NATIONAL QUALITY FORUM

+ + + + +

IMPROVING DIAGNOSTIC ACCURACY 2016-2017

+ + + + +

THURSDAY
APRIL 13, 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

HELEN HASKELL, MA, Mothers Against Medical Error

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
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, Senior Vice President, Quality Measurement

JOHN BERNOT, MD, Senior Director

TRACY LUSTIG, 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)

CONTENTS

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

9:04 a.m.

CO-CHAIR GRABER: Good morning, everybody and welcome back. I talked about the miracles that NQF is capable of and it's happened again. We've left them with this big mess when we were done yesterday. And they transformed it into a gorgeous document that summarizes all our concepts that we were happy with. At the back are the areas that need more discussion, the areas where there are still gaps or there is still more refinement needed.

So the plan for this morning is to start off with the gaps which are at the back of the handout that you have. We will go through it group by group. Each group will talk about the gaps that they discussed individually in their groups yesterday. The rest of us will have an opportunity to point out any other gaps that we identified then or now and would like to see included.

After that, we will get to the main

part of the document which is the concepts that we all like and agree upon and we will rank those according to their importance, their feasibility and their potential for cost-savings.

That is the plan. Christy will tell us more about the paper and the process.

MS. SKIPPER: Yes. Thank you, Mark, and good morning, everyone. Everyone should have in front of them this concept rating sheet which pairs down all of the measures that you all pointed out as important and to move onto the next level to move onto today for review.

We have around 70-72 measures concepts down from over 230 yesterday. So we organized this sheet so that you can see the measures per subdomain that were categorized.

I just want to note that I didn't finish all of my homework. So Groups 2, 3, and 4 do have their gaps in the presentation slides that we are going to show you this morning. Group 1 as I said their measurement gaps or ideas for measures that need refinement are toward the

end of that document. I'm sorry that I didn't make that more clear, Mark.

As Mark said, we will review the measure prioritization criteria with you this morning after we do the presentation of gaps. And you will use this paper document to begin to do your individual prioritization or your individual ratings of each concept according to the criteria, importance, feasibility and cost savings. We'll go over this again once we get to that activity.

But for now, we'll start out with the group report backs starting with Group 1 on any measurement gaps or as we call it concepts or ideas that we have for measure concepts and they need further refinement.

For Group 1, if you could hit this document here and turn to the second to the last page. You should see it up on your screen in front of you. Measure concept ideas that need further refinement. I'll turn it over to Mark to run through those.

CO-CHAIR GRABER: Is everybody on the next to last page? It says Measure Concept Ideas that need refinement. These are some things that we talked about in our group that didn't make it into the list of measures that we all agreed would be great. But they had something in them that seemed like it needed a little bit more discussion or fleshing out.

For example, the first thing there is about discussion support in the medical record to make diagnosis easier. And we have one in our medical record in the VA. If you're seeing a patient with diabetes, it pulled in all the information that you want to know for that clinic visit. It saved you from having to tab around and find the 20 things that had been aggregated for you.

We were wondering if that kind of thing might be helpful for diagnosis. If so, it would be great to have a measure that encouraged it.

Are we just mentioning these, Christy

or?

MS. SKIPPER: Yes.

CO-CHAIR GRABER: Vote on them? Get a group reaction?

MS. SKIPPER: Group reactions. Reporting out the concepts from Group 1. So we had those listed in one through eight were the concepts or ideas that you all discussed yesterday. Have you ran through them all?

CO-CHAIR GRABER: Maybe just thumbs up/thumbs down. I don't think we have time to discuss each one in detail.

MS. SKIPPER: Just general impressions or.

CO-CHAIR GRABER: Okay. Just maybe one minute on each. Thoughts on that? Ways to make diagnosis easier using decision support.

MEMBER SINGH: There's already something in there.

CO-CHAIR GRABER: The only thing in there now is decision support with differential diagnosis. So it would be like a data aggregator.

MEMBER NEWMAN-TOKER: I think we have to be careful about getting too far down the road of trying to legislate every aspect of the specifications of what's in the EHR. I actually think using an approach where the measurement is a little bit more details agnostic. Asking the physicians and the other providers does the electronic health record, how well does it support your delivering accurate, high quality diagnoses as an aggregated just Likert scale or whatever.

CO-CHAIR GRABER: Yes.

MEMBER NEWMAN-TOKER: We can give them some examples in that question. Does it provide decision support pathways? Does it aggregate data properly as a good data visualizations?

I'm just a little worried about creating 20 measures that say does your EHR do this, does your EHR do that, does your EHR do the next thing. I'm not sure that's really where we want to go.

CO-CHAIR GRABER: Hardeep.

MEMBER SINGH: So I think we need to first think about we can't have good clinical decision support or any decision support without good structured data. Now we don't structure some key data elements including, for instance, family history on patients who have family history of colon cancer and we miss them.

I think maybe we need to just think from the point of diagnosis the key would be the EHR should be able to capture some key critical structured data whatever that might be, chief complaint or family history or whatever other data elements rather than being at that high level.

CO-CHAIR GRABER: Thanks. It sounds like this one is a little premature. The second one was about education. Shouldn't there be --

DR. BURSTIN: Mark, I'm sorry. I actually had a comment. I just wanted to maybe ask David a question.

Some of these feel like they may in fact really be -- I'm sorry. There are so many

Davids. David Hunt. It just seems like some of these really could wind up being certification elements for EHRs. Not everything has to be a performance measure just because we're talking about nails and hammers here. Not every one of these is a nail. I guess the question would be would some of these really be things better suited towards certification requirements for EHRs.

CO-CHAIR GRABER: I definitely think that's true and when you get down to the EHR level, we really need to be more granular in terms of the concepts. Taken very, very broadly, an EHR with clinical decision support that really supports diagnostic accuracy is Watson. And that's just not.

