts be re-categorized? Did you find in your review, individual review or in your group review, that any of the concepts should be in another sub-domain? That they're not in the right sub-domain?

And then, are the sub-domains themselves appropriate? Do they need to be modified, amended, renamed, rethought?

But the sort of main thing we want you to do is to review which, that set of the most important or top concepts, and make sure that you agree amongst yourselves in the group that those are the right concepts to pull out for further rating. Does that make sense to everybody? Questions?

MEMBER NEWMAN-TOKER: Yes. I'm going

to, having struggled with this myself, let me make a plug for really thinking hard in the sub-groups about, when you're talking about a given measure concept, that you're adjusting the level of granularity to the right level.

Because I think there are some good measure concepts that are buried in some of the individual measures. But the measures themselves seem too specific to be part of the conversation. And I think we should be abstracting those to the level of generalization.

But there are also others where it's the other way. Where, you know, it's like, you know, mom and apple pie. Like, we want, you know, everything to be nice, but there's no way to measure it. And those need to be brought down.

And I think maybe if we could agree that any disease specific one, the sort of base assumption should be that it's, you know, disease X, you know, of a certain class, like, you know, dangerous disease, or whatever.

And only if we specifically think it needs to be a specific disease, or a specific disease class, like cancer, or whatever, that we call it out.

Like, that we be clear in the discussions whether we're talking about a class of, you know, all diseases, this is a representative measure. Or no, we mean specifically breast cancer. And that needs to be a measure. That at least we're clear about that when we're talking to each other.

MR. LYZENGA: Absolutely. Totally agree. So that is definitely something you should be thinking about and talking about in these group discussions. And we'll do that a few times again as we move through this process.

I also just, I forgot to mention, we noticed that in these documents, and in the spreadsheets, it must have happened in an import into the Google Documents, or in the sorting. But all of the measure types, the outcomes, structure, process are completely mislabeled.

MEMBER NEWMAN-TOKER: Oh, thank God.

(Simultaneous speaking.)

MEMBER NEWMAN-TOKER: I'm like, you guys need to give us better definitions of what structure, process and outcome is.

MR. LYZENGA: Yes, yes.

MEMBER NEWMAN-TOKER: Because I'm lost.

MR. LYZENGA: That was a technical glitch I think they got an import or in the sort of the columns. So, Martha actually kindly went through and reassigned all of them. And we'll do that again on the back end. Maybe we can get it for our next round, get those corrected.

But, any more questions or clarifications on the group work? Or should we break out and start our work?

CO-CHAIR MCDONALD: Just, oh, sorry. Just a comment --

MEMBER NEWMAN-TOKER: Can I just do one --

CO-CHAIR MCDONALD: -- on David's

comment though first. And then you can comment some more. I think there's a little bit of a picture of kind of how the group work can kind of sort of move through this process might be helpful.

Because it's a little hard to know, are we going to have like all these separate group discussion, like you just had with us as a big group? And not profit from the big group kind of calibration, cross calibration?

MR. LYZENGA: We will do that after. We'll have the groups do their work on this, sort of these questions. And then we'll get back. The groups will report out their findings, and their discussion.

And then we'll pretty much address the exact same questions again in the full Committee, to have the full Committee sort of check what the group's thoughts and decisions were, add their own perspectives, make sure we're getting it all right.

And out of that we'll have sort of

that final Committee decision on which are the subset of concepts from each domain. Does that make sense?

CO-CHAIR MCDONALD: Yes, that does. And, I mean, I think it's helpful for us to know that as we go into our groups. That we get to come back and say, here's where we sort of wrestled. We could go this way.

Like, you know, worry about granularity, or do groupings, like you said. And that we're coming back to the bigger discussion, knowing that we may in our groups have found things that would translate to other groups. And we may all bump --

MR. LYZENGA: Right.

CO-CHAIR MCDONALD: -- into the same stuff, and resolve it in the same ways. But we may also bump into slightly different stuff, or resolve things in slightly different ways.

So, I wanted to make sure we all understand that we get to sort of step through this, you know, in our sub-groups, but then

together. Go ahead, David.

MEMBER NEWMAN-TOKER: Can I just ask a question about the target numbers of things that we're shooting for? So, we've obviously had this big expansion, you know, to 200 things. We, maybe we had gotten it going down. And now it's blown up, and we're coning it down again, which is all good.

But it seems like you guys picked the numbers, you know, five these and eight of these, almost as sort of like a proportionality of the number of measures that were there.

I'm not sure that that's necessarily fair or right. Like, there may be some places where, you know, for the patient experience, or whatever, there might be fewer.

Because there aren't that many different ideas to deal with there. But it doesn't mean that three or four of them shouldn't end up in the, kind of the final version, even though there were only eight to start with, or whatever.

So, can you clarify for us what each group is supposed to do in terms of target numbers? Are we sticking with the numbers you gave us? Are we trying to get the total package down to 20? Or, you know, what are we trying to do?

MR. LYZENGA: I don't think we have a target number. We're just hoping to get it down to a more manageable number for, again, the full rating exercise. You know, we're hoping to get it down to maybe 50 or 60 to do that rating.

But I think that's a totally fair point. And I think we're comfortable with, if you, you know, if you're, if we've asked you to take out two or three measures from some sub-domain, and you really think we should get three, five more, pick them out. We'll discuss it at the full Committee level. And we can take those in and, we want to be flexible about this.

CO-CHAIR GARBER: And one more question that we need some NQF expertise about. Just, David brought it up. But some of the

concepts are really general, but very appropriate, but probably not actionable.

Do you want us to work on making them actionable, or just note that this is relevant, and you guys will reword it to make it actionable?

MR. LYZENGA: We would certainly prefer that you reword them to make them actionable. If you have ways that you think that could happen, that should definitely be part of your discussion.

Take a concept and see if you can kind of, you know, reword it, describe it, flesh it out a little bit to be more actionable. That would be fantastic.

And again, we can do that at the full Committee level again too. And we can do that as part of our gaps discussion. We'll be looking at, you know, are there things missing?

And as part of that, you know, we can also look at the concepts that are there and say, do they need to be still tweaked a little bit, or

refined?

MEMBER SINGH: So, following that on. So, I think bullets 3, second third and fourth, can should and are, those are fine. The first one is have. And I think that's the most important one. Have the correct concepts been selected. So, I think Mark is sort of trying to get to it, actionable.

I think we need to just be careful about, we're not developing measures. We're talking about measurement concepts that can be further studied, evaluated rigorously and, you know, made into measures in the future. So, I think that's one important thing that all the groups need to realize.

The second issue is, what are the prioritization discussions going to be like? What are they going to be based on? So, one is actionability. Mark mentioned that.

But, you know, it's okay for feasible action, important. But I think this came up last time as well. And my comment on United Airlines

is probably even more relevant now.

But essentially, are we going to be focusing on just things that are preventable harm? Or are we sort of looking at it broadly? Should that be under the discussion?

MR. LYZENGA: That, so we would expect you to sort of incorporate that. We'll talk about our criteria a little bit tomorrow. Or maybe if we have time, later today.

We've pared it down to initially two sort of axis of importance and feasibility. And part of that importance, importance sort of includes a number of different dimensions.

Among those could be, I mean, you could interpret the importance to be how relevant is it to avoiding patient harm, among other considerations.

We actually added another dimension or criteria on, with the input of Dr. Graber, to think about which measures or measured concepts may contribute to cost savings. Because, the idea of that being that that's a very important

area and consideration for decision makers. And may help sort of highlight the importance of these measures as we --

MEMBER SINGH: Can you qualify that? Do you mean cost savings for the organizations?

MR. LYZENGA: Yes.

MEMBER SINGH: Or cost savings in terms of diagnostic testing and --

MR. LYZENGA: I think the idea --

MEMBER SINGH: -- resource utilization?

MR. LYZENGA: -- is sort of, I think this, it actually needs to be fleshed out a little bit more, I think, in --

MEMBER SINGH: Because I don't think any of these measures is going to save anybody any money --

MR. LYZENGA: That's right.

MEMBER SINGH: -- anytime soon.

MR. LYZENGA: Overall, we're sort of thinking overall system costs. Whether it would help to sort of drive down costs for the

organization across the healthcare system. Not necessarily for patients. But maybe that would be applicable.

That, maybe we should discuss that tomorrow as we're talking about the criteria. Because we added that on a little bit late. And it's worth talking through.

CO-CHAIR MCDONALD: Okay. And, Martha? Yes.

MEMBER RADFORD: Also, I just noticed after going through this list twice, for various reasons. A lot of them kind of overlap. And I think it's worth the Committee, the sub-groups calling that out. And maybe summarizing three measures in one that would get to all of those concepts, if you will.

MR. LYZENGA: Right.

MEMBER NEWMAN-TOKER: And one other thing, just in terms of thinking about the future use of some of these measures. So, there were some things where at face value they sort of sounded like a good idea. Like the number of

diagnostic errors reported by physicians and patients, or something like that.

And it seems like a thing that was easy to measure, and that was relevant. But when I think about ones like that, where the result of having a higher or lower number could be both, either a positive thing or a negative thing, in terms of the quantification. That those are probably not great measures.

So, like you could be increasing the number of diagnostic errors reported, because you're doing a great job of beating the bushes, and encouraging culture in reporting.

Or, and your diagnostic error rate, reported rate could go, you know, your total number could go up. So, it's a numerator only measure. Or if you're really amazing at your job, right, the number's going to go down in terms of fixing it.

And I think we have to be careful about using, recommending a direction towards measures that both up and down could be a good

sign.

MR. LYZENGA: Thank you. And please do. You know, those are exactly the kind of thoughts we want to sort of have come out of this. We'll be, you know, we're recording this all. We're going to be reviewing the transcript.

And those are the sorts of things that we can try to incorporate into the report. And sort of add to the richness of the discussion around these measures. And say, these are the considerations to take in mind, you know, just the sort of thing --

MEMBER NEWMAN-TOKER: It's basically especially risky with numerator only measures, where you just don't have a denominator, and you're just dealing with, you know, unsystematic reporting where you just don't know whether you're doing a better job or a worse job.

CO-CHAIR MCDONALD: And, Prashant?

MEMBER MAHAJAN: So, just to clarify my understanding. We are looking at the conceptual area? Like, so for instance, patient

involvement in diagnostic understanding of the process.

To me that is a conceptual area, which we could say, yes, it's important, no it's not important. And that's what we are doing, rather than percentage of patients who understood their diagnostic process, you know.

See, because that is a measurable event. But what I think you were alluding to, Hardeep, is that we are looking more at the conceptual approach. Whether it's important for us to think whether patient involvement and understanding of the diagnostic process is important or not. Yes, or no. And then the actual measure, right? I'm not sure.

MR. LYZENGA: Yes. I think that's fair. And I think we can, you know, we would like to get down to specific sort of concepts or measure approaches if we can.

But if you, you know, that is perfectly reasonable, to say we think this is an important sort of concept to measure. And then

we can sort of say, here's one way you might do that, you know, this X percentage or, you know, the numerator or denominator. There may be other approaches. But we think that in general is a very important concept that we want to measure.

MEMBER RADFORD: I would go further than that. I would say, if we think the concept that's listed is not important to the diagnostic process, we go no further. And we only kind of think hard about those that we think are important.

And then, you know, if they need tweaking, or they need specification, or if they need a little more clarity around even just the concept, well then, let's work on that.

MEMBER SINGH: I think that's where --

CO-CHAIR MCDONALD: Yes. So, I mean, it seems that we probably should get into our small groups, try this out. This kind of conversation is what we'll have in our small groups.

And the idea is that whatever we're

discussing about how we're thinking about this, is part of the process that you need. But we can do it in our smaller groups, with these smaller lists, instead of the list of 200.

And at the end of the time we'll have some, you know, answers to some of these questions from each small group. And if we have questions, we'll be able to ask them in our small groups. Because we'll have a NQF staffer with us. But I'll just take the last series. Kerm?

MR. HENRIKSEN: Is there any interest in distinguishing between quality and safety measures, where there's a harm involved? Going back to Hardeep's earlier comment that some of these are very broad based, and seem to be quality.

MR. LYZENGA: If the Committee thinks that would be valuable, maybe we could something like, you know, tagging measures. This is sort of related to quality. Or, you know, this one is related to safety, or is a sort of safety sensitive measure, or something like that. And,

I don't know if that's, or some other approach.

MEMBER SINGH: So, I have another sort of suggestion, and maybe take on quality. I think there are several areas which are important. They're clearly important. I mean, one that, I'm not going to pick on anyone, but health literacy one came to mind.

Yes. It's really important to do that. But is it measurable immediately? Can we do something about it, or not? And I think that's number one. So, that, is it quality, yes, for sure?

And the second is, has that been associated with preventable harm? That's sort of getting into safety. And I think we need to have that kind of discussion when you go into the small groups.

Because that's going to lead to some kind of prioritization better, if you ask the questions. And then, maybe we can mark the ones that came out be under lots of scrutiny because they were important.

If people felt this was a really important area, but was not measurable. Or it was a really important area, it was not directly related to quality, to safety. But it goes under some other quality domain.

And maybe we could just sort of re-market those measures and give it to you. I mean, I'm sure there are other workforces addressing this issue too.

CO-CHAIR MCDONALD: Ready to divide this up?

(Off microphone comment.)

CO-CHAIR MCDONALD: I mean, I think this is, I think we're just having kind of little talks that will be the same talks we have to have. And that the focus is, that you'll put those questions back, and our groups will --

MR. LYZENGA: Yes.

CO-CHAIR MCDONALD: -- have to sort of organize ourselves to move through those questions --

MR. LYZENGA: Yes, exactly.

CO-CHAIR MCDONALD: -- and come back.
MR. LYZENGA: Again, we'll kind of --
CO-CHAIR MCDONALD: Yes.
MR. LYZENGA: -- iterate on this --
CO-CHAIR MCDONALD: Yes.
MR. LYZENGA: -- sort of at the group level --
CO-CHAIR MCDONALD: Yes.
MR. LYZENGA: -- and at the Committee level.
CO-CHAIR MCDONALD: Yes.
(Off microphone comment.)
MR. LYZENGA: No, no. We split those out across, say that your group will be assigned a few of the sub-domains.
MS. SKIPPER: All right.
MR. LYZENGA: So, do we have our materials ready for the groups?
MS. SKIPPER: Yes.
MR. LYZENGA: We do? All right.
MS. SKIPPER: Okay. So, we'll split up. Group 1 is me. I guess we can gather here

by Nicholas and Marilyn. Group 2 will be with Vanessa. Just, you all can gather down there by Thomas. Group 3, Andrew and Tracy, over in the corner by Kerm. And then Group 4 I guess can go here.

And then, members of the audience, please feel free to join any group, and sort of float around. So, we can break into groups, 1, 2, 3, and 4.

(Whereupon, the above-entitled matter went off the record at 11:11 a.m. and resumed at 12:19 p.m.)

CO-CHAIR MCDONALD: We have a chance for public comment at this point. And I think we're going to go to the phones. Is anybody on the phone ready to give public comment?

OPERATOR: If some of you would like to make a public comment, please press *, then your number 1. There are no public comments at this time.

CO-CHAIR MCDONALD: Okay. Thank you. And are there public comments in the room? Paul

Epner?

MR. EPNER: Paul Epner, Society to Improve Diagnosis in Medicine. Love the domains and sub-domains. You have external environment under organization, which is sort of a catch all.

I would, but I think there's some issues that are both, that are in there, and maybe overlap with access to care, that might need to be called out in your definition or example. So, one, employer issues such as being allowed to have time off to go get tests, or go have doctor's visits, or whatever.

And payer issues, where they can put patients through a tremendous tech review cycle. They can deny, and you have to go through appeal process. And you have payers in here, or cost issues. You don't have payers, per se.

But then can insert a tremendous delay if a physician asks for a certain diagnostic procedure, and they don't agree. That process could cause a big delay.

So, just again, I think you've got the

right categories. But you may want to add some rich examples, to make sure it gets outside of the health system walls.

CO-CHAIR MCDONALD: Thank you. Any other public comments here? Okay.

(Off microphone comment.)

CO-CHAIR MCDONALD: So, we've done public comments. What next?

MS. SKIPPER: And so, we'll break for lunch. And we'll resume at 12:45 p.m., so that groups that need additional time to go through their measure concepts can do so.

CO-CHAIR MCDONALD: I was hoping we were going to hear that. Lunch time. Thanks.

MS. SKIPPER: Thank you.

(Whereupon, the above-entitled matter went off the record at 12:21 p.m. and resumed at 1:36 p.m.)

CO-CHAIR GARBER: Okay, thanks for coming back. I hope everybody enjoyed lunch. I will do the reporting or share the reporting responsibilities for Group 1 with Tom.

The approach we used was to first hear what items on our list were highly rated to make sure that, indeed, we thought they were appropriate to include. We then went over the items on our list that got zero or very low scores to see if there were any that we thought had been overlooked.

We then went through and tried to decide whether there was any overlap or measure concepts that could be combined or needed to be reworded. And basically, we ignored the several very specific measures that were included in our group, because they seemed to specific and already very well developed.

Although, you'll see one example where we went backwards and took a measure that was very specific and translated it into one that was more general.

So we are reviewing the topic called Information Gathering and Documentation. And these are the measure concepts that we thought should be included. It was Number 5, that the

problem list is accurate and up-to-date. And people thought that was an important enough concept that it should stand on its own.

I'll just go through them all? Yes? Number 6 had to do with the percent of cut and paste in notes. And we combined that with several others into a measure concept that speaks to the adequacy and accuracy of documentation in general.

So our suggested revision is that clinical documentation should support quality in the diagnostic process. It should be clear, complete, and accurate. It should discourage inappropriate use of copy/paste, it should include the rationale for making a diagnosis, and it should include a differential diagnosis for new complaints and an accurate problem list that is reconciled with the patient and at transitions of care.

There's a little typo in the bottom there. We want an accurate problem list, not an inaccurate one. Yes.

(Off microphone comments.)

CO-CHAIR GARBER: Let's go through them all and then we'll just go back and make comments. So I think those were -- we definitely want EMRs to be able to capture the chief complaint and the reason for the visit.

Next? And we combined 11, 12, 13, and 14, so our suggested revision is down at the bottom there. It says that EMR should not require documenting a diagnosis before it is appropriate to do so. It should allow designating patients as being not yet diagnosed.

It should allow providers to assign a probability of a diagnosis being correct. And it should allow for ways to distinguish an initial or an admitting diagnosis from a working diagnosis, from a final diagnosis.

I think that's our last -- okay. And about communication, we combined several. So our suggested revision is down at the bottom on the left. Communication to patients and their families should be documented, and patients

should be aware of their diagnoses.

And 38, 38 is an example of an existing measure that was very specific, that within 60 minutes, if you're being transferred to another healthcare facility, you should have complete documentation.

So we're suggesting a more general concept, and that is that complete information about diagnoses may, during an in-patient stay, including tests pending at discharge, should be available in a timely manner to the clinical care team who will be seeing the patient subsequently.

I think that's it. Okay, so those are our recommendations for the measure concepts for information gathering. David?

MEMBER NEWMAN-TOKAR: Just a quick comment about granularity. Can you go back a couple of slides there to, yes, stop, this one? So you have -- so there, clearly the percent of cut and paste in notes is either all the way down at the measure level or a measure concept, certainly.

We called this higher level thing that you've got here aggregated as 19, the adequacy of documenting initial findings, the clarity and accuracy of documentation, a measurement theme. We found that we were missing that kind of one rung between sub-domain and measure concept.

Because underneath it, your bullet points are all measure concepts, in some sense, right. Like, you could talk about, under copy and paste, your measured concept could be the percentage of cut and paste in notes. Or it could be the measure concept underneath the measurement theme of adequacy of documenting the initial findings.

So in case that helps as we sort of struggle with keeping things at kind of the same levels, I think maybe we need an intermediate between sub-domains and measure concepts, just to lump things in a sensible way.

Because clearly what you were saying was we need to find some way to make sure that people are getting clear and accurate

documentation in the EHR.

CO-CHAIR GARBER: Right. I think as we see how each group did it differently, we'll get some feel for those different levels. Thank you, good suggestion.

CO-CHAIR MCDONALD: I have a question for your group. In terms of when I look at the list versus what you've pulled, I was wondering about all these things that are actually quite clinically narrow and specific, whether there was any discussion of batching those in some way.

So, you know, this was the list that had a lot of measures that already exist, percentage of patients with esophageal biopsy reports for Barrett's esophagus that contain a statement about dysplasia and if present the grade of dysplasia.

So it had a lot of these, like, very, very specific things. And I just wondered if there was any thought about, like, grouping those as a concept, that when there is a specific thing that could be looked at that it would be on a

list of measures to --

MEMBER NEWMAN-TOKAR: Yes. We didn't have time to go through each one of those and see whether there was some more general concept that needed to be spelled out except for this one example we had on the last slide. But I think you're right, I think that we should do that, go back and do that.

MEMBER SINGH: Do you have more, or just that information?

CO-CHAIR GARBER: That's it from information gathering.

MEMBER SINGH: But you've got more other concepts to go through?

CO-CHAIR GARBER: We have one more concept.

MEMBER SINGH: Oh, okay. Okay, so I just wanted to sort of just reflect on a couple of things. So we're still -- some of the intent, I'm getting confused. So some of them, the language -- or maybe that could get cleaned up -- is more about this is what you should be doing as

a good practice.

It's just a good clinical practice to have, you know, good notes and not do copy and paste versus actually a measurable concept or a measure concept. So I think we need to get a little more defined.

And if you look at your slide, the one that you just showed, not your actual last slide, the timely one, you know, somebody started with 60 minutes. Fine, you could have X-minutes. But I think we should think about having some defined parameters.

What is timely? What is accurate? We don't know. I mean, problem list, how do you know it's complete and accurate? How would we ever know if it's complete and accurate?

So I think we need to sort of think about what we are proposing. Is it just a good clinical practice, which is a good thing, or is it something like a measurement concept that we can actually do something about.

So I think that kind of distinction

we're going to have to think about doing. So I was getting confused between, yes, this is all great stuff, but Helen's going to --

DR. BURSTIN: I actually think it's a really good thought. And I think some of this gets at this question of what's a really interesting idea, broad concept, and what becomes a performance measure? Are they really talking about here a measurement concept, meaning what's the target focus, what's the population you're most interested in?

I mean, this is a great topic. And it's really important to this. But I guess the question is how would you actually get to a level of detail more granular? For example, where I practice, none of my residents have any idea how to take a problem off the problem list. There's no dating, there's no time, there's no -- I mean, I'm very serious.

So the question is are there elements of this where you create what are the things you would measure with a bit more specificity of what

is, in fact, an accurate and timely problem list? Because otherwise, just simply stating it isn't a concept, it's just a nice idea.

MEMBER SINGH: And, you know, that's exactly what I was getting to. I think Helen put it much more nicely than I was trying to get to. But the point is, you know, like, differential diagnoses you've got on there. Absolutely. We've got -- we've seen patients who are completely misdiagnosed, no differential diagnosis, right? I mean, you quote that study as well.

Can we try to make it some type of a measurement concept around differential diagnosis? I mean, just start. More than half of the records seen on a, you know, triggered patient list, I mean, just make up something to make it more measurable. That's what I'm getting to, rather than having a way documentation is related.

CO-CHAIR GARBER: But is there a measure? It sounds more like a measure than a

measure concept?

MEMBER SINGH: We could make it a measure concept. I'm just thinking. I think we're going to have to narrow down --

CO-CHAIR GARBER: Well, let's take an example and see if we could do that. Because I'm really struggling with the difference between a concept, and a measure, and what's in between.

DR. BURSTIN: Why don't we stay on that example. If you want to, I'll try it. So for example, it said that you had a, where was it, a complete and accurate problem list. Is that still on this slide somewhere? It's the one -- but can you go back one, wherever that is? Okay, here we go.

So for example, problem list is accurate and up to date is really, I think, the goal of what you're trying to achieve in the measure.

You could do something like percent of problem lists that include time stamps, you know, something that allows you to actually qualify

what that is among all those using electronic health records that include a problem list. So something that specifies exactly what it means to, in fact, be accurate and up to date.

MEMBER SINGH: Can you go a little bit more extreme and think about something like number of diagnoses found inaccurate on a review of a hundred problem lists? I mean, that would be a measure, but --

DR. BURSTIN: How would you know what's inaccurate among --

MEMBER SINGH: Well, that's what I'm saying. You will have to get to a review, which is exactly the point. How are we going to operationalize any of these areas?

CO-CHAIR MCDONALD: Okay, we have a lot of signs up so make a list.

MEMBER HUNT: I was going to say, one way you could do this is by having the patient or family compare the problem list to what the doctors have and make sure there's agreement. And that would be one way that you measure this

pretty easily, I think.

CO-CHAIR MCDONALD: And, David?

MEMBER SINGH: There's already a measure on that.

CO-CHAIR MCDONALD: Oh, what, David? This one, and then -- I'm going back and forth.

(Off microphone comments)

MEMBER SEIDENWURM: So we use just reconciled at each visit.

CO-CHAIR MCDONALD: You said --

MEMBER SEIDENWURM: The problem list, we use, in our place, reconciled at each visit. There's a million ways you can operationalize this.

MEMBER SINGH: Well, then it becomes one of those medication reconciliation things that we could never, ever do for the last decade.

CO-CHAIR MCDONALD: David, it's David Hunt?

MEMBER HUNT: Yes. No, I was going to say that so many of us in practice know the problem list is a cesspool. It's just something

--

(Laughter.)

MEMBER HUNT: It's just something that's just continually added to. But some things are never taken off. And who is ultimately responsible for the problem list? Traditionally, it should be the primary care physician, I would say. Because they're the coordinator of everything.

I would like to also offer up support for a concept of the differential diagnosis. In many EHRs, there's not a place to really put it in a structured format. And to be really useful and manipulatable, and that's not even a word, but it would have to have a structural place in the EHR.

And so I would advance the ability to have -- at least to record what is your differential. Because my residents nowadays, they don't even list the differential, usually, in their records. Because, you know, well, where am I going to put it and that type --

CO-CHAIR MCDONALD: Okay, Marilyn?

MEMBER HRAVNAK: So in defense of Team 1, I think that what we were trying to do was -- I thought that our first assignment was to try to separate the wheat from the chaff a little bit. So to take this -- and we had two very big areas to take out and just kind of pull out those concepts that we thought were thought were the most important versus those that were lower on the list.

I don't think that we -- you know, that was sort of like the first pass. And then I think the actual measurement concepts under them, it could be that some of them might be very generic for that particular concept and would apply to all settings and all populations.

Whereas some of them might be more specific, you know, some of the ones we were looking at specific to a cancer diagnosis or a, you know, a heart disease diagnosis.

So I think that's why we kind of had to step back, just sort of looking at what are

concepts that we should be -- that should rise to the top of this list to begin the winnowing process, was my view of what we were doing.

CO-CHAIR MCDONALD: Reasonable, Lavinia, you had yours up. Okay, no. Okay, David?

MEMBER NEWMAN-TOKAR: So I think there're actually, as a dimension, several different ways that one could operationalize this, including some that haven't been mentioned.

So, for instance, you could say, okay, I'm going to take five or ten diseases where I have, you know, common diseases, where I have a known gold standard of some kind.

Like, you know, most of the cancer diagnoses have pathology reports somewhere in your EHR that say lung cancer, or breast cancer, whatever. And your stroke patients have, you know, stroke written on their MRI report, or whatever.

You can electronically compare how often patients with, you know, Disease X, as

confirmed by such and such testing that everybody can agree on is a gold standard, in the cases where we have gold standard -- we don't have a gold standard for everything, but we have gold standards for a bunch of things -- and you could say, okay, how often, when there's a gold standard diagnosis of Disease X, does it show up in the problem list? And how often does it show up in the problem list without a gold standard diagnosis of Disease X?

So you wouldn't get an all-encompassing view of whether the problem list was perfect, but you could still measure whether, you know, where you were at, like, if you were at 50 percent sensitivity and 50 percent specificity for the top ten, you know, things that you had gold standards for. That would be telling you something.

MEMBER SINGH: Yes. Actually, I think Adam Wright from -- and, Tom, you may want to add to this, Adam Wright from Harvard has done some work in the area.

So you're looking at patients who have definite diabetes, because they've got hemoglobin A1c showing up on your EHR measurement as over ten, or they take oral hypoglycemic, but diabetes is not on their problem list.

So we could come up with a bit more specific measure, saying definite presence of disease by other criteria, but absent on problem list, as a little bit more specific measure concept that we could push forward.

MEMBER NEWMAN-TOKAR: And just one more issue on that granularity. Like, I would take what Hardeep said and say, okay, look, the measure concept is, you know, Disease X diagnosed by gold standard test, present or absent on -- you know, percent present or absent on problem list.

The individual measures are one for diabetes, and one for stroke, and one for cancer, whatever, and this is a measurement theme. This idea is I want the problem list to be accurate and up to date. How am I going to operationalize

that as a measure concept and then individual measures?

CO-CHAIR GARBER: Yes. That's very helpful. Could we go ahead then, Tom, with your section?

CO-CHAIR MCDONALD: Put your microphone on too, Tom.

MEMBER SEQUIST: Okay. So this is the section on information integration. I think it's the next one. Yes, okay. So our top seven choices for information integration, so I guess in this section we did a little bit more of this. We kind of changed the wording of some of the -- how it was actually displayed in the Google doc.

And we had a little bit of -- took a little bit of liberty with trying to reinterpret what was being meant. So this one, medical record sharing among non-economically related entities, we sort of re-translated that as participation in health information exchange across institutions that support diagnostic quality, such as transmitting test results and

disease diagnoses. So that was, again, that was sort of our interpretation of what it meant to medical record share across non-economically related institutions.

I'll just keep going, and then we can do comments and at the end. The next one was proportion of diagnostic evaluations with appropriate team involvement. So we had a lot of discussion about what this meant and what it was to have team involvement.

How we rephrased this was, proportion of diagnostic evaluations with appropriate patient and inter-professional team involvement, such as nurses, physicians, pharmacists, and everyone in the medical neighborhood.

So the next slide, yes, all right. So, proportion of patients diagnosed with a specified targeted disease of interest who received a second opinion. So this probably had the most debate in our group about whether this -- basically, I think a lot of the debate was about whether this was separate from what I just

described on the prior concept which was around appropriate involvement of all members of the team.

And I think that the debate surrounded this concept of what does second opinion mean. Because I think you could take that as a very broad generic, kind of asking anyone, asking the nurse who works with you, or the pharmacist.

But what people are -- my concern and some others, the concern was the term second opinion is, to many people, going to mean you asked another physician specialist, like you referred them. You placed a referral.

And we weren't entirely sure that that was a good measure of diagnostic quality, simply the act of involving more physician specialists. So there was a lot of debate on that one.

But we thought we would just put it out here and get group discussion. Because there is this other important concept that it is good to get second opinions in general, to get advice from other people, like, even within your primary

care practice, let's say.

If you're struggling with, you know, a person comes in with a rash on their foot, and is it cellulitis, is it contact dermatitis, and getting to somebody else to say, hey, what do you think that is? And so trying to capture that spirit without saying that I'm asking you to always refer patients to specialists was the sort of debate here.

So the next one, Number 10, we lumped a bunch in this area of information integration. We thought Concept 10, 13, and 19 were all related to each other. Ten and 13 were almost exactly the same thing. One was can you track closed-loop referrals, and the other one was your performance on that, the percentage that are closed loop.