I think the thing that I'm hearing more and more from our colleagues is they really need good cognitive support such that data is aggregated and that you get all of the data.

So its inoperability is absolutely key. That test, that scan, the other information

that was done last week is in front of me now so I can make an appropriate decision. I think the data aggregation and collation functions are probably key.

And strictly speaking it already should be baked in there when we look at what we've asked to have done just isn't done very well. So a long way of saying, I think to some extent some of this fits into what we call the category of usability more so than actual new functionality.

MEMBER RADFORD: I'd just like to iterate a focus that I have in this group which we are developing measures that will be reported nationally. Some of the measures' suggestions predispose the interventions that need to be made. This is one of them.

And I really think it's very premature to put these as measures. These are potential ways to improve diagnosis that some providers will experiment with, trial. And after they become more common, maybe they can become

measures perhaps via certification. I like that idea. But I personally would not put this on the short list today.

CO-CHAIR GRABER: Anything else on that one?

The second one was about education. We don't have a measure that says maybe there should be some education about diagnostic error either in medical school or for people in actual practice. Should there be a measure related to providing some education about diagnostic error?

MEMBER SHERIDAN: Yes, I was thinking about this last night just thinking about the measures that we spoke about yesterday and wearing my patient advocate hat, thinking about the errors that took place in our family. Would these measures reflect that or prevent some of the harm that happened in my family?

And you brought up the education. Something that we learned in my son's case who suffered brain damage from his newborn jaundice that the nurses had not -- They looked back for

13 years in the records of the education of the nurses and they could not find any evidence of the nurses being educated about the dangers of jaundice.

When I testified in AHRQ in 2000, that's one of the things I asked to be researched. How could nursing staff not know in the newborn section about the dangers of jaundice going back 13 years? There was no indication of education.

That's a very specific in a very specific area. But I think that if there's some way to communicate that we need to know that there is adequate education in certain areas where there is a high risk of misdiagnoses.

MEMBER HASKELL: Well, I would agree, although I don't know how specific it necessarily needs to be. But it needs to be a big part of medical education since it's obviously a big part of what young doctors do. They often are not aware of what they don't know.

And the other issue that Sue just

brought up is education for nurses which I think is something that's an aspect of diagnosis that we need to really develop some guidelines for of how do nurses intervene in diagnosis. What do they do? Maybe there are guidelines that I don't know about. But I think that's an important issue.

MEMBER DUNNE: Is this limited to medical school? Nursing school and medical technology school, radiology/technology school, this is all part of health care. I mean if we're going to cover it we've got to cover it as part of the education for every health care related field, pharmacy. Don't limit it to medical school.

CO-CHAIR GRABER: I guess it's a question for the NQF staff. Are we allowed to make measures that address education needs?

DR. BURSTIN: It's a really interesting question. Obviously, there aren't accountability measures per se for that sector which is why again -- back to my hammer and nail

thing -- not everything has to be a performance measure. It could be this is a recommendation of this committee that you would like to recommend. These get incorporated, for example, into ACGME and whatever the equivalents for the other broader health professional schools.

But I don't see it as a performance measure per se. But I don't want to lose Sue and Helen's point. This is really important and it could be important recommendation of this committee. I don't know whether IoM already did that as well as part of theirs.

MEMBER SINGH: This was sort of touched up already by IoM. And we're going beyond a little bit of what we're supposed to do here.

If you want to try to say here's a measurement area that there is a gap and you want to do something about it, then we need to make it more specific and say "I hear that you have a Macy Foundation grant to develop a curriculum." Andrea Olsen has been developing curriculum.

Maybe we could say something specific. Curricula such as X, Y and Z is the recommendation of the Committee, should be incorporated and somewhat measured and trainees should be measured on his performance in the future as a gap area that we need to address and could be addressed in the future. But it has to be a lot more specific than right now.

MEMBER HRAVNAK: I was just also going to say that we could probably put the word interprofessional in here. I think that that would help to broaden it and maybe something about both preparatory and continuing education.

MEMBER NEWMAN-TOKER: I think the other thing is that we could potentially wrap some of this is into the notion of the learning health system rather than couching it in terms of what's taught in medical school and what's taught in nursing school. We should be thinking about it as part of what the institutions are doing when they're learning from diagnostic error and then retraining and educating people around the

errors that are found. Then we could be essentially closing the loop on the educational front as well.

CO-CHAIR GRABER: The next one is about differential diagnosis and we already had a measure that we had agreed upon. It was number five back on the very first page. It says allow the clinician to document the differential diagnosis.

And we had some discomfort I guess about that. Did it effectively address not only structure but process? We felt like we didn't quite nail it.

Is it enough to say that your EMR should allow you to do it? Or should there be a measure that says we'd like to see one on every new patient that's got a new complaint?

MEMBER SINGH: This one doesn't say EMR, the one we have here. It just says allow for the clinician to document. It doesn't say the EMR should, does it?

MEMBER NEWMAN-TOKER: No, it's under

specified, implied, but obviously we would need to state it. But in this situation again for any given idea or theme you could have a structure, a process and an outcome measure. And part of the goal of the non-proliferation of measures is to try to pick the level at which you can get there.

I think if you have a measure that says what percentage of the cases have a differential diagnosis documented effectively you've covered the structure measure of are you able to document in your EMR a differential diagnosis. And it's a better measure of whether they're actually doing it rather than whether they have the potential opportunity to do it. So I would just switch that from a structure measure to a process measure and move on.

MR. LYZENGA: And actually we have a structure measure that's in the list already to allow for the clinician to document.

MEMBER NEWMAN-TOKER: Number five. I'm saying take number five and just make it percentage of clinicians who are able to document

their differential diagnosis or do document their differential diagnosis in the electronic medical record. And you'd kill two birds with one stone.