But Number 19 used the standardized communication techniques between consultants. We thought it was sort of included in this concept of how well are our referrals being managed from sort of soup to nuts, from the placement of the

referral, to the occurrence of the visit, to the communication of the treatment plan, and any results back to the referring providers.

So we came up with the wording of, close-loop referral to specialists, including completion of visits and communication of test results, and treatment plans, or treatment recommendations back to the referring team. That's sort of encompassing 10, 13, and 91 together.

The next slide. So Number 12: use of structured hand-off programs in the hospital. We left that one as is and thought that was an important concept.

Number 15, so there was -- all this said was diagnostic reconciliation. So several of us in the group didn't know what that meant or hadn't heard that term in particular before.

So we sort of thought about what it might mean and, actually we were all just talking about this a couple of minutes ago, are we going back and confirming diagnoses? And is the

problem list actually accurately listing what the, you know, what conditions the patient actually has?

Similar to Med Rec, and again like the group was just saying before, it comes with all the same problems as Med Rec, good concept, nearly impossible to measure in any, you know, meaningful way.

So, again, we liked the concept of saying, you know, revisiting, making sure if you're saying someone has COPD that they actually have COPD. If you're saying somebody has, you know, X, Y, or Z disease that they actually have that disease.

Although our sort of caveat with that is it does -- and then we put it on here, Med Rec, it does sort of lead you down that same pathway as medication reconciliation.

Number 23, so this said, correlation of histology and molecular findings. So, we weren't necessarily supporting that particular measure, but it got us into the conversation

around the general concept of are you looking at all the sources of information that you have in the record and making sure that they all lead you to a concordant diagnosis.

And really, we got into this discussion of saying, okay, well, do we have any concepts in here around information integration that helped counter our heuristic biases?

So if I have a confirmation bias -- and I'm just, like, I think the patient has X disease, and any information that gets presented to me that supports it, I'm going to use it, and any information that gets presented to me that doesn't support that, I ignore it -- do we have a process where you are sort of forced to reconcile outlier information?

Again sort of, we thought, an important concept, didn't get to the point of how would you operationalize that, you know, and maybe start to feel a little bit more like Med Rec again.

But we just felt, like, in the

information integration there wasn't anything in here that addressed our biases that we bring into the diagnostic process as part of integrating information. So that's why we had that one in there, again, open for discussion.

And then the, let's see, the last one was more general. So I just talked about the first bullet on the last slide around biases. And then this is all actually relating to that last concept. I think I covered it all, but --

(Simultaneous speaking.)

CO-CHAIR GARBER: No, good job, absolutely. So we definitely need some input from the group on whether there needs to be a measure concept related to getting second opinions. Do people feel that's something that's valuable and we should include or not?

MEMBER SEIDENWURM: We know that in the radiology field, and I believe this is also documented in pathology as well, that there are improved outcomes, for example, in mammography. It's been well documented.

CT for oncology follow-up, I think, is also documented in certain pathological circumstances as well, I think, in the lymphomas and in the distinctions among the different grades of breast cancer, where second opinions are valuable.

The problem is, they're expensive. And they're difficult to administer, especially in areas where there might be only one competent specialist available or expert in that particular area. So it would be great if we could do it, but the other impediment is that there's often not -- it's not a chargeable event either.

CO-CHAIR GARBER: Hardeep and then David.

MEMBER SINGH: You know, I was going to say, I think we're going to need to do something about second opinions, especially because I think it's an important concept.

What we need to do is debatable. Because my view on this would be oftentimes it's the fact that, I know it's hindsight, but a lot

of times people should have got a second opinion or, you know, we should have got a second opinion for them, depending on what side you are, but it was not done.

So I think we may have proposed something like this similar before. But something like -- there are known diagnostic dilemmas, you know. Celiac disease comes to mind. There is, you know, ankylosing spondylitis. There are certain conditions for which it is almost essential that you need to get second opinions for or specialists to refer them. Oftentimes you just can't tell which one is which.

And if you've got a condition such as celiac disease which took ten years to get diagnosed, where a second opinion just happened one year prior, and for nine years somebody never got a second opinion, or their physicians never got a second opinion for them, that's a problem.

So I'm just wondering if we can re-frame our measurement or measurement concept in

some way. Or, for acute disease, make it about, I mean, we had this paper, spinal epidural abscess. People came in recurrently to the emergency room with multiple red flags, and they were sent back.

So if you've got a patient who just got diagnosed with spinal epidural abscess but had to have seven emergency room visits in the last, you know, two months, or three weeks, whatever it might be, that becomes a measurable or measure concept. So either do it in absence of or too much, you know, depending on what encounter you're referring to.

MEMBER NEWMAN-TOKAR: So I agree with Hardeep that you probably need to get a little bit more specific. Because if you have a too blunt tool in this situation, you're going to end up with the is it better if I had more or better if I had fewer kind of problem.

I think Hardeep's onto something there, if I could paraphrase. He's saying in situations where we know the diagnostic failures

are common or, you know, sort of a potential pitfall, we could pick those situations the same way in the last scenario.

We could say what are the diseases with a gold standard diagnosis where we could reference where the problem was. What are the situations where we know these diagnostic errors and delays are happening, and can we reference how frequently those people are being sent for second opinions or how early they're being sent for second opinions?

So I think if you narrowed the focus and didn't just say, you know, what is the total percentage of our population that gets a second opinion, which I think is not going to tell you very much, I think if you start getting specific into situations where there's diagnostic uncertainty and known pitfalls, I think that's the right idea.

CO-CHAIR GARBER: There was one concept that was raised by, I'm not sure who, that it would be nice, for example, in a group of

primary care providers if there were somebody who raised their hand and said I'd be happy to provide second opinions. So that's a kind of a structure measure. Is there a person who is available to provide second opinions for patients who would like one.

MEMBER NEWMAN-TOKAR: One other thing is I would just make sure that we distinguish between second opinions and the sort of quality improvement that hopefully we'll talk about activities where you deliberately do second reviews of some sample of, you know, like, you double-read ten percent of the pathology slides, or the radiology, or check clinic charts, or whatever.

Like, that's a little different though. You're talking about second opinions, like, I don't really know what's going on and, you know, I need somebody else to help me rather than just test, retest.

CO-CHAIR MCDONALD: So on this one there is sort of the patient-facing side, and

there's kind of the payer-facing side where sometimes second opinions are, you know, we often think of second opinions with treatment.

But for these types of situations where the diagnosis is going then produce, you know, maybe a high cost treatment, a second opinion could be used to verify the diagnosis, and it would find some that were missed.

If we think about the payer side, maybe payers would get involved in figuring out where it's appropriate to have second opinions around specific diagnoses.

And then if you think about the patient side, patient's being able to get second opinions when they feel like they're not getting a diagnosis that matches or comports with what, you know, the explanation has not turned out to resolve anything, perhaps because they have been misdiagnosed.

MEMBER HUNT: I mean, the payer getting involved is a little chilling to me. But that's my own issue.

One thing I would like to see supported in the second opinion is some institutional activities that may be included as a second opinion, but it's not a formal one.

And I'm thinking of Tumor Board. I have gotten a tremendous amount of feedback when cases are presented at Tumor Board. Because you have a group of your peers that really weighs in, similar to almost Mortality and Morbidity Conference also. In some way or another being able, if the measure concept or if the measure could include that type of activity, I think it might be good.

CO-CHAIR GARBER: Tom?

MEMBER SEQUIST: So I think that was the spirit of what we were talking about. It was less about can we measure referrals to specialists, but does the environment in which you're practicing sort of enable you or facilitate you getting additional input on what you're doing?

And we've given the example in my

clinic, or our clinic I was in a while ago, actually someone had started a sort of an online, although within our firewall, blog where primary care doctors, among 20 or 30 primary care doctors could all say, you know, hey, I saw this patient today who had X, Y, and Z. Has anyone else seen that? I did this. What would you do?

And then, like, the 20 of us could write back and say I would have done this, or that's what I would have done.

I'm not saying we have a measure around who's got blogs in their clinics. But it's more, like, the spirit was, like, is there a way that it's being enabled that you don't practice in a silo, that you are somehow able to get -- now I worry about that being too -- I was going to say touchy-feely, not touchy-feely, well, maybe I should be worried about it being touchy-feely. But it's to, like, yes, squishy maybe. I don't know, but very hard to turn that into a measure.

But it's something really tangible,

right, that a practicing clinician feels, right, that even though I'm surrounded by 3,000 doctors in my hospital, I'm basically practicing medicine alone.

(Off microphone comments.)

MEMBER SEQUIST: Sure, yes. So if you use, like, the arc, you know, sort of model of --

(Simultaneous speaking.)

CO-CHAIR GARBER: Last call for comments on this section. So Prashant, David, and then Hardeep.

MEMBER MAHAJA: So my only question is after this comes out it does appear that the second opinion is getting some traction. I'm just wondering that should we also be looking at the downside effects of the decisions that come out of this committee?

For instance, I'm just looking at this, like, the second opinion were to be given more importance. And depending upon how it is received, it has implications on excessive investigation, or more diagnosis now in a

different cost, and psychological impacts. So I'm just going to throw that out. Do we need to think about it at this time or not?

CO-CHAIR GARBER: Hardeep?

MEMBER SINGH: I was going to say I found, you know, that I was scared. Because David said just the opposite. When Kathy was onto something about payers, and David said no, no, no, keep them off -- sort of --

MEMBER HUNT: It just chills --

MEMBER SINGH: Yes, chills. Yes. We don't have a payer representative here, right? We don't have anybody from there. But they have a lot of the data. And they know what the final diagnosis is, because then they are paying for treatment. Well, hopefully it's correct.

But if you've got a patient with, let's say, multiple sclerosis, and you're paying for that diagnosis, you can go back and look how their, you know, utilization was and how many second opinions they got, or how many neurologists they went to, and how many primary

care visits were for, you know, for a totally unrelated neurological condition before they got diagnosed with MS. So they could be a source of rich data that you want to consider some utilization metrics on.

CO-CHAIR GARBER: Great. Thanks, everybody. Let's move on to Section 2, Group 2.

MEMBER NEWMAN-TOKAR: So Group 2 used a very similar process, the ones that were described earlier that Mark articulated, the top ones, and then we looked at the other ones that didn't make the top, made sure there wasn't anything. And then we tried to consolidate.

And we sort of went with this kind of intermediate kind of measurement theme idea underneath the sub-domain. Because we could all quickly agree that the theme was either important or unimportant. And then we sat around and argued about what the measurement concepts would be within that theme.

So the first thing we agreed on pretty easily under the diagnostic efficiency sub-domain

was the timeliness of the diagnosis of priority Disease X. And we said priority just meaning, you know, whether you're believing that that's diseases that kill people, like cancer, and sepsis, and whatever, or whether priorities are public health oriented things, like, you know, a lot of patients with asthma, or diabetes, or whatever.

So people could kind of define that the way that they wanted to. But then we wanted to get something about timeliness. So we then broke that up into two domains that would get us to kind of a measure concept here, sorry, domain is not the right word, two sub-themes or two measurement concepts within this theme.

And one was the timeliness of initial diagnosis, that is from essentially the symptoms to the explanation.

And the second sort of phase of that, the timeliness of from the explanation to management, recognizing that the group felt that there was, even though some of this is a

continuum, that there was kind of this difference between measuring whether you had, you know, gotten to the point of a lung cancer diagnosis, as opposed to getting all the subsequent staging, and testing, and whatever else, to get to the point where they actually knew which chemo to get you after they did the molecular diagnostic tests, and so on, and so forth.

And we've actually seen at Hopkins some of this. You know, our lung cancer folks have done a great job of squishing down the time from biopsy proven lung cancer to the point of actually chemo infusion. By bundling all the diagnostic tests up and happening quickly, they cut out a month-worth of wasted time.

But that's a little different than that time to first get to the point where you've got the cancer diagnosis in the first place. So it was enough feeling that those were discreet that we should have two separate measure concepts there.

Obviously, the timeliness concept, as

everybody said before, is really around -- is really disease-specific. And we didn't want to get too hung up and particular about what would be considered acceptably timely.

But I think that, obviously, it's going to be different for subarachnoid hemorrhage than it is going to be lung cancer, than it is going to be celiac disease, or whatever, conceptually. And some of that's just going to have to be worked out with the science of measurements or figuring out.

And there are nice epidemiologic studies showing, you know, sort of looking at the relationship between how long a diagnosis took and the likelihood that there was kind of a missed opportunity. So that's those two.

And then we moved to the next sub-group which was really around -- this theme was around value in the diagnostic process.

So there was really a concern expressed in the group that we needed to make sure we didn't just deal with the under-

diagnosis, if you will, or missed diagnosis type problem. That we also thought about the issue of over-testing, and over-diagnosis, and that there needed to be somewhere in there, some kind of measures that sort of dealt with the specificity problem, not just the sensitivity problems, so to speak, if you want to put it in diagnostic test terms.

So here we had one of the sub-themes was related to over-testing, the other to over-diagnosis. And then we had three measure concepts in the over-testing.

One was that whether there was a policy in place at the organizational level or standard operating procedures for some kind of gatekeeper function for tests that are known to be overused.

So if there were specific things where it was just known that it was being done too much, molecular diagnostics for cancer, or whatever it was, that there was some mechanism for tamping down on that.

The percentage of patients with Symptom A, or Disease X, who are tested inappropriately, and we gave you a couple of examples there, that actually is -- there's an extra bullet there that's not there. There are only two concepts there. I think it's just -- that's a mess up on my part. I hit an accidental character return there. Yes, there you go.

So, like Lyme disease serology ordered in a patient with non-specific rash in a Lyme endemic area, you can subtract stuff like that.

Now, on the over-diagnosis side, we have these two bullets here as sort of measure concepts, disease-specific. And these are really, they're kind of the standard ones that are kind of out there in the over-diagnosis space.

One is essentially measuring whether you're diagnosing this stuff a lot more frequently than everybody else who's got a similar patient base to you.

So the idea would be sort of case-mix

adjusted peer organizational comparisons, like a health system might have a -- they might be diagnosing prostate cancer, you know, at a population prevalence of five percent. And everybody else is doing it at one percent. And then they'd be an outlier in the percentiles. They'd be in the 90th, you know, or 99th percentile for the prevalence of prostate cancer.

And as long as you had sort of roughly comparable groups or case mix adjustment, which everybody already does all the time in all these institutions around payments, you could have a similar kind of measure there.

And it could include sort of, you know, the disease/illness spectrum of severity if we were concerned that it wasn't just the total but that it was a shift to everything being diagnosed. And it was sort of early stage or not really, you know, not really breast cancer or whatever, that same kind of deal.

And then finally, the last one there, looking at the relationship between incidents and

total morbidity and mortality, again, another sort of over-diagnosis metric that, if you're making a lot of excess diagnoses without benefit, it shows up in the form of unchanged morbidity and mortality but an increase in disease incident. So again, relative to peer organizations having some kind of percentile rank, just to see if you are way far an outlier, essentially.

All right, next group. So then we had diagnostic error. We lumped this into a couple of big themes for ways that one would think about identifying possible or likely diagnostic error-type scenarios.

One is around this unanticipated change in level of care. You can see that there are several measures that are all the sort of same thing. But it's, you know, one set's from primary care to the emergency department, one set's from emergency department to ward, another one says from ward to ICU.

They're all the same, in my view, and

I think the group's view, the same measurement concept, which is that if you got bumped to the level of care, and at the same time you had a change in diagnosis that for the sort of same symptom presentation or problem, that that was potentially a sign that not only had there been a diagnostic error, but potentially that there was some harm associated with it, especially if there was a care escalation.

So this would be the measure concept percent of patients with discharge diagnosis X, where X is benign, subsequently diagnosed with, you know, Disease Y, where Y is dangerous for the same index symptom sign or test result example.

You know, some of the stuff we've done, for instance, with the discharge from the Emergency Department within nine days in a history of men with stroke. But you can do the same thing with any other similar kind of pairing.

And then the percentage of patients harmed by the diagnostic delay is defined above,

just so you get a sense for sort of the magnitude of the harms. And obviously, that would be potentially a little bit harder to measure. And you'd have to, you know, have people systematically looking into that.

But we thought that at least trying to get some sense of whether people were harmed or not from these kinds of problems would be significant.

And the group also wanted to make sure that we covered the reverse, the sort of de-escalation idea that, like, oops, you know, we admitted the patient to the ICU as an MI, and then they were de-escalated to a esophageal spasm within 12 hours. And they went to the ward, and then they went home that afternoon or the next morning. Again, balancing both sides of the coin, you know, we missed important stuff or we over-called stuff that wasn't important.

The second bin here, so the second measurement theme was this idea that an outcome that happened to the patient, either a lost to

follow-up, or adverse events, or explained deaths, might be a marker of misdiagnosis.

And here we tried to get the concept of -- we had these two concepts of percentage of patients with, you know, Symptom A, lost to follow-up prior to a confirmed diagnosis.

So basically, if somebody's sort of, you know, they go to their pediatric clinic, they've got a headache. They get neuroimaging ordered, because somebody's worried. They never follow-up, or it never happens that they're sort of lost to follow-up. You could have, you know, sort of the fraction of those kinds of headache neuroimaging ordered that never followed up or disappeared.

That would be potentially a sign that there were -- that, at the very least, you weren't closing the loop, and at the worst that patients were suffering some adverse consequences.

And the same for the second one, same basic idea, Symptom A, suffering major health

event or death prior to a confirmed diagnosis. And then in particular, obviously, if you have patients who, you know, have then a known cause of death or the adverse event that you can link back to that, we never closed the loop or finished the diagnostic process.

Next one, so the last tab that we were asked to deal with was information interpretation. I think some of this actually harkens back to one of the other discussions about information integration, this issue of sort of reconciliation of conflicting results. There may be someplace to harmonize there.

But the general theme here was there ought to be some way that one is sort of monitoring and managing when Report A says you have cancer, and Report B says you don't have cancer, or whatever, you know, whatever diseases are things those are.

And we wanted both kind of a policy and a procedure in place for identifying and reconciling those discordant interpretations or

findings. You know, like, the radiology diagnosis of brain tumor, and a pathology diagnosis of the biopsy of the lesion as a demyelinating lesion, that there's some way of actually tracking that, and sort of feeding back into the system, and reconciling them, and learning from them.

And then the other two measure concepts there were the sort of fraction of discordant diagnoses that might be resolved through those type of SOPs and a percentage of patients where there was a discordant result of some kind associated with -- where it didn't match their clinical outcome. So, like, percent of patients with clinical normal colonoscopy diagnosed with colon cancer in a short timeframe, or something like that.

The second theme in the information interpretation section was use of decision support. So we suggested two concepts. One is availability of sort of an EHR-integrated, evidence-based decision support pathways for

diagnosis of common symptoms.

The second was the percentage of encounters in which decision aids, those or otherwise, are used. And we suggested three potential ways that that could monitored or measured, either with click tracking on the EHR or using administrative data.

Seeing whether, you know, if somebody's scheduled to be on Pathway X for chest pain that they're always supposed to get an EKG, or troponin, or whatever, looking at the percentage of cases where that actually happened, or surveying people and seeing whether they're following the pathways.

And then finally, the EHR supports high quality diagnosis. And really here the demand was that we go above and beyond the simple idea of kind of meets the minimum standard. Like, there was some goal that it actually worked rather than it met specs.

And I think you'd ultimately need to measure things like that by surveying providers

who use the EHR and things like that. I think you could do it that way. And say, look, does this actually work? You know, do you actually feel like -- And, you know, you think about historically that's been done in some places, like, the people at the Brigham who really had their electronic health record up, you know, pretty much first and had a really nice one.

Were unbelievably upset, you know, and had to go to Epic. And, you Know, they felt the loss. So there is clearly the ability for people to discern between, you know, not so good and good in terms of this.

So those were our main things. We did have a couple of things at the bottom that we thought -- that sort of came up in our discussion, that seemed like they fit in some of the other buckets, that we needed something in there about systematic second review of diagnoses, whether it's radiology, pathology, ten percent, second reads, or whatever. But also the clinic records, the same kind of thing, making

sure that we have the same kind of comparison to a reference standard, not just inter-rater reliability but to some gold standard or whatever, and then making sure that people are available, consultants or diagnostics, you know, radiologists, pathologists, to interpret the results. I think that would show up in some of the other groups.

(Off microphone comments.)

CO-CHAIR MCDONALD: Microphone.

MEMBER NEWMAN-TOKAR: For which one? Tell Vanessa which one to switch.

MEMBER MAHAJA: The first one, go to the top one. And the ones that we use, the IOM definition for including diagnosis. So we use explanation of the patient's problem and then do so that they stayed consistent with that.

CO-CHAIR GARBER: And, David, you had something on specificity here. But it seems that you dropped the sensitivity measure in regard to screening. Did your group not want to include recommendations on screening tests for cancers,

for example?

MEMBER NEWMAN-TOKAR: Which tab was that in?

CO-CHAIR GARBER: Efficiency.

MEMBER NEWMAN-TOKAR: So the question is, yes, I don't think we were entirely -- it wasn't one of the ones that made the adequacy of screening procedures. It wasn't one of the ones that made the group top list. It was actually one of the lower ones. So I don't know if it's an efficiency issue. It's more, to me, I guess it's closer to --

CO-CHAIR GARBER: Yes, I'm not sure where it goes, because I don't think it's mentioned in any of the other tabs.

MEMBER NEWMAN-TOKAR: If it's not mentioned in any of the other tabs --

CO-CHAIR GARBER: So if cancer is the number one condition that's misdiagnosed, shouldn't we have something about trying to address that through appropriate screening?

MEMBER NEWMAN-TOKAR: I do think that

the adequacy of screening for whatever diseases are relevant to the screening issue -- I mean, there are a lot of screening topics, right, there's cancer, there's, you know, neonatal hearing loss screening, there's screening for diabetes, and so on, and so forth. But there are -- yes, I think that the adequacy of screening is something that's worth doing. I just don't know where it fits.

CO-CHAIR GARBER: Please use the microphones.

MEMBER DUNNE: We kind of address that in population-specific testing that takes into account prevalence, disease prevalence, and so forth. So buried within there, I think, are the specifics that are necessary for any kind of screening test. It's just not specifically addressed.

(Off microphone comments.)

MEMBER DUNNE: Well, right under over-diagnosis. I mean, we discussed this in terms of, for example, testing for Lyme, using

screening testing or testing that's appropriate for populations, looking at diagnosis or particular diseases within your particular population relative to others.

CO-CHAIR MCDONALD: So are you saying that maybe the label shouldn't be over-diagnosis, but it should be under and over-diagnosis? Because those kinds of --

MEMBER DUNNE: Sure, yes.

CO-CHAIR MCDONALD: -- assessments could see either side of it.

MEMBER DUNNE: And there was another part too. Let's see. Go down a little bit.

MEMBER NEWMAN-TOKAR: So I think it needs to be there. Like, I mean, I don't know whether it needs to be on this tab or on a different tab. But I do think it needs to be there. There needs to be some issue of, you know, how well we're adhering to guidelines for screening for cancer and other diseases that are --

CO-CHAIR GARBER: Or we could present

the counter argument.

MEMBER SINGH: I strongly recommend that we do not put screening related stuff in this diagnosis related report. It is a different animal, if you will. It is a different concept.

We can't even get diagnosis right when we have obvious signs, and symptoms, and tests of patients who are being misdiagnosed at an alarmingly high rate. Why would we go for asymptomatic populations where the evidence is much weaker about what we're going to do right and what we're going to do wrong?

I mean, look at the controversies around any cancer about screening. There's just very few cancers where the evidence is very strong and the harms from screening are pretty, you know, it's coming out --- And I would encourage you to just look at the recent literature. Some of it has been forwarded to Mark for his consideration. But I don't think the evidence is strong that we should be including in our group. That's what my

suggestion would be.

CO-CHAIR GARBER: David?

MEMBER SEIDENWURM: So I think that there's about a billion metrics out there for uptake of screening. So I don't think that we need to go there.

I think that that there're -- and I am sympathetic to the argument that's just been made regarding the harms of screening, but there is one area that I do think that we should address with respect to screening, or two areas, actually.

One is inadvertent screening which is the phenomenon of incidentaloma. You know, for example, if you do a MRI targeted to the lumbar spine, and you see a lump on the adrenal glands, that's essentially screening for adrenal carcinoma, which no one would advocate as a stand-alone procedure, yet we do it inadvertently.

So I do believe that we should address, you know, incidentaloma/inadvertent

screening. And we should clearly define that as a concept, you know, in this report.

The second thing that I would like to address is, and I do believe that this is one area of screening that should be part of this report, and that is when a screening test is, we'll say, mandated even, right, because, for example, mammography is an ACO metric, you know, a mammography uptake is an ACO metric, we should also have corresponding metrics for the quality of that service.

If colon cancer screening is, I think that's also an ACO metric, then we should have corresponding quality metrics for all of the various modalities that are employed for that purpose, you know, both the optical colonoscopy procedures, the CT, the various fecal tests, and so forth. So I think that -- and I think also with respect to the follow-up pathways for cervical cancer screening.

And I do believe that that should part of our agreement here. So inadvertent screening

and then the adequacy of the screening procedures themselves but not, as Hardeep said, you know, whether someone should be screened for a particular diagnosis. So that's well covered.

MEMBER SINGH: Yes. And I want to qualify it. Abnormal screens needing follow-up is a totally different ball game, because it comes under diagnosis. Because then there is signal to go forward to do something. The person is not asymptomatic anymore which is sort of the whole argument about doing this. My other sort of --

MEMBER SEIDENWURM: Yes, I totally agree.

MEMBER SINGH: Yes. My other sort of comment was has a lot of the language been changed in some of the concepts? Because this is -- a lot of this is new language. And it's -- I don't know, I mean, I was just sort of -- it was hard for me to keep at it and sort of try to understand where things were falling, having done the, you know, the things.

So I'm not sure what we lost and what we gained. A couple of the things I was going to mention, so the EHR one, for instance, and then there was something about over-testing and over-diagnosis. I mean, there's over-diagnosis stuff right here. I don't know how you would measure that, to tell you the truth.

But something you could, you know, do, something which is more measurable, I mean, there is decision support engines to cut down unnecessary testing, for instance, or lab tests was mentioned. So we over-use, you know, and we do want to see lab testing.

Well, there is now additional support that works to prevent that. So can we combine things like over-diagnosis or over-testing in one area and, you know, we know we want better EHRs and sort of combine the clinical decision support and make that more of a measurement concept. Because some of this still appears a little loosey-goosey to me.

MEMBER HUNT: I keep going back and

forth and ping ponging as far as the screening. But I'm persuaded by what Hardeep says. If we think about diagnostic accuracy as the timely explanation of the patient's health problems, then that obviously precludes screening, because they didn't have a -- this speaks diagnostic tests for something, for a problem rather than screening.

MEMBER SHERIDAN: Yes. Just looking at this through a patient perspective, first of all, I don't know how to define over-testing. And then as I was reading this and hearing you, David, I was thinking about -- I reached over to Kathy, and I said what about under-testing. And under-testing of things that are, you know, the standard of care now or embedded in guidelines that still aren't getting done.

And I'm thinking of the jaundice, you know, testing and screening where there is, you know, the whole population, children in the 90's that emerge with kernicterus because of the acts of omissions of necessary tests. So I was

wondering where that fit in here.

MEMBER NEWMAN-TOKAR: Yes. I think that's a good question. So if we just scroll up a little bit to the timeliness, so you can break down the issue of appropriateness of diagnosis in any number of ways.

What we were actually trying to do was keep it simple in the sense that the thing that people ultimately care most about is whether they got a timely, accurate diagnosis, right.

So what we're saying here is that if you didn't get a timely or accurate diagnosis, we don't necessarily care whether it was because the tests were underused, or what percentage of them were underused, or whatever, right. Like, we're just saying the outcome is what matters. The outcome is whether you got a timely diagnosis.

We weren't asking to measure all of the steps along the way. You could, and we mention it there, like, the timeliness of the first key test. Like, did you get to that, you know, first, I guess, in a symptomatic patient

with a potential, you know, with a headache and a potential brain aneurysm. Did you get to that CT scan within such and so many hours, or whatever, or not do it at all, right?

But we have the notion of correct diagnosis and under-diagnosis, if you will, to sort of use a loose term that it counters the over-diagnosis term. It's all in this idea of whether you've had a timely diagnosis of your disease. It's not, as opposed to breaking it out and saying did you get this test, did you not get this test?

Now, you could do all those things too. There's no reason why they can't be measurement concepts. And I'm not averse to including them. But it wasn't that we didn't consider under-diagnosis. We just wrapped it up as you need -- what percentage of the time are you getting a timely, accurate diagnosis? That was the roll-up of getting the right diagnosis.

And the other piece was just to make sure that in the process of encouraging people to

get timely, accurate diagnoses, we didn't just willy-nilly encourage them to, in a profligate way, over use tests and end up with, you know, bad results as a result of over-testing.

That was the logic. It may not resonate, it may feel like we need something that feels a little more analogous to the over-testing. But this is where it is.

MEMBER SHERIDAN: This, in the body of this information here, is there -- and again, from a really simplistic patient point of view, you know, patients, we tend to believe that guidelines are followed and that that's going to keep us, you know, evidence-based guidelines, and that's going to keep us safe and from harm. And we know that a lot of guidelines are not followed.

And is there any such measure of the percentage of clinical guidelines that are followed for certain things that are, you know, misdiagnosed? I mean, something as simple as that.

You know, hospitals do, in their bylaws, I mean, say they, you know, follow the American Academy of Pediatrics guidelines. But do they really? Is there a percentage of that, any measurement that would be something that the lay audience would understand?

MEMBER NEWMAN-TOKAR: You mean is there a measure on this page of that? Yes. So if we scroll down, I don't know if it's something the lay audience would understand or not understand, but we have here, excuse me, where's the -- if we go all the way down to the bottom one, the interpretation, I think, one, so what I guess we don't have adherence to guidelines specifically in there. We have --

(Off microphone comments.)

MEMBER NEWMAN-TOKAR: Well, we have, you know, we have the availability of pathways, we have the percentage of encounters where some kind of decision support was used. But we don't have the percent adherence to guidelines in diagnosis. And I think we could -- that's easy

enough to add. I think we should.

CO-CHAIR GARBER: David, there were a couple other things that seemed important that were missing.

MEMBER NEWMAN-TOKAR: It's very similar to the timeliness idea. But it's different. It's a process measure rather than an outcome measure. Yes?

CO-CHAIR GARBER: A couple other things that seem to be missing, one was in the diagnostic error bucket were measures relating to asking patients and physicians about reporting diagnostic errors. I didn't see anything on that.

And under information interpretation, there seemed to be a lot of value in the kind of trigger tools that are being used by Michael Cantor. I didn't see a measure concept being advanced in regard to that.

MEMBER NEWMAN-TOKAR: So there's no measure concept that specifically has the word trigger tool in it. But that's what all of these

things are.

MEMBER SINGH: Well, I thought there was a trigger one. Where did it go? It was down here?

MEMBER NEWMAN-TOKAR: No, no, no. There is one on the list, but there wasn't one that rose to the level of making it to the document.

CO-CHAIR GARBER: Well, you're not opposed to that concept?

MEMBER NEWMAN-TOKAR: No, I'm not opposed to concept of trigger tools. These are all trigger-based ideas.

MEMBER SINGH: You know, I was just going to emphasize, I think, Sue, what you're referring to is, well, we would imagine that there would be guidelines to make sure that we diagnose patients with colorectal cancer with some timeliness, such as 60 days, or 90 days, or even 180 days.

But oftentimes for diagnosis issues, there's no clear cut guideline as to what to do

next. Because it's all a lot of gray, a lot of gray zones.