MEMBER SINGH: If you're going to go down that route, we have to make it a gain more specific. You don't need to put a differential diagnosis for every single patient who comes in for a routine medical visit. You would do that if they're coming with a new symptom that you're uncertain about. So I think we're going to need to put some parameters if you're going to make some recommendations that aren't something that broad.

MEMBER MAHAJAN: Just as a part as a clinician who would see a patient as part of that evaluation it would be a part of the differential diagnosis. And I'm just looking at the EHR world and the feasibility of initiating a new step in the EMR which is mandated versus requiring the clinicians to document differential diagnosis as a part of their known. I don't know how feasible that is but I'm just saying from that

perspective. So because One is changing the EHR, which may be a harder battle, versus forcing the clinicians because we need the portion of the clinician, not the fact that they can document.

CO-CHAIR GRABER: Kerm.

MR. HENRIKSEN: Just an observation. Many medical offices are starting to use medical scribes for documentation purposes. And this is a growing industry it looks like. And it frees up the clinician in terms of the problems they have with the EHR to begin with.

But at the same time, there's opportunity to introduce new forms of error as a result of someone not as well trained perhaps as a medical scribe into the system. So it's not just the clinician that may be doing the documentation. It's the medical scribe. And that's still something that needs to be evaluated I guess in terms of the long-term outlook for that. As long as EHRs are not very useable, this is likely to be a growing industry.

CO-CHAIR GRABER: The next one is

about whether there should be a measure that addresses quality monitoring and laboratory and radiology services. So many health care institutions, they have a policy that a certain fraction of the x-rays are re-read as a peer review process and a quality monitoring process or certain fraction of the biopsy gets a second reading. And we don't yet have a measure that says that that's a good idea. Should we?

MEMBER DUNNE: There are accreditation agencies for both of these specialties. And it's pretty rigorous. So if the CAP certifies the laboratory as being accredited on a routine basis, then it's not something that we have to redo.

CO-CHAIR GRABER: Does the CAP have criteria that says a certain fraction should be re-read?

MEMBER DUNNE: You can develop your own quality improvement markers and so forth so that they allow for a variety of different. For example, in microbiology, what's the percentage

of contamination of blood cultures? How many identifications don't correspond with the correct antibiogram and so on.

MEMBER NEWMAN-TOKER: We filled that gap in the revision after the discussion. So it's on the list.

MEMBER RADFORD: I would just like to point out that there's a lot of cardiology testing that needs to have the same type of concern, EKGs, ECHOs, nuclear, blah, blah, blah.

CO-CHAIR GRABER: Looking at number six, I'm not recalling our discussion on number six. Anybody from Group 1 remember the issue about OPPE and FPPE?

MEMBER SINGH: We addressed it in ours, but it wasn't anything like this. I'm not sure what this means. I mean you've got to talk to patients and providers. I think that's so broad.

CO-CHAIR GRABER: And I'm having a similar problem with seven. I don't remember the issue there.

MS. SKIPPER: And I think measure concept seven that was one that Thomas had come up with and he is not here to perhaps provide more clarity today.

CO-CHAIR GRABER: Okay. And number eight --

MEMBER KUZMA: I think it was just the idea that when there's data that doesn't support the diagnosis are you documenting that there's a conflict that exists for this work-up or something like that.

CO-CHAIR GRABER: That's right. Yeah, wouldn't it be nice to have a Watson like process running in the background to say your diagnosis is not consistent with this lab finding over here or something like that.

MEMBER NEWMAN-TOKER: At the very least we have some in ours that talk about in reference to Gold Standard.

CO-CHAIR GRABER: Okay. And the last one was about I think in each of our groups we had some very specific, disease specific measures

that had already been developed and approved elsewhere. And we were looking at those and we just had the comment that it would be nice if every professional subspecialty came up with measures that are very specific for their diseases like the ones that we all had examples of.

The thought that the work on developing measures shouldn't stop with this group, we should ask some subspecialist and their societies to go through a similar process to come up with measures that would target their specific areas of expertise.

David and then Prashant and then David, David.

MEMBER SEIDENWURM: With respect to number eight, I think that's really an important one, particularly in areas where there are clear criteria for the work-up to stop. And I think in radiology we see endless follow-up. I think if there are clear criteria that define benign positional vertigo rather than an infarc, if

there are clear criteria that define a benign adrenal lesion rather than a rare adrenal cortical carcinoma, these should be defined and everyone should be given permission to declare victory and go home.

MEMBER NEWMAN-TOKER: And this would be a recommendation. You're not proposing, Mark, that eight would be a measure or even a measure comment. You're just saying, right, it would be one of the recommendations made in the report.

CO-CHAIR GRABER: Yes.

MEMBER SINGH: I think the recommendations could be very strong in this area because there are areas that need potential measures and some of which already exist. So United Kingdom, they have very strong measures around timeliness of cancer diagnosis. They don't exist in this country. They have two week, four week. They're very strict about it. They have lots of papers about it.

I think it would be good to outline some high risk areas and get the specialty

societies to do some work around it.

CO-CHAIR MCDONALD: I think actually there is a measure concept here though in the generic thematic level. So I think the concept is that there's sequences of testing that are understood by each profession to be appropriate under particular conditions. And there should be measure that fit that concept. We're not saying what they are. We're saying that it's appropriate for the specialty societies to look through what they have and then create measures around that.

MEMBER NEWMAN-TOKER: I think we have that measure concept in ours in the appropriateness section.

CO-CHAIR MCDONALD: Okay.

MEMBER NEWMAN-TOKER: So I think we have the --

CO-CHAIR MCDONALD: The sequence of testing and the idea that you didn't know the sequence of testing to say that we're benign.

MEMBER NEWMAN-TOKER: You may not like

the individual measures that we've suggested that might be associated with that measure concept. But we have the concept.

CO-CHAIR MCDONALD: Okay. Good.

CO-CHAIR GRABER: Does anybody have any other gaps or things that they think are missing from the Group 1 areas which were information gathering and documentation and information integration?

And, Christy, if we think of things as we go along or after we leave, will there be an opportunity to say I just thought of this one that we missed?

MS. SKIPPER: Yes, there will definitely be time for that.

CO-CHAIR GRABER: Okay. So we don't have to necessarily identify every gap right now. But it would be nice. Okay. Group 2.

MEMBER NEWMAN-TOKER: We took to heart all the suggestions that were made by the large group which I think were great. And we buffed up, changed, filled in gaps. The main thing we

changed just so you can find it is under diagnostic efficiency.

So we really completely refrained the issue we heard, Sue's concern and others concern, that we missed the boat on the under/over issue. We were over weighting or over concerned about over diagnosis and over testing and under concerned about under diagnosis and under testing.

What we did was -- I think this was a really, really strong point, Sue, that got us to a better place in terms of the measurement concept which really is around appropriate diagnosis and appropriate testing, essentially, the measure being linked to being an outlier whether you're a high outlier or a low outlier around your performance in this space.

And that gets us really where we want to be which is if everybody's under diagnosing, we want them brought up. And if everybody is at the other end we want to bring them back down. But we can get that in one measure that's linked

basically to your percentile ranking. I think that was a great suggestion.

We also took Mark's concerns and brought up a couple of things. I don't know if it made it onto this page. We put some text in there about trigger methods. And really trigger isn't a measure. It's a way of getting to a measure. It's a method for getting to a measure. And we put that in there predominantly as an appropriate strategy for getting to the measurement.

And then we added things/gaps that were either brought up in our group or by others including this diagnostic reliability thing which could easily go in the diagnostic QI subdomain. But it also could go in the diagnostic air subdomain if you wanted to, this whole idea of second reads and whether it's radiology, pathology, clinical charts, EKGs, or anything else. The idea that you might sample a certain percentage of these and have a systematic way of making sure that you're at least getting

consistency, if not, accuracy. We buffed up the ones on accuracy that are related to diseases where there's a Gold Standard for comparison there.

Anything else that we filled in gaps on, guys, that I'm missing? I think that's all we have to report that's different from the first time you saw it. I'll just make sure.

Oh, patient reporting, sorry. There was one other.

CO-CHAIR MCDONALD: Tell us where you are.

MEMBER NEWMAN-TOKER: Sorry. I'm not sure where I am. Here it is. I'm under diagnostic errors number 13. This is one that there may be worth some discussion. The first thing we did was put a structure measure in there. Mark, you were concerned that we didn't have anything in there about patient's reporting diagnostic errors which I think was an important oversight.

We have a structure measure that says

there's a way for people to actually report the stuff. Clearly, I don't know if it will make the top ten list, but it's clearly something that would be relatively easy to do and sensible in terms of making sure there's a portal for that to happen.

Then what we also added was the idea that you might actually want to try to get a rate of diagnostic error out of the contact with the patient. Essentially, the frequency of diagnostic error is hard to measure across a broad range of conditions. But probably the single, best source of information on whether a diagnostic error has occurred is probably the patient.

Although it's fraught with lots of problems -- people don't have the health literacy and they don't know whether their prior problem was related to their current problem and there's confusion about causality and relationships and they don't know enough about the details to be sure -- we talked about how it might still

nonetheless be valid for individual institution. If they were systematically testing or robocalling their patients at 30 days and saying, "You were seen on such and such a date with this symptom and we gave you this diagnosis. Is that still to the best of your knowledge the right diagnosis, just get a yes or no. Then you could follow up the nos with a human call and get a little bit more meat on the bone as a potential way of tracking your total diagnostic error rate. That would be otherwise very difficult to aggregate.

We put all that in there because you expressed that concern. I think it was a very legitimate one.

CO-CHAIR GRABER: Any gaps in David's areas? So, David, your areas were efficiency and diagnostic error.

MEMBER NEWMAN-TOKER: And one more. Information interpretation.

CO-CHAIR GRABER: Maybe we could just take a minute to look over it. That's a lot of

material. Can we give people a minute to just look through that and see if there's any gaps?

(Pause.)

CO-CHAIR MCDONALD: Go ahead, Martha.

MEMBER RADFORD: Is that working?

CO-CHAIR MCDONALD: Yes.

MEMBER RADFORD: Sorry. Okay. So I'm seeing a lot of overlap with the last three and maybe the last four on our discussion before about the fact that we need to have differential diagnosis. And very likely there are health systems that will use the EHR and decision support to do that. Are we focusing a little bit too much on the solution rather than quantifying the issue?

You know these all may be very fine. But I really see there down the road they're not the first set.

MEMBER SINGH: Are we going to have a later discussion on ranking these?

CO-CHAIR MCDONALD: Yes.

MEMBER SINGH: I just want to make

sure.

CO-CHAIR MCDONALD: There's going to be a whole priority thing along different criteria. So that's where some of this will sort out.

MEMBER SINGH: This is just a gap session, right?

CO-CHAIR MCDONALD: This is the gap session, yes. Are there extra gaps? Comments though about thinking about any of these are helpful in terms of priming whether things are of priority or that we're missing something that we haven't gotten to.

CO-CHAIR GRABER: Are we ready for Group 3?

CO-CHAIR MCDONALD: But there were no more comments on this group's three tabs. No more gaps.

CO-CHAIR GRABER: So for Group 3, could you tell us your areas and any gaps that you would like to discuss?

MR. LYZENGA: Let me see. They may be

down at the bottom. I can't remember if we -- It may actually be best to pull up those slides and talk through them that way. I'm not sure if any of those actually warrant being pulled into the ratings sheet at this point. Many of them were just raising issues. Let's take a look.

MEMBER SINGH: We have built all the language around the comments that we received.

MR. LYZENGA: Yeah, we updated some of the concepts. Keep going. It's just like a couple of notes on external environment. It's difficult to measure but important. We're not sure.