So there are a lot of screening related guidelines. But if you look at the guidelines for diagnostic issues, algorithms, protocols, pathways, they are probably just less in number. And they're more, you know, because every patient is different.

So the one patient with abdominal pain, you're going to go this way, but the next patient with abdominal pain, you might go a totally different way. So it all sort of depends.

People are coming up with more. I know there's somebody working on, like, an abdominal pain guideline in the ER. But it's not easy, and they're not a whole lot of them available, sort of my take is.

CO-CHAIR GARBER: We need to move on to Group 3. I'm sorry. Yes, we could talk about this one for a lot longer. We're suggesting that we do Group 3, and then take a break. All right

with people? Okay. Group 3?

(Off microphone comments.)

MEMBER SINGH: Our approach was fairly similar to what Mark was describing. So we looked at all the measures, you know, even the ones that had zeros in it. And then we integrated and consolidated concepts.

Basically, we were trying to look for areas of overlap, and synergy, and we changed a few things. So I'm just going to quickly go through and you can read the rest.

Also the first one was key activities, the diagnosis of the key activities that an organization could do. So we changed the language, and we combined.

So if you'll look, each of these sub-bullets was a different measure. And we ended up combining it to one, or whatever, measure concept. And we made one out of these five or six. So we thought we were being efficient.

So that then becomes examples of the types of things that the organization can do to

measure diagnostic performance. So they need to have an EHR data analytic infrastructure to measure diagnostic performance. So we thought we got all -- we gained a lot of ground without losing anything, keeping all the concepts together and then merging it into a high level measurement concept.

The second measure concept here is, organization has established mechanics of providing feedback. Each of the feedback ones got three or four each. But when you combine them, that's what it became. And it became, like, seven out of -- you know, so it was pretty high. And so we combined that. So that's the second one that, again, was prioritized a lot.

Next one. So then we had a big discussion about, well, you can have all of these infrastructures in place, but unless you're learning from all of this measurement we're not going to gain any ground.

So we thought that the learning part itself deserved a separate measurement concept

away from sort of the, you know, the measurement part earlier and the infrastructure around it.

So this is the one around learning. And then we put language about RCAs, autopsies, someone brought up, very rightly, number of autopsies performed per number of deaths. It's a great measure, but unless you learn from it, there is no value to it at all. And so a similar concept applies to M&M conferences.

Next one. So we had one that gained, I think, three or four words. I can't remember. But it was important enough for us to consider for a later time. And I think somebody else is doing patient access, correct, mixing it?

And so this was important. And actually, it has already come up in the previous discussion. On the recent study for Open Notes, 20 percent of patients actually discovered errors in their own documentation and reported it to the organization.

So building on that concept, I thought that was an important concept. We all thought

that it should be kept. It just wasn't in our area. And we wanted to hand it off to somebody. So whoever is doing patient stuff, should take note for this one.

CO-CHAIR MCDONALD: So, Hardeep, actually, before you go on, maybe we should do one tab at a time. I was noticing last time when we did three tabs and then we had to get comments, it was hard.

MEMBER SINGH: Oh, sure, sure. I can stop.

CO-CHAIR MCDONALD: So maybe, like, stop --

MEMBER SINGH: Yes, yes.

CO-CHAIR MCDONALD: -- stop here, get a few comments, and then do the next tab.

MEMBER NEWMAN-TOKAR: Is leadership engagement anywhere there?

(Off microphone comments.)

MEMBER SINGH: So we had the thought of involving, I think it was actually in a separate one. There was a separate leadership

measure that probably got disconnected. But I thought we actually put it back.

MEMBER RADFORD: Oh, this was part of it. I mean, having the Board get these measures, et cetera. We thought all of that was important. But it was sort of wrapped into organization-wide measurement.

MEMBER NEWMAN-TOKAR: Yes. I think just adding, making sure that leadership is bought it where -- because almost all of the kind of culture, behavior change stuff, ultimately if you don't get leadership buy-in, it doesn't work.

CO-CHAIR MCDONALD: David Hunt?

MEMBER HUNT: Typically I'm going to say the same thing, that this really rolls up to the question of does your organization support a culture of diagnostic performance quality.

CO-CHAIR MCDONALD: Culture and leadership. Mike? Are you -- no, okay. David?

MEMBER SINGH: So I don't know if anybody's taking notes, but I would just say, you know, make sure that we address leadership and

culture amongst one of those bullets in there --

MEMBER SEIDENWURM: There were several things in there in the list of proposed concepts about imaging autopsy. I'm not sure that we want to go too heavy on that. I think that literature, I think those ideas were ahead of the literature.

MEMBER SINGH: Yes, we left them out. They didn't even come to high enough.

(Off microphone comments.)

MEMBER SINGH: Just number of autopsies in general.

MEMBER SEIDENWURM: Sure.

MEMBER SINGH: Yes. Just to make sure that people are still doing them.

CO-CHAIR MCDONALD: And then what about -- so on the RCAs, I know there was one in here that said RCAs, like, with patients involved. And there's the organization, like, learning and getting feedback.

I don't think it's patient engagement. Like, the patient engagement tab to me, or that

thing is more about patients and their own individual care. This QI activity would be more about what the organization's doing. So a handshake with having patients involved in something here seems like it's a missing piece.

MEMBER SINGH: Yes, you could just say something like the organization expert needs to conduct a comprehensive RCA in cases involving diagnostic errors. And these RCAs involve patients when appropriate.

MEMBER SHERIDAN: Hardeep, on that bottom bullet: The organization has an established mechanism providing feedback when there's a significant change in diagnosis. Feedback to whom?

MEMBER SINGH: Well, you know, I think the spirit was supposed to be care teams.

MEMBER SHERIDAN: Yes.

MEMBER SINGH: Providers and care teams. There was a separate one for disclosure to patients that we didn't really go into. It didn't even get enough of a priority. Because

people already have disclosure programs. And we didn't think -- and it was post-error, and we didn't want to sort of mess with that.

MEMBER SHERIDAN: So it's post-error. Okay.

MEMBER SINGH: Yes. This is mostly for care teams and providers. That's what we understood. And that's why --

MEMBER SHERIDAN: So it's not like it's during active care, like, when there's a change in diagnosis or a change in a pathology, that there is a mechanism to change the -- to inform the care team and the patient?

MEMBER SINGH: Yes, yes. Yes, that would be included. Yes. That should be included. If there's a change anywhere, I think -- we didn't think this would specify a real time or, you know, post-one-year feedback. It was just anytime, anytime there's an opportunity for feedback and learning. That's probably what the intent was.

MEMBER SHERIDAN: Okay.

MEMBER SINGH: I mean, we didn't discuss exactly what you're getting to.

MEMBER SHERIDAN: Yes.

MEMBER SINGH: I think I know what you're getting to, yes. But it's included though. It's sort of encompassed in this.

Okay, so two stood out from the external environment very strongly. We changed the language for the second one. The first one you might be familiar with. I don't think we changed the language for the first one. But in the second one, we changed -- we combined the two that stood out.

And we basically said payment incentives should be aligned to promote timely and correct diagnosis. And that basically was trying to get to the external environment piece of it. There was a measure concept already in there. And we just changed the language a little bit.

In access, we changed some language. This wait time was already there, stratified by

specialists. It stood out quite, I think it was actually the number one.

The second two is what we talked about a lot in our discussions. We changed this a bunch. Availability of rapid or point of care testing for critical diagnostic decision making. It was in a way that was not phrased correctly. And we had people in our team -- and feel free to jump in, everybody from our group, to give your input as well.

And then there was one about having advanced imaging and lab available. And we changed to access to appropriate testing for the most common conditions. And in fact, some of you who may have been on the call in March, this was discussed on the call quite a bit as well.

The measure concept, at that time, said there should be access to advanced imaging and lab. And we didn't know that meant. So we changed it to access to appropriate testing for the most common conditions that you see that are most relevant to you.

MEMBER MAHAJA: I just think, you know, just from the ER perspective, I keep on looking at it, patient-reported communication isn't getting calls back. Like, from the patient perspective, would that make any --

MEMBER SINGH: Was that there? Did it reach up to --

MEMBER MAHAJA: Patient access, right under patient access. Like, from the patient's perspective, the ease of --

MEMBER SINGH: Yes. I think it didn't raise up to the --

MEMBER MAHAJA: It didn't --

MEMBER SINGH: Can you read it, because I think it probably didn't raise up to the --

MEMBER MAHAJA: So I think it's Number 7: patient-reported communication ease in getting called back by medical team when reporting concern.

MEMBER SINGH: Yes. Actually, I think this was an important one that -- I actually

ranked it. But it didn't reach up. But if you think it's --

CO-CHAIR MCDONALD: I had that ranked too and --

MEMBER SINGH: -- one to include, lets' just add it.

(Simultaneous speaking.)

CO-CHAIR MCDONALD: Yes, you can add in.

MEMBER SINGH: Let's just add it.

CO-CHAIR MCDONALD: Yes.

MEMBER SINGH: It's the number --

MEMBER MAHAJA: Number 7.

MEMBER SINGH: Number 7. We should include it. Anybody from our group feel differently?

(No audible response.)

MEMBER SINGH: I think it just didn't get enough. Because, you know, it got one, because I know I gave it a one.

CO-CHAIR MCDONALD: And I know I had marked it, but I didn't get mine all the way

through. I didn't get --

CO-CHAIR GARBER: Could we all --

CO-CHAIR MCDONALD: I knew you had one. You have more votes for it.

MEMBER SINGH: Oh, we had more votes than we think? Oh, okay, that makes sense.

(Off microphone comments.)

CO-CHAIR GARBER: Could we also talk about Number 10, that patients have a long enough appointment if you're a new patient? A lot of doctors say that's the number one problem. They just don't have enough time.

MEMBER SINGH: Long enough for the new patients?

Yes. We were asked to, you know, keep it to three. I mean, we're going to make it to five. Is that okay with the NQF team?

CO-CHAIR MCDONALD: Five's okay, I think, from our vote and this conversation, yes, that --

(Simultaneous speaking.)

MEMBER SINGH: No, I think I marked

that one too, actually. It was a good one.

CO-CHAIR MCDONALD: Yes. Let's add those two. I mean, the comment was that each of these tabs, you know, just because there's only a few doesn't mean they're not -- they should --

MEMBER SINGH: Yes.

CO-CHAIR MCDONALD: Yes.

MEMBER SINGH: Okay?

MEMBER SHERIDAN: Can I just ask a question about the payment model slide? Can we go back to that, payment? So, payment incentives should be aligned to promote timely and correct diagnosis. Physician payment form must recognize the importance of frontline diagnosis --

MEMBER SINGH: It's quite high level.

CO-CHAIR MCDONALD: Yes. It's very high level. But I'm wondering if we should add a word, like something that -- a collaborative or something. Because, you know, it's something that -- I've been engaged in this dialogue about ensuring that pathologists and ordering physicians get reimbursed for the time talking to

each other. And right now, that's not in payment bundles.

So I don't know if it gets embedded in that, some type of payments that should be aligned to promote collaborative, timely, and correct diagnosis, or something to -- it's not really a second opinion when you --

MEMBER SINGH: So team-based? You know what I'm saying?

CO-CHAIR MCDONALD: Yes, team-team.

MEMBER SINGH: Team-based, timely, and correct diagnosis?

CO-CHAIR MCDONALD: Yes.

MEMBER SINGH: Yes. I think we have another one that talks about teams. I don't know whether --

CO-CHAIR MCDONALD: In the payment -- yes.

MEMBER SINGH: Okay, yes. That's good.

CO-CHAIR MCDONALD: Yes. And that's part of the IOM report, if that was, like, you

know, the --

MEMBER SINGH: Okay.

MEMBER NEWMAN-TOKAR: So could I just ask on this slide, these are obviously important topics, a little like mom and apple pie, but what did you envision as the actual measures for something like, payment incentives should be aligned to promote timely and correct diagnosis?

MEMBER SINGH: So I think it's also sort of reflected on maybe another -- the workforce one. You know, right now we're rewarding quantity and not quality. So we didn't actually come with a specific measurement or measure area.

This was one of the ones that's bordering on something that needs to be done, but exactly what we're not sure. And so when we started having a little bit of discussion, Andrew said keep that discussion for the sort of the next stage rather than, you know, hashing it out right now.

MEMBER NEWMAN-TOKAR: I do think this

is one where you could potentially get -- you could survey providers. I mean, providers might be able to answer these two questions, and it might be a valid way of approaching the issue of whether they feel, in a generic sense, they're being supported both legally and payment-wise.

Could I just see the one more slide forward. Oh, that's what it was. It was just brought up, this issue of the length of the appointment. And that's one I would just make sure we're careful about.

I think everybody agrees that it's impossible to provide high quality diagnostic care unless you have a sufficient amount of time to spend with the patient. And it's certainly an important complaint.

There's a lot that's been written about visit lengths. And over the last three decades, the visit lengths have actually increased by about one minute on average in primary care clinics. But the face time has dropped by 50 percent or whatever it is.

So we have to make sure that, when we go there, that we've gone there in a way that's sensible. If people are, you know, if the way people are doing better diagnosis is by changing the care delivery model, and then happens to shorten the apparent visit length but it increases the face time, or whatever, we have to be careful about not incentivizing against creative care approaches that actually improve diagnosis and maybe even improve the time with patients but don't look that way when you do the math.

That's a very dangerous measure, in my mind. It's not that it's not important. It is. But I think we have to be really careful when we go there.

MEMBER SINGH: So I think we have -- if you look at the next one, people are going to get a little --

CO-CHAIR MCDONALD: Actually, so wait, though, before you go there. Because we've got Lavinia, and Tom, and Martha. And then we'll

just have you go on. So, Lavinia?

MEMBER MIDDLETON: Just on the next slide, I think the word we were looking for perhaps is integrated diagnosis. Because that's initially what the IOM report states and, I think, is best care integrating the pathology, the radiology, any other complex molecular tests that are there, or clinical symptoms brought in from the patient.

So integrating all of the available information in order to create the diagnosis and having time to do that, either with the patient or outside the patient room is, I think, where we're trying to go.

CO-CHAIR MCDONALD: Tom?

MEMBER SEQUIST: So I just wanted to, I guess, second what you were saying about the visit times. We used to measure that, sort of using EHR data. And it's really fascinating data, but at the end of the day, we weren't really sure what to do with it after a while.

I mean, we'd see things, I would see

patterns, like, not to pick on specialties, but you would see, like, dermatology office visit lengths. And you would look at it, and I would say that can't be safe for some people who were outliers. But I don't know to -- I didn't know that. I just sort of was, like, about --

PARTICIPANT: They were 30 seconds.

MEMBER SEQUIST: Yes. When you're less than 10 seconds, it couldn't have been a good full exam. But it never felt right, it never worked as a way to assess whether the diagnostic quality was good. So we sort of abandoned it for -- and I just wanted to sort of second. I agree that I don't know about that particular one.

CO-CHAIR MCDONALD: Martha?

MEMBER RADFORD: Yes, just a couple of things. First, on that last comment, we were most concerned about, you know, two mills.

(Off microphone comments.)

MEMBER RADFORD: Mills. You know, and trying to figure out the outlier physicians who

were seeing, you like, 100 patients a day or something that you can't possibly make the right diagnoses in that period of time.

And we also acknowledged that this is specialty-specific, so things like dermatology might have a different profile. So I think that that's the concept we were trying to attack first.

I just want to also say something about somebody said, well, how would you measure, you know, the payment for making -- the teamwork payment for making a diagnosis. And I'd say you make it a measure of the health plan. Does the health plan reimburse for it? And that's, you know, fairly simple. You just look at their reimbursement rules and figure it out. So that's how I would address that.

CO-CHAIR MCDONALD: Okay, now there's a few more cards. But, Hardeep, why don't you go on and do the rest of your assignment so --

MEMBER SINGH: Yes, because a lot of what you're talking about is actually addressed

in the workforce slide, which some of you might like, some of you may not like. So let's -- have we finished patient access? We have, right? Okay.