Similar to what we were talking about with education. Who do you hold accountable for having policies that facilitate good diagnoses, payment policies and things like that? These are difficult to apply to any given entity or body.

MEMBER SINGH: I think the point was a lot of these are HCO health care organization centric measures. And some of them were so beyond what the organization could do. I mean

who would be accountable? Would some of these like CMS and other policy making bodies be accountable for some of these measures? ACGME? Whoever? So that was a big problem especially for leadership and cultural issues.

MR. LYZENGA: Skip forward a little bit. Somebody mentioned why didn't we see the length of time seeing the patient come up and we just noted that that was addressed in the work force section. Skip forward. One more. One more after that.

Again, we had some. I don't know if you want to talk to it, Hardeep or anybody else. We had a measure related to incorporating measures of diagnostic performance and to credentialing and re-credentialing processes. These were just some questions about how do you go about doing that basically and how do you link them to internal dashboards and link them to learning. Anything else to add on that?

MEMBER SINGH: Lavinia, do you want to add to that?

MEMBER MIDDLETON: I think that there is a push to put a lot of the assessment measures from the specialty societies in both the ongoing professional practice evaluation and the focus practice evaluation. I think there's an opportunity to use some of the six competency areas recommended by The Joint Commission which include diagnostic knowledge and either refer or use that vehicle that's already in place and recommended by The Joint Commission to look at individual providers' ability to make appropriate diagnoses and to document when that ability falls short. It's also to link that with credentialing and privileging and making sure that we're privileging people based on not assumption of knowledge. But it's also documentation that they are able to diagnose correctly. I think it's another opportunity to pull in information that's also being collected. That was a recommendation.

CO-CHAIR GRABER: I'd like to make one comment on that. As much as I would love to have that happen, we don't have really any way now to

know who's good at diagnosis and who isn't or even to assess how good they are of clinical reasoning. I would love to see some tools developed over the next decade to be able to say that somebody is competent at this.

MEMBER MIDDLETON: What we use is our rate, rule and roll metrics. And basically we're comparing each provider to their peers based on predetermined and a lot of times society-specific measures.

Actually, I did a little homework last night and we were talking about it. I do have a list of many of the specialty societies. There are URLs websites that we could perhaps share as other reference points to look at where this information is being collected or recommended to be collected.

MEMBER SINGH: One of the examples we measured was the colonoscopy withdraw time that gastroenterology has as a quality measure which already exists. Just an example of how the specialty society stuff could be operated.

MEMBER RADFORD: Right. And the point here is that this group is going to be recommending some measures that are going to be appropriate for national presentation somewhere, national display reporting eventually. Some of those will be in this sphere of control of the physician and could be drilled down upon by physician and then reported in their credentialing process as would the subspecialty measures as well.

So it's a recommendation that those things start to appear in credentialing and that be used in OPPE and FPPE and also that it's a way to use the measures at a local and a national level. But is there is a measure here? Well, maybe it's yes. You use these measures as much as possible in your credentialing process which is a structure measure. But another way to work this is to say "Joint Commission, you should have this in your standards."

MEMBER MAHAJAN: Also this is giving the concepts. I think we could recommend that at

the minimum there should be one measure or one aspect in the OPPE and FPPE that is linked to diagnostic accuracy or diagnostic performance. I mean we could come out with that statement as a strong statement.

CO-CHAIR GRABER: Sounds like a great way to word it. Anything else for Group 3?

MEMBER HRAVNAK: Be sure to include the interprofessionalism here, particularly for credentialing and privileging. Remember that it's not physicians. I think The Joint Commission language has other licensed, independent providers. So they changed the wording a little bit to reflect that.

CO-CHAIR GRABER: Anything else for Group 3?

MEMBER SINGH: I'm going to just bring that up. So that's come up before. So is it going to be okay when we do that if some of the nursing community feel that they don't make the diagnosis. So why are they going to be made accountable for something?

(Off microphone comment.)

MEMBER SINGH: Yeah, but are there any diagnosis that affects things or any of the team-based things that we could include?

MEMBER MIDDLETON: Absolutely. And following up and communicating results -- I'm sorry to speak, but absolutely. And these are metrics that can be evaluated. And they're part of the diagnostic team. We talked about that the first time we were together in the whole diagnostic continuum. And they certainly play an important role that can be measured and reported.

MEMBER SINGH: So I think we should make a determination or at least some distinction that we're not saying they're going to make the diagnosis in maybe some of the nursing community, but things like follow-up. That came out in our discussion.

That is so important that if the nursing community can assist with follow-up of test results. It's such a huge step up.

MEMBER HRAVNAK: So there are probably

two roles here which is their participation in maybe the final medical diagnosis. But there are also independent nursing diagnosis that are made. For example, we would stay within the scope of nurses to diagnose pain and then to apply.

That's where we got with the HCAP thing yesterday. That's why I was asking that question. There is usually a standing order for pain medication, but it's up to the nurse to really assess the patient for pain, diagnose that they're having pain and then apply the intervention. If they don't make that diagnosis correctly, then patients aren't satisfied with their pain.

MEMBER SINGH: I think these are great examples and it will avoid some of the pushback that I've received about we're not the person making the final clinical diagnosis. So I really think we should give specific, concrete examples just the way you guys have listed. Anytime you mention specific things like including them for certification related things.

MEMBER NEWMAN-TOKER: We've had good success at Hopkins engaging nurses. They started out feeling like they had nothing to do with medical diagnosis And now I've got roomfuls of hundreds of nurses raising their hands saying that they think that diagnosis is part of their shtick.

We've just written a manuscript and submitted it, but the three areas are one of them is a measure that Hardeep mentioned, this idea of improving and closing the loop in the follow-up of test results.

The other two areas are around medical diagnosis in addition to nursing diagnosis that Marilyn mentioned, one is around the notion of what we call diagnostic triage. Triage is a very comfortable idea for nurses. They involved in the diagnostic process all the time. The medical diagnosis process, all the time.