So workforce one, we spent a lot of time on this one, especially on the ones that are red. So the first one was we were trying to make this more actionable, this, providers operate at the top of their license.

It did get a lot of votes, by the way, you know, sort of at the top, to free up cognitive liability. We thought it was important. This is where the team stuff that you all were talking about, is where what we discussed. You know, how do you make sure that the teamwork principles, and there's, whatever, eight or nine them, are integrated.

And so this is where I think we're going to have to think more about this. And we, obviously, in the short time we had, we didn't do that. But I think this is really, really important. How do leverage teams? How do you

leverage a nurse? How do you leverage the other specialties, other patients? So this is one of the ones that needs some development.

The second one is, I think we're getting close to what you were talking about. Do we have adequate time to gather, integrate, interpret all the data that we need? And you could say providers are care teams, for that matter.

And this is where the face to face time and the non-face to face time came up as well. We spend a lot of time taking care of patients through the EHR, you know. So it's just not the visit length. I agree with Tom, I'm not sure if it's just the visit length.

But then there's one that we -- I want to just highlight, because it's related. If you look at the bottom one, we did think it's not safe. If you're an internist, and if you're seeing more than 50 patients a day, there's something wrong.

MEMBER SEQUIST: That's back to that

one that I was talking about. It definitely feels wrong.

MEMBER SINGH: And so we didn't know what we would do, but we wanted to put it out there. Because we discussed this a lot. And that's why we said, okay, you know, it could be specialty-specific.

And what we wanted to do was to change the thinking that we don't want to award quantity. We want to think of something else. Because that's what's rewarding quantity. The reason they're seeing 100 patients or 75 patients a day is because that's how they make money.

And we want to change that thinking. I thought some area of measurement in that sort of thematic area would be really important.

So this is what we best came up with in the 30 minutes we had. Identification of potential outliers related to the number of patients encounter per day.

Let me finish, and then we can take more questions. We thought burnout was important

-- no, no, no, I didn't finish. That's why I dropped down. If you look at Number 3, rate of burnout, there's measures to do that. There are surveys, everything, so people are already doing it. The radiologists have already stood out. We talked about that.

Here's the next new one that we came up with: diagnostic performance is included in professional practice evaluation for credentialing and re-credentialing of clinical providers.

We wanted to see if we could get to -- you know, frankly, some physicians who may not be ready to practice anymore, or may be close to retirement, or something close to that, and we were, like, how do we identify these people? And so that's the best, most diplomatic way we thought we could do that.

The last one is, I think, maybe just an additional Number 8. If you go to the next slide. Oh, yes. David, do you want to talk about vacancy rate? Because this is the best lab

and primary care thing we came up.

MEMBER GRENACHE: Right. So I was the only one in my group that advocated to have this on here. But I seemed to convince the others. So this started because one of the proposed measures was vacancy rate in clinical lab personnel.

So, you know, the laboratory's a source of quite a bit of diagnostic data. Many people in healthcare outside of pathology are unaware that there's a shortage of clinical lab personnel, up to projections of 40 percent vacancy rates in the coming five years. Yes, yes. So tremendous job growth, you know. And I want us to go into it. But it's going to be a problem.

And so addressing that is important. But then we ended up talking about, well okay, it's an easy thing to measure. But how does it improve diagnosis, right? How does it make things more accurate?

So one of the concepts here was that

awareness could raise -- could be addressed by, oh, you know, healthcare systems recognize, hey, we don't have enough people working in our lab. We should put incentives in place to assure adequate staffing. And then it expanded beyond just a focus on labs to talk about other specialties and primary care physicians.

MEMBER SHERIDAN: Dave, I noticed that Number 3: attending staff are onsite to supervise trainees 24/7, didn't make it. And I'm going to channel Helen Haskell here, because I think she would probably bring this up.

And I'm also bringing it up, also in light of the ACGME going back to the 16-hour resident hours, or going back up to 28 hours from 16. I'm just wondering if this is something that we should consider?

MEMBER SINGH: WE talked about it briefly. Martha, you can help me here a little bit. I mean, one of this is -- this is the game. All the second is, you know, oftentimes it's not just the attending is not there, but it's the

resident never bothered to call the attending and doesn't that they need an attending.

It's unbelievable to see the calibration of -- sorry, I have to sort of put my personal views up here -- but some of the residents that are coming out are very confident. They don't think they need the attendings.

And so I don't know if having that is useful or not. You could debate. We did discuss it. And we thought it was useful, but then we decided, Martha, you --

MEMBER RADFORD: Yes. This is kind of an important area. But I don't think that's a good measure of it or for it. It's really a process measure.

Honestly, I think it needs to be -- we're really focused on a lot of process measures here if you look at it, and structure. The outcome is, you know, timely and accurate diagnosis or lack thereof, for the safety aspect of it.

And we addressed this actually in our

health system through looking at the number of codes outside the ICU that were not preceded by a response team call, for example.

I mean, so that's looking at the potential failure to diagnose because of this failure to interact with the senior people on the team. We're a teaching hospital. So I personally don't like anything having to do with duty hours. I think that they're terrible measures of anything. And I'd like to look at the outcomes myself.

CO-CHAIR MCDONALD: We have more -- so, Lavinia, did you have something? Or that's up, okay. David, I've seen you put that up for a while.

MEMBER SEIDENWURM: I wanted to get to the idea of too many patients per day. I think that that's something that we have to be very careful about. Severity adjusting or something.

Because if someone said, if an intern sees 50 patients in day --

PARTICIPANT: I did.

MEMBER SEIDENWURM: Okay.

PARTICIPANT: Huge responsibility.

MEMBER SEIDENWURM: Right. So it depends on what they're seeing them for.

You know, if they are very simple one problem quick things or if a radiologist is looking at a 100-lymphoma follow-up, CT abdomen, chest abdomen, pelvis CTs in a day, that's too many. But if they're looking at a hundred ankle films in a day, we'd fire them. You know what I mean?

For underperformance. So you have to be very careful.

CO-CHAIR MCDONALD: Yes.

MEMBER SINGH: I think we could make it easily 60 a day. Or 70 a day. Pick a number.

But I think we're going to have to, and you can adjust based on the type of visit. If you have 9921, is it 99213 still or whatever, the middle one? If you're having, because it's supposed to be for, I think a 15 to 20-minute visit, and you can say -- you can tie this to

other stuff too. So, the other thing we discussed is, some of these measures were, this looks important, but it looks important if something else is also not going right.

And so, if you've got lots of patient complaints, and if you're one of those more than 50 patients a day kind of guy, we'll, we got to look at you. Just to see what you're doing.

So that sort of was the intent. Or you didn't pass your OPPE for instance. Which is the evaluation.

MEMBER NEWMAN-TOKER: So a few comments on this, and the other things we've been discussing.

So, I'd be careful about putting too many numbers on exactly the number of visits, the number of people seen.

I do think that it's right to talk about looking at whether people who are underperforming on outcomes, like they're getting the diagnosis wrong, the diagnoses are delayed, to ask the question, what is the distribution of

their number of visits and how much time are they spending with patients? That's a great question to ask as a causal factor in predicting, is that the reason why they're getting the diagnoses wrong?

But I think it's really treacherous to just looking in an unadorned way at the number of visits or the time of the visit. If people are getting it right and doing it in two seconds, because they're retina specialists who are just looking at choroidal melanomas all day and whatever it is, that they're just really, really good at it, they should be encouraged to get it right in a very short amount of time.

At some level, we shouldn't be discouraging them from moving quickly. If they're getting it right consistently.

What we really need to be measuring is whether they're getting it right or not. And then we can argue about whether they're overworking or seeing patients too quickly or not spending enough time.

Second thing is, on a couple of these things, I think it would be helpful if we made a measure concept that was a less specific. Like the radiologists are available 24/7 to read stat diagnostic imaging studies in real time, seems like a good measure, but it seems like there are others in that measurement concept umbrella.

Like you ought to be able to get a stat CBC read in your lab or whatever else you might need. Like, the idea that you have stat access to diagnostic services that are ordered that way I think is important --

MEMBER SINGH: We have that too, by the way. It's in a previous slide.

MEMBER NEWMAN-TOKER: Okay. So, and the other thing I would suggest is, around the burnout, another potential measure there, other than the survey-based measures, which I think are good, would be actually to look at turnover numbers. Which would be another surrogate for that same concept, but that's very easily operationalized and one that is probably the

strongest indicator of how bad things are.

MEMBER SINGH: Yes, we could change it to rate a physician -- or other measures of turnover. Just for the record.

CO-CHAIR MCDONALD: Okay then, Prashant and then Tom.

MEMBER MAHAJAN: So I just want to nuance the last point. You know, from the 50 patients. And I agree, specialty-specific.

But the way I was interpreting that is when -- we should have some measure where we should be able to put a hard stop where the provider is not further exposed to more workload. Like cognitive workload.

So again, going back to my specialty, you just keep on seeing a lot of patients. And you see by the end of your shift, you are either overtesting, or the patients staying there for a longer time; you're not giving a proper explanation. So I looked at it from a different perspective.

And they have that for the floor,

right? They have a cap for inpatients. So that is how I was looking at it. So I think there is a value in having a hard stop.

MEMBER SEQUIST: So I had, recognizing I had initially said the number of visits worried me, it didn't worry me because I think it was dangerous. It worried me because I wasn't sure where to draw the line. Seventy-five patients for an internist is too many patients.

Even, I can't imagine any internist who is seeing patients at an acuity level low enough that it's okay to see 75 patients.

I just want to sort of clarify two things in my mind. So one is, we're not talking about comparing ophthalmologist to internist. I mean, ophthalmologists would all be compared to themselves, and I recognize that dermatologist can do visits faster than an internist because they're doing something different.

So you're comparing dermatologists to dermatologists. Which inherently controls for a lot of the confounding we've been talking about.

But the other thing that, that it's not to me, the reason why we were looking at the number of visits was not just about the amount of diagnostic time during the encounter; it's really interesting to look at the diagnostic pitfalls that occur downstream of having high volume.

So the providers who see lots of patients, not surprisingly, generate twofold more lab test results that need follow-up. And we know, not in the workforce section, but the follow-up of lab tests is a giant diagnostic error problem. And you just can't manage that information flow.

So it's not just about the time in the office with the patient and whether or not you're spending enough time to get the diagnosis. It's just the high volume.

If we have providers who are on average reviewing 85 test results per session or, not per session, oh my gosh, but let's say per week, with a result being a whole panel of labs, right, that it just, it creates all these other

downstream problems.

So I guess I just want -- so I'm not saying that I am against the concept in general. I don't know how to implement it. But I believe that the concept is measuring something real. That it's not safe to see high outlier volumes of patients over time.

And just like you said, that is why we have inpatient teams where we cap you at 15 patients or 18 patients. Not because we just sort of like want to be real nice in the hospital to the hospitalists. It's because it's not safe ultimately.

MEMBER HUNT: Well first, I always, my derm friends always said they see so many because they are so much smarter than I am. That's what they always tell me.

MEMBER HUNT: So much richer, too.

MEMBER SEQUIST: They made the smart choice in medical school.

MEMBER HUNT: That is true. But I do want to put out a couple of things. It's sounds

as though we're revolving around the concept of like the batting average. It's not just how many times you're up to bat, but how many times that you actually get a hit.

So if you're seeing a large volume and your diagnostic accuracy rate is relatively low, then that combined. So it's a bit of a composite measure I think that we want to begin to think about.

But finally, I don't want to completely dismiss the idea of total time, because there are other fields, gastroenterology, colonoscopy, time, that is actually a measure. How quickly or how long it's taking you to do that colonoscopy.

MEMBER SEQUIST: Yes, I wasn't saying, it's not just about the time --

MEMBER HUNT: yes.

MEMBER SEQUIST: -- you have a downstream implication of seeing high volume.

MEMBER HUNT: Exactly.

CO-CHAIR MCDONALD: So we're about

ready for a break, but go ahead.

MEMBER NEWMAN-TOKER: One small point on this issue of specialty-specific. So, an ophthalmologist is not an ophthalmologist is not an ophthalmologist. And a retina specialist isn't a retina specialist is not a retina specialist.

So it has to do with scope of practice more than it has to do with what the name is on your discipline. Because if you're a retina person, we had a guy who all he saw was retinal choroidal melanomas questions. It was either it was a nevus or it was a melanoma, and that was all he did, all day, all the time. It's a very different life in terms of diagnosis than an ophthalmologist or even another retina specialist who is dealing with medical retina, and they're trying to figure out whether the patient has central serous retinopathy and all this other, you've got a list of a thousand things it might be.

So I think, even within the specialty-

specific idea we have to be a little bit nuanced about the notion of, what's too much volume. I mean, obviously, there is some point at which it is impossible to see the patients. And you could drive the number up as high as you want, and eventually you'll get to a place where everyone agrees it's too many.

But I do think it's harder than it looks to just do that on a specialty by a specialty basis.

MEMBER SINGH: Well I know. And again, none of this is easy, and a lot of this has never been done before, but the point is, the key word out there is potential outliers.

You've got to be an outlier. So you could be in the 5th percentile, 95th, whatever you want to call it.

And all that means is somebody needs to look at these people in context of other things. Rather it's a downstream lab testing or lab follow-up or patient complaints or some other measure. And I think that's the point, the

spirit behind the measure concept.

CO-CHAIR MCDONALD: Okay, so we have one group left, but we're going to take a break first. I think we have to figure out how much of a break.

Ten minutes? Ten minutes. Run outside, get warm, run back.

(Whereupon, the above-entitled matter went off the record at 3:17 p.m. and resumed at 3:35 p.m.)

CO-CHAIR MCDONALD: Come on back. And we're going to be having Group 4 report back. Jen is going to report back for our group. There is our group. And you can take it away, Jen.

CO-CHAIR GRABER: And the goal is to be done by 4:00.

CO-CHAIR MCDONALD: Yes. So we have 25 minutes.

MEMBER CAMPISANO: I think I can do that. Okay, so the first tab that we tackled was patient engagement.

We decided to keep three measure

concepts. Number 2: patients understand actions they can take to improve diagnostic performance. Number 6: information on red flags provided to patient. For example, included in after visit summaries, discharge summaries. Number 10: whether the organization has a documentation system that captures informal caregivers' roles for each patient, and do they reconcile it with the patient and their caregivers, at some interval or every encounter or on any regular basis.

And then we came up with a couple of new consolidated concepts. So we took some of the proposed measurements that were on there and combined them and maybe tweaked them a little bit.

So the first one was timely patient access to medical record, including test results, in and out of hospital, and available to the patient electronically or otherwise. And then process to assure that diagnosis and diagnostic information is communicated in an understandable

manner to the patient. So without jargon.

Do you guys want to talk about this tab before I move on to the next one, or are those pretty clear? Okay.

CO-CHAIR MCDONALD: Yes.

MEMBER CAMPISANO: Oh.

MEMBER NEWMAN-TOKER: Just quickly. I think Number 6 is really important. I think we need to sort of improve the wording of it --

MEMBER CAMPISANO: Sure.

MEMBER NEWMAN-TOKER: -- so it clarifies. It was sort of like the same thing as the Number 5: percentage of patients with presumed benign condition given explicit instructions for how to recognize dangerous symptoms should their condition evolve or something.

Like, I think we need more meat on the bone there, but I think that's a really important thing.

They basically need to know what to expect and know that if things go off script, it

may be not because their treatment isn't right, but because their diagnosis isn't right. And that's such a critical thing that I think most patients don't leave with.

MEMBER CAMPISANO: A clear understanding of.

CO-CHAIR MCDONALD: I'll just add to that. That was one that didn't receive that many votes in the broad group, but we talked about it and wanted it. So yes.

MEMBER CAMPISANO: Yes. Okay.

MEMBER HRAVNAK: I had a question about the health literacy stuff. Because there were a few health literacy things originally, and it looks like they fell off. Or are you consolidating them under something else?

MEMBER CAMPISANO: I believe that they were consolidated. And if we can go to the next slide?

MEMBER HRAVNAK: Understandable.

MEMBER CAMPISANO: In understandable manner. So this process to assure that this

information is communicated in an understandable manner to the patient, with jargon-free communication.

We wanted to take, one of the things that we discussed was taking the onus of literacy off of the patient, because there can be very health-literate patients who still don't understand medical language, per se. And that they want -- that we wanted to ensure that basically this information is being communicated in a way that is understandable to an average patient.

MEMBER HRAVNAK: I don't disagree with that, but I just worry about us losing the health literacy. I mean, it's just so well-established in the evidence-based literature. There are measures, actually, to measure it. And it's pretty finite.

CO-CHAIR MCDONALD: Do you think it's been linked with the diagnostic side of medicine very well?

MEMBER HRAVNAK: I don't know about

the diagnostic side as much as like follow-up and people being able to follow their post-discharge instructions and things like that. But, I think that it could be, but I just don't think it's been looked at.

If people -- the level of health literacy. And you can explain something crystal clear and free of jargon at an 8th grade level of printed material and so on, and yet, just based on people's health literacy and experience, they still won't get it.

So I just think, if you say that it's jargon-free or understandable, it still doesn't get to that piece of that own patient's individual level.

CO-CHAIR MCDONALD: I mean, my thought on that is that there's a little tension here. Because even health literacy is almost a jargony thing. So from a patient-, sort of, facing perspective, the jargon-free is more helpful.

I think you're right though, from the perspective that there's a stream of research,

there's a stream of activities in the healthcare delivery system that try to pay attention to health literacy, it probably makes sense for us to sort of incorporate that in some way. As a sort of supplementary piece of this, but not take away understandable, the understandable piece.

MEMBER HRAVNAK: I'm pretty sure it's like a mesh term. And if you search on that in PubMed, or whatever, you will get a body of literature surrounding that.

MEMBER CAMPISANO: Yes, I don't personally have a problem including it in some way. I just, I think what Sue and I were getting at was, again, taking the burden off of the patient. For having to be the one to learn, to come up to speed to learn the certain terms to decipher their own diagnosis.

CO-CHAIR MCDONALD: I would just add to that, too, that in some ways a person can be fairly health-literate in one domain, based on having had a lot of activity, and then they like move to different domain where there's this new

diagnostic journey.

And I just don't know if the research is there that says how it can be different when an individual -- is it a static health literacy ability? So we should look at that a little. I mean, it should be looked at some more from a diagnostic lens.

MEMBER HRAVNAK: If we're losing something by leaving it off the table.

MR. EPNER: Just as part of the group, so I'll intrude for a second if you will let me. I think the concern was that, if you switch the discuss too aggressively to health literacy, you have to write at the 5th grade level or something, it becomes, you check a box, and you don't know if the patient understands.

So we were at a very high-level trying to, recognizing that the measures people will have to operationalize it later, but trying to make sure that the issue is understanding, rather than hitting a certain level of literacy.

CO-CHAIR GRABER: So, several of us

have written about the Rory Staunton case. So this was a 12 year old boy who was seen in the ER, told he had a benign gastroenteritis, but really was septic.

And part of the problem was, his parents realized he was deteriorating at home, but didn't know how to get back into the healthcare system. So I'm wondering if we need some concept measure that would deal with ensuring that follow-up is facilitated and that you can get back in to the healthcare system if things aren't going well.

And it may be related to the patient access domain, which we already talked about, or maybe it could be here. But I don't see it in either place.

CO-CHAIR MCDONALD: Tom, are you waiting?

MEMBER SEQUIST: Yes. So I had a comment on the first bullet, but I had a comment on both bullets. You didn't go over the first bullet yet, did you?

CO-CHAIR MCDONALD: I read it.

MEMBER SEQUIST: Oh, okay. So two comments. One quickly is maybe the wording is around -- so the health literacy one in the second bullet, I worry a little bit about us setting too high of a bar, I can't believe I'm going to say this, but setting too high of a bar in the evidence base. Because there is -- just this space in general of diagnostic error is hard to study. And health literacy is clearly linked to so many health outcomes.

But so I wonder if what we're looking for is that you somehow confirmed an understanding on the part of your patient. Regardless of their health literacy or any other communication barriers that may have.

But to me it's sort of, what's lacking in that statement is that we have confirmation that they get it.

The first bullet, I guess I wanted to put -- get folks input on, obviously, the patient should have access to the medical record. It's

their record.

You emphasize test results, which I think are somewhat important and the diagnostic process. I think they should have access to the note.

And so these initiatives around note transparency, which everyone has access to, but you just have to go to your medical records office and make a request.

If I were to emphasize something in that statement, I would like to emphasize access to the note. Because that's where the physician is thinking out loud, right? And that would be, I think, helpful in the diagnostic process.

Test results are important, and you want to know if you have a lung nodule on a chest x-ray.

But I would like for us, thinking forward, to emphasize that people should be reading their notes. Because then they're reading the mind of their, not in a weird way, but reading the mind of their physician, that

that's --

CO-CHAIR MCDONALD: That's a good point. We actually did want it to be expanded. It was just we didn't want to limit it to test results, so that's why we said including. It was coming off of the concept that was on the list. But the notes, really good call out too.

On the understandable, this one gets at the provision of something that's understandable. We have in the next batch I think, or maybe we had it in our gap, the receipt in the -- but it was actually -- it arrived and was understood. So we parsed that. I think it's coming in another tab.

I think we have to kind of keep moving to keep to the 4:00.

MEMBER CAMPISANO: Sue, do you want to take over or do you want to --

MEMBER SHERIDAN: Oh, you're already there.

MEMBER CAMPISANO: Yes.

MEMBER SHERIDAN: Yes. Great. Sorry

about my absence here, but ironically, I was out in the hall talking to CMS about the importance of Open Notes. And so I walked in the room and you made the comment about notes.

And I don't know if there is evidence to show that having access to notes reduces diagnostic errors, but if any of you have it, please sent it my way.

MEMBER SEQUIST: There was an article that was published in the, I think it was either in the Joint Commission Journal for Patient Safety or in the Journal for Patient Safety, where it was a qualitative study by folks at the, I think the BI guys in Harborview. Where they present a lot of great case examples of it.

MEMBER SHERIDAN: Yes.

MEMBER SEQUIST: But there's not been a quantitative study of it.

MEMBER SHERIDAN: We could talk later. And any kind of evidence would be great.

Okay, so we're talking about, now, patient experience. These are two measure

concepts that -- patient-reported experience of diagnostic care, were problems explained, et cetera. And patient-reported understanding of diagnoses.

What's our, are we just going through all the slides now? Continue?

Okay. So the new consolidated concepts, and I think this is, I'm refreshing my memory, but I think this is where we collapse several of the proposed recommendations. You know, package them all together.

That patient satisfaction within the diagnostic process, including the decision of the diagnoses, e.g. patient had the opportunity to give input to the process.

We had quite a biter of dialogue, actually, around this. In that we wanted to make sure that if patients were questioning the diagnoses, that they had the opportunity to give input, give additional information.

We can go to the next slide. Are we asking for comments after each slide?

CO-CHAIR MCDONALD: After each batch.

MEMBER SHERIDAN: Yes.

CO-CHAIR MCDONALD: So this is one of our domains --

MEMBER SHERIDAN: Yes.

CO-CHAIR MCDONALD: -- so go ahead on this one. If people have --

MEMBER SHERIDAN: Any comments on that?

CO-CHAIR MCDONALD: And our last one is follow-up, which is a little different.

MEMBER DUNNE: I can wait until we're finished.

CO-CHAIR MCDONALD: No, go ahead.

MEMBER DUNNE: All right. My question regards bias in terms of patient reporting satisfaction or happiness with the whole process.

There are going to be a number patients who don't feel enabled to critics the healthcare system and/or a physician. And there's no way of knowing whether they're being silent, and they're unhappy with the process or

whether they were just pleased as punch.

So there is inherent bias into that system, and I don't know you can get rid of that.

MR. HENRIKSEN: One thought, looking at the wording, regarding the decision of the diagnosis, if a patient, you know, the process could be quite satisfactory. But if the decision is related to the outcome, and if the outcome is terminal disease, I'm not sure one can be very satisfied with the decision. And so just the wording, how that gets worded, would probably need to be adjusted.

CO-CHAIR MCDONALD: Good points. I guess, any other comments on those? Yes, David. Sorry.

MEMBER SEIDENWURM: Well, two things. I mean patients with serious diagnoses are often the most satisfied with their care, so I'm not sure that would necessarily be a problem. And Fenton, I think, was the author of a paper about that a couple of years ago.

Was this where you, did you guys

address the measure concept around, was the juice worth the squeeze? Was the diagnostic process worth the effort? Is that where this concept was dealt with?

CO-CHAIR MCDONALD: It was on the list of like the cost of the diagnosis. It was on this list.

MEMBER SEIDENWURM: And time to diagnoses, I mean, is that where this is?

CO-CHAIR MCDONALD: It was on this list, but we didn't --

MEMBER HRAVNAK: It didn't get much attention.

MEMBER SEIDENWURM: Oh.

CO-CHAIR MCDONALD: Yes. It hadn't got a lot of attention from the group nor had any of us put -- so would you like to make a plug for it?

MEMBER SEIDENWURM: I think it's interesting.

CO-CHAIR MCDONALD: Like more of a plug for it?

**MEMBER SEIDENWURM: I think that one of the things that we see is the diagnostic cascade or the diagnostic vortex. And we see patients get sort of sucked into these kinds of processes only to emerge at the end.**

**Now, some of them are paradoxically quite satisfied with this. They think that their doctors were very concerned and effective individuals who have done a good job for them.**

**I mean, I remember one time, I made a mistake on a film and then I got a call back from the, you know. We followed it up, and it was nothing, of course, but it shouldn't have been called back to begin with, and the patient was happy.**

**The doctor called me back to tell me what a wonderful radiologist I was. And I looked at the pictures again, and I had completely screwed up. So there is that side of it.**

**And then there's the other side where the patient has gone through a valid diagnostic process and sometimes isn't satisfied with it.**

But more is down the middle where we see a patient with some incidental finding, and they go into a two-year process, a follow-up, for something that's a one in a thousand, one in two thousand event.

And so I think that if we can get at some of that, from the patient's point of view, not the least of which we could probably bankrupt them, it would be of interest, I think, to our constituents in this process.

MEMBER HRAVNAK: I just had a question. I was curious about the HCAHPS information, because I didn't see that it was reiterated here. And I was just wondering what the conversation was.

MEMBER SHERIDAN: Yes. If you really read through the HCAHPS it doesn't really address. You can have a wrong diagnosis and die from it, and the HCAHPS doesn't capture that kind of information. So we thought that that wasn't really the right tool to capture diagnostic, the experience in the diagnostic process.

MEMBER HRAVNAK: Okay.

CO-CHAIR MCDONALD: Okay. Well, let's do follow-up. And just --

MEMBER SEQUIST: Can I just make one comment?

CO-CHAIR MCDONALD: Oh, sorry, go ahead. Yes, Tom, sorry.

MEMBER SEQUIST: I think this is a really great space to get into. Can we use the word patient experience with the diagnostic process?

Satisfaction implies ratings, in the CAHPS world. So it would seem less helpful for patients to rate their diagnostic experience from one to ten.

Like rate your hospital, rate your doctor one to ten. Because that's very subjective, and we don't know what's leading to that reading.

If we use the word experience, then it implies that what we're really looking for from them is, do you think too much money was spent on

your evaluation? It becomes much more objective.

I know that's like nitpicky, but in the CAHPS world, like it means a lot to say satisfactory --

CO-CHAIR MCDONALD: Yes, that's actually quite, I mean I'm interested in the rest of the group, but that's probably a pretty good idea. Because that would broaden this, which is the idea; it's a consolidated concept

Because when it's consolidated, it means we wanted to cover a few of the measure ideas. So that's another, I mean, we've had a couple of ideas right here that are in line with, I think, our desire to broaden it.

So it would be patient-reported experience with the diagnostic process. And that could then have sub-measure concepts. David.

MEMBER NEWMAN-TOKER: One of the things that you have to think about, a little bit, is whether we're going to advocate for patients to be inquired on this issue systematically or in the sampling schema or

whatever.

Because at the moment, for instance, most of these follow-up calls that are done are done for just like a 10 percent sample or whatever. And depending on how you plan to use these, right, that may not be enough for identifying diagnostic errors in subgroups of patients or with particular providers or anything else.

So there's also, in addition to kind of honing the wording, there is also the sort of honing the sampling frame.

And I think what Mark said earlier, about asking the patients whether they got the right diagnosis or not I think is an important source. It should probably go under the tab that we had, the diagnostic error tab. But I think you could kind of combine some of these ideas together.

MEMBER KUZMA: I can't remember if the slide before did this, but I think in addition to satisfaction or experience, we should have

patient understanding of the diagnostic process. With a teach-back or something like that.

CO-CHAIR MCDONALD: Right. We didn't have it here, but yes, there it is.

MEMBER KUZMA: Okay.

MEMBER GRENACHE: That reaches into the health literacy issue that we were just talking about earlier. So I'm half guessing that.

CO-CHAIR MCDONALD: Meaning here again, we should probably make sure we're keeping the link to health literacy, to the extent --

MEMBER GRENACHE: Yes.

CO-CHAIR MCDONALD: -- that it is applicable in the diagnostic space in some way. Yes. But that should be examined.

MEMBER GRENACHE: Yes.

CO-CHAIR MCDONALD: Okay. Can we do the next tab?

MEMBER SHERIDAN: Sure. Okay, this is follow-up. This is the fun topic.

Okay, so we've got the -- again, we

consolidated. There were many concepts that were similar, so we consolidated several of them.

And then we have the percentage of tests that were pending during a transition of care. Yes, we really thought about transitions of care, home or to another location.

Gosh, I'm having a hard time reading it from this angle. Why don't you read it, Kathy. Go ahead.

CO-CHAIR MCDONALD: Sure. Yes, so percentage of tests that were pending during a transition of care, so it had been hospital-only, are documented, and have adequate and appropriate hand-offs.

So pending results includes awaiting final read or final interpretations. So we just wanted to make that clear.

The second one is rate of critical test results that are acted on in a timely manner. And then the rate of noncritical, actionable test results that are acted on in a timely manner.

So we kind of divided up the test space into three main categories. There is critical test results, which are, they're on some list. Joint Commission's, or they're defined as critical.

Then there is noncritical but actionable. In other words, test results that clinicians would be interested in seeing and care about.

And then there might be other tests that nobody is really trying to act upon once they've gotten a quick look at them. So those are the three.

So we ended up with six actually in follow-up.

So this one is that there are sort of processes in place, standard operating processes that are in place, to ensure closed loop communication of actionable test results to the patient. So this would be both critical and the non-critical.

The same thing, standard operating

procedures are in place to ensure monitoring of abnormal findings, broadly construed. Standard operating procedures are in place to ensure that results are communicated to the responsible communicator.

For example, the primary care doctor or other responsible, sort of organizing physician if a patient's care is mostly being coordinated by another physician.

And those were our six that were pulled in from what we had --

MEMBER GRENACHE: Can you go back to the previous slide?

CO-CHAIR MCDONALD: Yes. Okay. Yes, go ahead, David.

MEMBER GRENACHE: So you mentioned, for the critical test list, use something that might be well-adopted as these are critical tests. Is it up to us to define that, or are we going to defer --

CO-CHAIR MCDONALD: Yes, so our understanding is that there are lab tests out

there that are already called critical test results.

MEMBER GRENACHE: There are --

CO-CHAIR MCDONALD: Okay.

MEMBER GRENACHE: -- but they're not standardized.

CO-CHAIR MCDONALD: Okay.