They decide post-op patient whether they're just going to give them meds off the list or they're going to call the doc because this

patient has a swollen leg or whatever it is. They're involved in the initial process all the time and the dichotomy between saying you're never to make a medical diagnosis and then trying to tell them that they're also part of the medical diagnostic team is causing trouble. So we have to just get rid of the old logic and move to the new.

And the other place they can really help with in patient education around diagnosis. We talked about whether the patient understands their diagnosis. That's another critical domain.

CO-CHAIR GRABER: Helen.

MEMBER HASKELL: Along these lines, I just want to reiterate some of the things I had said on the phone earlier about the idea of working to the top of your license. I think it does go against that. And I see what you have here, the idea of replacing it with the idea of teamwork principles which I think is a much better concept.

I think that promotes hierarchy and it

goes against the idea of teamwork so that you have the nurse, for example, saying I can't do diagnosis. It puts people into slots.

CO-CHAIR GRABER: Great segue to our last group which is Group 4.

MR. LYZENGA: Just maybe Hardeep or Martha or somebody could talk. Some of these were just little notes. But a few, maybe those last four or five bullets there were possible measure concepts.

MEMBER RADFORD: These were more specific measure concepts that organizations might want to consider. That's basically it to add to the OPPE business, yeah. They weren't really necessarily measures for this group, although I suppose they could be.

MEMBER NEWMAN-TOKER: On this list, the last one, unexpected transfer to the ICU is covered in our section very specifically. And I do think it's worth talking about failure to rescue since it's an existing measure that's reported by a lot of people. It incorporates

more than just diagnosis, but it's the closest thing we have to a diagnosis-oriented measure that's actually widely used in clinical practice. What do people think about that?

CO-CHAIR GRABER: But it looks like you had a measure on that, David, in this revised list somewhere.

MEMBER HASKELL: There is PSI 4 is what you're talking about, right?

CO-CHAIR MCDONALD: There's two versions. There's PSI 4 and there's Silver's version.

DR. BURSTIN: PSI 4 was just withdraw from AHRQ.

CO-CHAIR MCDONALD: Because of --

MEMBER HASKELL: Why was it withdrawn?

CO-CHAIR MCDONALD: Because of the amount of work that it would have taken to put it back through. I think it was a resource decision.

MEMBER HASKELL: And there is a new failure to rescue measure that was approved by

the Patient Safety Committee.

DR. BURSTIN: There is the existing measure that's Jeff Silver's. And there was also I think a version of Jeff Silver's measure for pediatrics just brought through the Pediatrics Committee.

MEMBER HRAVNAK: I would just reiterate. I think we strongly need to support the existing failure to rescue measures and metrics. And it is viewed as a nurse-sensitive measure as well.

MEMBER RADFORD: There are issues with the current measure. It is not specific enough in my view. In fact, for national reported measures, we really need to err on the side of specificity relative to sensitivity. I'd better write a paper about that because it's all there is.

MEMBER RADFORD: It could be improved, but it's all there is.

MEMBER HRAVNAK: I know.

MEMBER SINGH: And I'm not sure how

much of that is related to a diagnostic related issue. We could say some of these failure to rescue measures should be developed to evaluate for diagnostic-related issues rather than what they might be looking at right now.

MR. LYZENGA: Do we have any guidance on how you can focus these diagnostic issues?

CO-CHAIR MCDONALD: One possibility would be to just say failure to rescue reconfigured with more of a focus on diagnostic sensitivity. You have to think about the timeliness of where would there be more preventability. Because if diagnoses were more delayed rather versus less delayed and it's possible to have it less delayed, then you would want to be looking for failure to rescue for those situations. And you'd want to give them primacy over situations where it's much, much harder to expect that there could have been a diagnosis in time to then rescue.

MEMBER SINGH: I think the point of this is this is a gap area. There are a lot of

unknowns about it. The same thing when we were talking about readmissions yesterday, so many of these are nothing related to diagnosis. I think this is just a gap area. That's how we should couch it.

DR. BURSTIN: Just one comment. It may be that rather than keeping to use. The term failure to rescue at least the AHRQ measure for example is surgical. It's very specific. And maybe it's really an opportunity to think about a broader framing of this.

Is this really the idea of getting towards a trigger measure that would make you look areas where there may be potential for diagnostic issues? Then you could list that what some of those could be. And those could be good fodder for developers.

MEMBER SINGH: Actually we just did a study in the pediatric ICU transfers where RRT's failure to rescue when the patient went to the ICU and then deteriorated subsequently in 24 hours was better than just looking at autopsies

or patients who had a visit previously before they came to the ICU. So it has actually some fertile ground, but it needs to be looked at much more.

MEMBER NEWMAN-TOKER: Just in that vein then, Helen, that's a great suggestion. If so, we could make the argument in the diagnostic error section where we talk about death, adverse event or loss to follow-up as a means for essentially investigating or triggering a search for diagnostic error is in there in that sense. And maybe we could allude to the fact that it's a refinement of the failure to rescue so people don't think we just ignored it.

CO-CHAIR MCDONALD: There are new opportunities with ICD 10, too, that didn't exist with ICD 9. That's why there was the surgical restriction in the last round of coding availability.

CO-CHAIR GRABER: Anything else with Group 3?

(No response.)

CO-CHAIR GRABER: Okay. Group 4.

CO-CHAIR MCDONALD: I have that. So go to the follow-up session. In that we have on concept two we've just elaborated a little bit more. We went back and talked about this idea of the two different types of tests we talked about before, critical and non-critical, and didn't like where we can came down. So we did some massaging there.

The idea would be that the definition for critical results are those that fall significantly outside the normal range and may suggest a life-threatening situation. That's what The Joint Commission calls critical test results.

The idea for critical test results is that definition that's been legitimized by a national body. I think there was some discussion of VA having critical tests. Perhaps other national bodies have something they call critical tests. This concept relates to those situations where there's a national body that has a

definition for critical test and there's sort of a standardized process for what the follow-up time frame should be.

Then we took our rate of actionable tests. So instead of saying rate of non-critical actionable test, now it's just rate of actionable tests as the other area. This is one where probably specialty societies and other groups would need to work pretty hard on developing the right lists of tests and findings that would be appropriate to be measuring and assuring that the rates of getting those test results to be acted upon are what would be expected. So, that was sort of a refinement.

Then down at the bottom, the gap was to have a process in place to identify the responsible clinician -- that gray area, seven -- a process in place to identify the responsible physician for tests. And we had a lot of discussion about this and how difficult it really is to know who the responsible physician is. Yet, ultimately without having a responsible

clinician responsible for whatever the next step in the diagnosis is, responding to the tests, etc., that's a big crack where there could be many failures.

It's not a fully developed idea. We kept it at a pretty high level. We had a lot of discussion on that. I just wanted to know if there were any comments on that one as a gap. Paul.

MR. EPNER: And just to expand on your definition, we looked especially at pre-op testing where a primary care physician may order a battery of tests for the surgeon, not really for themselves. But the surgeon isn't the ordering physician. That's an example of where knowing who needs the results may not always be the ordering physician as opportunities for things to fall through cracks.

CO-CHAIR MCDONALD: Marilyn.

MEMBER HRAVNAK: I was going to say the same thing. I think it needs to not be physician-limited. And maybe we need a multi-

layer system as well. If you're going to send it just to the physician who is in the operating room, it will be hours before anything will happen. Frequently this is placed in the hands of nurses to track people down, but it's not fail safe. So I think multilayer.

CO-CHAIR MCDONALD: Thank you. That's great. I know a lot of us are thinking about in situations that we've seen go awry for family members and such which measure concepts cover. And I've certainly seen this a lot where all sorts of other clinical professionals are involved in the care and don't necessarily have the results when they need them or know that they have access to them and so forth. Good.

And then moving down to --

MEMBER HASKELL: Wait one second. And one thing I would add to that, Kathy. The patient knows which clinician is in charge of the tests. That's sort of a failsafe measure.

CO-CHAIR MCDONALD: Right. Yes. Thank you for adding that. We should include

that. That's even more aspirational, right, if the clinical team doesn't have an ability or a strategy to know where the assignment is except for norms and assumptions, then it's going to be even harder for the patient to know. Good point.

Sue, go ahead.

MEMBER SHERIDAN: Again, a midnight reflection about our conversation yesterday and just being a patient advocate you learn to just use what's in your heart. You have to say it. And sometimes it doesn't fit in.

But I really thought about our conversations yesterday and my role as an advocate -- not CMS, I'm taking my CMS hat off -- and thinking about the communication of critical test results. That's really been kicked around for almost 20 years. Who's defining what and is it critical or is it not? Is it actionable? Is it subcritical?

Quite frankly, it's unacceptable. People are still dying today from malignant pathologies that do not get communicated. They

fall through the cracks. And it's chilling to know that our health care system has not stepped up and done something really major and robust to stop this from happening.

Quite frankly, if somebody in my family gets cancer or if I get cancer, I'm going to go to a hospital and say, "What is the percentage of malignant pathologies that get communicated to your patients?" That's the measure I want to see. I want to go to a hospital and say "Do 100 percent of your patients know when they have malignant cancer?"

I would like to challenge us to be a little more specific about let's step in and propose some concepts and measures that's really going to stop some of these diagnostic errors that are happening. It's very clear in the diagnostic process that we can step in and do something that's really meaningful. Quite frankly, nothing has happened in 20 years since this happened to my husband.

DR. BURSTIN: I was going to say

something almost identical to Sue actually. I think in many ways we're actually way beyond the idea you would figure out who the responsible doc is anymore. Frankly, every pre-op test I order is my responsibility. I don't care. It's my name on it. I need to check it.

But the bigger issue I think is is it time for a measure that simply says all test results are available to a patient within seven days. Hard stop. It's what the VA has done. It is part of their process now. You don't have to figure out what's critical and not critical.

The patients get every lab result, every result available to them within seven days. And if you talk to people at the VA -- several of you are here like Hardeep and others -- there was a lot of consternation early on on the part of clinicians like oh no. What if I haven't seen it first? Do you know what happened? Clinicians made sure they saw it first and it worked.

I would very much like to be in Sue's camp. In some ways, I think it will push us

where health care needs to even if it's not where it is now. And I don't want to get into who is responsible. It just creates a quagmire and it will never get the job done. My personal feelings.

MEMBER SINGH: Can I answer this? Being someone who actually wrote the VA national policy on communicating test results, I really think Sue's on the spot. We've had enough of this. We also wrote ONC SAFER Guides in addition that have been out about three plus years which very few health care organizations around the country are using even though almost everything you have here is in the SAFER Guides.

So rather than reinventing the wheel over and over again, why don't we just have some kind of measure where organizations are doing test results reporting and communication of SAFER Guides. We actually did the SAFER Guides.

We updated them. The first SAFER Guide did not have a recommendation on patient notification of test results because of some very

convoluted reasons. But there is one now, 3.3, if I remember correctly.

So I think it's time for us to endorse something like the VA policy on communicating test results which actually is seven days for actionable and 14 days if it's not just to give clinicians adequate time. There was a lot of pushback as Helen mentioned, but it's the right thing to do.

There are two big resources. I have no idea when CMS will actually take it up to come up with a task force which will have some national recommendations on communicating test results. But we have two very good existing resources we need to be building upon. So I think a measure around just those two things will be good enough.

MEMBER RADFORD: I'd just like to share. I would like to agree very strongly with the last three speakers and just tell you that we have a standard at NYU that all test results are reported to patients within 72 hours. And we

measure that.