MEMBER GRENACHE: And I think I'm going to agree that potassium, troponin and things, certain things are critical. But then it quickly expands beyond, things that I would say aren't critical but are often considered to be critical.

Which brings me to my next point, the bullet point number 3. Is it, again, up to us to define what an actionable test result is, or --

CO-CHAIR MCDONALD: Yes.

MEMBER GRENACHE: So there's ambiguity there that --

CO-CHAIR MCDONALD: Definitely there's ambiguity there. And so the idea was just that there is, the thought was, what we were hearing

from folks in our group was, that there is such a thing that critical test results that labs already say are critical and that have to be communicated within a very confined time window. And that those should have their own category.

Because they've been blessed by someone somewhere. At the Joint Commission and perhaps others. We didn't know if there was any further extension.

And the other category is tests that clinicians would feel are, it's good to have them be acted upon in some timely manner, both --

MEMBER GRENACHE: Right. Okay, so --

CO-CHAIR MCDONALD: -- which would need to be defined. Yes.

MEMBER GRENACHE: Right. It could be left up to the healthcare system.

CO-CHAIR MCDONALD: Yes.

MEMBER GRENACHE: You know, this is our medical review board's decision on what a critical test or an actionable test --

CO-CHAIR MCDONALD: Yes. Hardeep.

MEMBER SINGH: So there are current recommendations that CLIAC, which is a CDC-based Committee, Federal Advisory Committee, passed on to CMS, which we're hoping that they would respond to in the near future, about standardizing exactly some of the things that you brought up, in terms of what tasks are actionable or not and what to do with them in terms of communication and sort of follow-up procedures. And closed loop reporting.

The results of that will not be out by the time NQF has to produce this report. So we can look at other types of recommendations that have come up, by other governmental agencies, including ONC and the VA.

So within the CDC recommendations, both of those documents, one was ONC SAFER Guides, that David can talk to a little bit more, that we helped develop. Which addressed communication of test results.

And in fact, I think John, there is some measures that we discussed briefly in email,

that get exactly to that. So rather than proposing things that would come up with debate or action, can we just point to some of the SAFER measure concepts that are in the SAFER Guides already?

And David, the new SAFER Guides have some of these measures that we could just say that, to adopt. And then also reflect on some of the VA's work on communication of test results. And using some of the measures that were proposed. As an example. It doesn't have to be --

That's what CLIAC did. They just give examples, and they said, VA and SAFER Guides are examples of the types of things that this workgroup ought to be working on, which should come out of CMS. I don't know what the CMS action on that's going to be though.

CO-CHAIR MCDONALD: Okay, Mike.

MEMBER DUNNE: Yes, once again, as we get into this area, it's like peeling an onion.

So positive blood cultures are a

critical result and need to be communicated directly to the responsible physician. Antimicrobial results, susceptibility testing, or the organism that's recovered from the blood, is not.

To me, it's just as important as the organism. But there have been studies, there are passive results. They go right into the LIS and the HIS. But physicians need to retrieve the results themselves.

But there have been interesting studies that have shown that all of the antimicrobial activity, in terms of therapy, occurs after the blood culture is reported positive. And very little happens after the antimicrobial susceptibility results are entered into the LIS.

So there's a sense there that there probably are critical results that need to be called, but it would be impossible to comply with that.

CO-CHAIR MCDONALD: Interesting.

David. And then David and then Sue. Oh, Sue, did you want to reply to any of that because --

MEMBER SHERIDAN: No.

CO-CHAIR MCDONALD: Okay.

MEMBER SHERIDAN: No, it's an additional comment.

CO-CHAIR MCDONALD: Okay. Okay.

MEMBER SEIDENWURM: So this idea, this whole concept of closing the loop, is extremely important. And I think that one of the main areas, and I think this would be great if we could address this, is a patient comes into the emergency room, for example, has a big work-up, and there's stuff that needs to be dealt with that isn't necessarily germane to their acute event. And then two weeks later, they're discharged, and that information has been kind of lost in the sands of time.

So if there could be a dedicated EHR functionality for following some of these problems that was actually kind of user-friendly and so forth, that would be great.

The other thing that would be great, and we've tried to implement this in our system and have had a great deal of difficulty integrating this into our EHR system, is just the whole mechanism of loop closing and making it a humane process and not something that depends upon people's memories and looking back through the chart when the patient comes back.

And then just to say that radiologists are, we own part of this problem, because we don't give useful information sometimes, so we are putting in place a -- well, we're proposing a performance measure now through our process of specifying the time interval for follow-up and the specific modality of follow-up, to the extent that we can.

And so I think that each other, some of the other specialties perhaps could do some of the same things. And specifying exactly what that would be. And that would somehow be integrated into these EHR functions that would actually make it more likely that these processes

occur.

Because right now, even though the vendors purport to have these functions, they're oftentimes difficult to use.

DR. BERNOT: David, could I just respond to the very first thing? Can you go to the next slide, Vanessa?

We actually were hitting that exact same thing on the abnormal findings in the second bullet point over here. We may need to clarify the language, but that's what we had intended to do.

You go to the emergency department; something strange comes up. How does that get followed up? Some abnormal finding. So we'll clarify the language.

CO-CHAIR MCDONALD: Okay. And then, we'll just go, Lavinia.

MEMBER MIDDLETON: Just a quick comment to you. We could perhaps use a decision support tool that either the clinician who put in the positive -- put in the antibiotics for the

positive blood culture would be notified, that very clinician or the pharmacy, once the susceptibility comes out.

And either taking the clinician out of the ordering loop or notifying the clinician and the pharmacy that the susceptibility of whatever the positive blood culture is. I mean, I think that's the type of closed loop and decision support that we need, in order to improve patient care.

And to assume that the pathologist and the laboratory will be able to identify the appropriate person for both the positive blood cultures and the susceptibilities is less likely to happen. But to have a decision support where it's either automatically updated or automatically contacted and updated is, I think, where we want to go.

CO-CHAIR MCDONALD: Yes.

MEMBER DUNNE: Well yes, I can foresee a time where a pharmacist is available 24 hours a day and is looking at results coming off on a

tickertape like Wall Street. And when something is incompatible with a drug that the patient is receiving, then it would flag that result.

A lot of this stuff has to be automated. You know, it just can't be done manually.

CO-CHAIR MCDONALD: That's cool.

MEMBER NEWMAN-TOKER: So if we go back one slide. Just one comment and then sort of a comment/question.

I would make sure that we're a little bit specific about what we mean by test results. So, like it's easy enough to think about lab tests and then people start talking about, well, does that include radiology. Okay, probably it does, but then what about other tests.

What if you get an EEG, and it shows that you're in status epilepticus. Is that a critical, just make sure that we kind of clarify the spectrum of what is included in that umbrella.

And on the last slide, one more down.

Forward, whatever you want to call it.

Is there a reason why, I mean, because we struggled with this a little bit, too, like, when do you say our measure is whether standard procedures are in place, as opposed to just measuring whether we're doing it?

CO-CHAIR MCDONALD: Yes. Yes, we've actually said we're not sure that it's the standard procedures exactly that should be measured. So the stuff to the right is more the critical piece, and whether it's some measure around standard procedures or whether it's some other measures, that's still to be determined.

MEMBER NEWMAN-TOKER: I think part, for me anyway, the sort of threshold is, if I can't think of any way to actually measure whether they're doing it --

CO-CHAIR MCDONALD: Yes. Then you do this.

MEMBER NEWMAN-TOKER: -- then I resort to a standard operating procedure.

CO-CHAIR MCDONALD: Yes.

MEMBER NEWMAN-TOKER: But I almost don't care about the standard operating procedure if I can find a way to measure the actual thing.

CO-CHAIR MCDONALD: Yes.

MEMBER NEWMAN-TOKER: Whether they're doing the behavior.

CO-CHAIR MCDONALD: Yes. Yes. Do we have to be wrapping up? Okay, we have to be wrapping up. So maybe quick last comments. Hardeep, Susan.

MEMBER SINGH: I was just going to add, and I think it's sort of implied, but it probably needs to be --

One of the single most important breakdown points, why results are not getting -- at least the sub-critical ones, and I think those are the ones that are even more vulnerable to follow-up, not the potassium, is an institution coming up with who is the responsible person for our follow-up.

And working with so many institutions now, just sort of trying to come up, everybody

has their own thinking, and I'm wondering if we should actually call that out as somewhat sort of measurable.

We can put it as a separate SOP, but that assigning who's the reasonable clinician in your institution is one of the hardest things for anybody to do.

CO-CHAIR MCDONALD: Yes. Hard for the patient to know, too. Okay, so we're at the point where we are going to break into our small groups again, and I think we have instructions. Who's giving us the instructions?

MS. SKIPPER: Yes. Yes, so we'll break back into the same small groups we had this morning and review any measurement gaps.

And I believe we have a slide that will sort of guide through some of the discussion questions for this breakout. But we will come back at 4:45 and then go to public and member comment. And then we'll let you all report out notes on gaps first thing tomorrow morning.

So we are going to --

CO-CHAIR MCDONALD: Yes. So go sit in your groups again, see if there's any gaps that you feel that your area might need to cover because it wasn't covered somewhere else. And it seems to be related to your area, or you just want to expand your area because you've heard from people, and write those down.

And we will be reporting back on those tomorrow. We're not going to report back on those today. Since we'll come back and have the public comment and be done for the day.

MEMBER SINGH: Kathy, are these slides, all of the slides going to be available as we discuss, to just recap and say, hey, what are they saying about that measure and that one.

CO-CHAIR MCDONALD: Yes, that would be helpful, wouldn't it.

MEMBER SINGH: That's number one. And number two, do we have any sense of, you know, we started off with more than 200, do you know how many we have come down to?

Because we know, in our group for

instance, and we were down to, what, 17 right? So we're down to 17 in our group.

CO-CHAIR MCDONALD: It looked like everybody had like four to six on average for the 12 tabs. So we must be, roughly, at 60. Which was the goal.

(Off microphone comment.)

CO-CHAIR MCDONALD: Yes.

MEMBER SINGH: If we could count them, it would be really useful. Please. I mean, just seeing how much more we need to weed out.

CO-CHAIR MCDONALD: Yes, I think we've made good progress.

(Whereupon, the above-entitled matter went off the record at 4:11 p.m. and resumed at 4:53 p.m.)

MS. SKIPPER: Okay, everyone, if we could make our way back to the large group, we'll open up for public and member comment and just a couple of announcements.

Operator, could you open the line for any public and member comments?

**OPERATOR: If you would like to make a public or member comment, please press * 1. And there are no public comments at this time.**

**MS. SKIPPER: Thank you. Comments in the room?**

**MR. EPNER: Paul Epner. It's tough to make member comments four hours after the comment was made that you're responding to, but I have just a few.**

**There was a notion this morning on the, or earlier this afternoon, on capturing the initial diagnosis, the working diagnosis, and the final diagnosis. All of which are really important points.**

**I think that I would just ask the committee to keep in mind that, especially in primary care, sometimes the first visit is a diagnosis sort of, and then the patient never comes back, and you don't know if it was the right diagnosis or if it went away or anything. So does everything get coded as an initial? When does it become a final? What's a working?**

It's always a snapshot in time, and how it would get decided, I don't know.

I think the most important thing of the notion is the audit trail. Being able to know when it moves so you can look at why it shifted and who changed it and things.

So my comment there was, think about a structural requirement to have an audit trail as things move from initial to working to final.

The second opinion statement, I think a possible measure concept is around second opinions that match the first opinion as focusing on the positive there. Because getting the second opinion that is discordant doesn't mean the second opinion is right any more than the first opinion was right, unless there's an adjudication process and that's documented. It's still just overall knowing your rate of confirmation from second opinions. But also, it's just sort of a little of the reverse.

Something we've probably beat to death, the notion of overtesting and

overdiagnosis. I think on the issue of overtesting, there is no question; there's a ton of overtesting.

But as was pointed out, there is undertesting. A1c among diabetics, people on drugs that have liver toxicity that aren't getting their liver tested. You know, live enzymes.

So there's lots of examples. And so to put influence on, more emphasis on one more than the other in the absence of evidence, which I haven't seen that one is more important than the other, I think I would recommend against that.

And again, when we talk about overdiagnosis, there may be diagnoses that we want to put in that category, but I think ever misdiagnosis is an overdiagnosis of something because it's clicked up one more in the diagnosis that shouldn't have been there.

So again, being careful of what are we adding in the measures world, that doesn't mean

people in the drug world or people in the policy world don't want to keep using overdiagnosis, but in the measures of safety and quality, what are we adding from this world overdiagnosis that misdiagnosis or overtreatment isn't capturing? So just something to think about.

We talked a lot about pathways and guidelines and measuring compliance with them, and which ones are the ones we're going to measure compliance with. A structural measure might be, have health systems adapted, adapted a certain set of pathways.

However they define what's acceptable, but that they said, this is how we wish to practice medicine. And then measuring compliance with that.

Whether it's some specialty organization that did a systematic review and made a recommendation. The onus should be on the system to say which ones are we going to live by. And then measuring compliance with that.

And then, I don't know if that came up

in gaps or whatever, the comment this morning from me about delays due to payers forcing patients to a denial or a petitioning cycle. So again, whether it's access or whatever, just that notion that picking up those other aspects I think is important.

And I think those are my comments. Thank you.

MS. SKIPPER: Thank you. So just a couple of announcements. We will be taking your discussions from this last breakout group and having those ready for you to review and report back out first thing tomorrow morning.

We'll also be cleaning up the measure concept lists based on your discussions today to have a clean slate to sort of work from.

And then also, I just want to note that we do have a reservation for dinner tonight at Siroc. You can see the instructions for the address on the slide here.

The reservation is under NQF. And we just ask you to please keep your receipt for

reimbursement following the meeting.

MR. LYZENGA: I forgot to mention that last time. We're not allowed to pay for you for dinner. We reimburse you, but we can't pay directly. So pay for it, get your receipt, bring it back to us, and we'll reimburse you.

And apparently, you're only allowed to have one alcoholic drink per the federal government rules. Or we're only allowed to pay for one.

MEMBER MAHAJAN: Is the size of the alcoholic drink --

(Laughter.)

MR. LYZENGA: No. You may have found a loophole.

(Laughter.)

MEMBER RADFORD: So those of us that are at the hotel now have a dead hour and a quarter, essentially, if we're going to go eat. So many can we start a little sooner maybe?

MS. SKIPPER: Do what?

MEMBER RADFORD: Could we start eating

a little sooner, because usually restaurants don't care if you show up early.

MS. SKIPPER: Yes, I'm sure we could reach out and arrange that. So I guess maybe push it to, what, 5:30?

MEMBER RADFORD: Yes, something like that.

MS. SKIPPER: Get out of here and maybe --

MEMBER RADFORD: Yes.

MS. SKIPPER: -- go back to your rooms and set your stuff down and come back or --

MEMBER RADFORD: We can't.

MEMBER NEWMAN-TOKER: The hotel is too far away.

MEMBER RADFORD: The hotel is too far away.

MS. SKIPPER: Okay.

MEMBER RADFORD: That's the thing.

MS. SKIPPER: Oh.

MEMBER NEWMAN-TOKER: It will take us 15 or 20 minutes just to get back to the hotel.

MS. SKIPPER: Okay.

MEMBER NEWMAN-TOKER: So it's just a little bit easier to just move on.

MEMBER RADFORD: Yes.

MS. SKIPPER: Okay.

MEMBER RADFORD: Just go right to the restaurant. Let's eat, and then we get back to the hotel.

MS. SKIPPER: Okay. We will take care of that. And just to note, we are starting tomorrow, 8:30, breakfast; 9:00 a.m., the meeting will begin again.

And we'll let you know, I'm sure the reservation will be find, so we'll see you there.

(Whereupon, the above-entitled matter went off the record at 5:01 p.m.)

**E**

**U**

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Improving Diagnostic Accuracy 2016-2017

Before: NQF

Date: 04-12-17

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Court Reporter