We're living through the doctor pushback about I have to know first and all that. We're living through that and you're right. They just go in and check. I think it's a very good thing to do.

MEMBER CAMPISANO: Just speaking from a five year cancer patient perspective, I know a number of patients whose doctors just give them access so that 24 hours after their test they can go in and check their results whether they've talked to their doctor or not. Especially in the metastatic setting where women and men, but in metastatic breast cancer, it's primarily women. But they're living with this and going in for scans every three months.

They know the process at this point. They know what to look for on their scan reports. So they are able to go in the next day and read them themselves whether they've had the chance to talk to their doctor or not.

MEMBER SEIDENWURM: There are some

state rules with regard to disclosure of cancer diagnoses. California has some rules about that. I don't know if they're statutes or administration. They're not to be disclosed in this automated fashion. The idea being that they should be presented in some face-to-face manner. So there are embargos on some test results in that regard.

MEMBER SINGH: Yes, and I think most of the stuff that we've written via policies as well as ONC SAFER Guides don't say that sensitive test results such as HIV and cancer should be immediately communicated either. I mean there needs to be some kind of institutional hold on some of the results that need face-to-face discussion. All that can be built in.

So rather than again coming up with these nuances, this is a very complex areas. It seems very simple. In fact, we used to be called the Low-Hanging Fruit of Diagnostic Error. It is not low-hanging. It is really a hard and complex area.

CO-CHAIR MCDONALD: Lavinia.

MEMBER MIDDLETON: Just to share at M.D. Anderson. We first started with a seven day rule to release results. We've decreased that now to immediate release of results. And the patients can access them real time in their electronic medical record. I think it helps everybody.

CO-CHAIR MCDONALD: Paul.

MR. EPNER: And I think it's just important that we explicit about whether we mean make available or confirm receipt because not all patients have easy access to IT tools and things. It doesn't reduce our obligation.

CO-CHAIR MCDONALD: And that's number four, right? We have rate of close of communication actionable results to the patient.

MR. EPNER: But in this discussion.

CO-CHAIR MCDONALD: That was right there. Yeah, number four. An important point. Okay. Others? Did I miss anybody else who wanted to comment on that? Helen.

MEMBER HASKELL: I would say it would still be helpful for the patient to know who the relevant clinician is.

CO-CHAIR MCDONALD: Yes, that should be in addition. So process in place to identify the responsible clinician for tests and have the patient know who the responsible clinician is.

MEMBER NEWMAN-TOKER: If you're got that information in the EHR, then it ought to be available in the portal. Yes.

CO-CHAIR MCDONALD: Yes, people are taking notes.

Then our next category was patient engagement. And on that one, we just have added based on the conversation yesterday to number five. We had said that we wanted diagnostic information and the diagnosis communicated in understandable manner to the patient. We added while recognizing the impact of health literacy. We added that based on the discussion yesterday.

For that one, we hadn't identified gaps. I'll go on to patient experience and then

we'll open it up to see if we've missed other gaps. In patient experience, the gap that we noted and added, it's number four in the gray. That's patient experience with a diagnostic process.

We have patient satisfaction with a diagnostic process. And now we've added patient experience with a diagnostic process which included: Was it worth the effort? Did the patient feel like they didn't incur a lot of extra expenses that weren't appropriate after the whole diagnostic process was experienced? I think that was your point, David. But there would be other subdomains of this possible as well.

MEMBER SINGH: Kathy, I thought we were going away from the word satisfaction.

CO-CHAIR MCDONALD: Some people were not liking it and some people were liking it. So we left it on there for right now to go through the prioritization process.

MEMBER HASKELL: I would really say something more along the lines of assessment,

the patient's perspective on the diagnostic process that they went through.

CO-CHAIR MCDONALD: So as a separate theme or under one of these?

MEMBER HASKELL: When you're talking about the word satisfaction, I agree that that's not a good word because it has so many implications. For a bland word, it's pretty freighted. So I would say something like perspective, which gives a little more weight as well to the patient's opinion.

CO-CHAIR MCDONALD: So maybe it would be that this could be in the structural process about was the patient's experience of the diagnostic process invited and checked on and known.

MEMBER HASKELL: I interpret this -- and maybe I'm misunderstanding because it's just a few words -- to mean did the patient feel it went well. They were diagnosed in a timely and efficient way. Or was it a mumbling process that they weren't happy about?

CO-CHAIR MCDONALD: Yes.

MR. LYZENGA: Just to remind us that we have a few like the measure concept above this of similar patient experience with the diagnostic process or different aspects of it.

MEMBER HASKELL: Yes, I think they can fruitfully be combined.

MR. LYZENGA: I don't know if we want to bundle this together.

MEMBER HASKELL: Yes.

MR. LYZENGA: A sort of composite of patient reported experience survey that includes these elements and others or something like that.

CO-CHAIR MCDONALD: Yes, this one seems like it really needs more work. I don't know what the process should be by which it gets the extra work it needs.

MEMBER SINGH: I'm just going to add. I think this is a very under researched area that I'm not sure that we could claim. I mean all of these are nice things. But I'm not sure they are getting to the point of trying to create even a

measurement concept.

What does it mean? If I have cold, I go to the doctor. My head hurts or whatever. I get a steroid shot. I come back and I feel great. My diagnostic experience was so good. I feel beautiful the next day.

Was that good? Is that what we're trying to catch? I don't know. I mean I just think we could say for high risk conditions maybe, for certain aspects of breast cancer, colorectal cancer, something to make it more specific. I just don't know if we know enough about this area.

The other thing we should think about is -- and maybe this is in the next session -- are some of these things just good concepts that need to be put in further for development and measurement. But they are good ideas. But we just can't call them measurement concepts yet. And we should put them separately as a little addition section or not. Can we do that?

MEMBER SHERIDAN: I'm circling with

this as well because also what is the timing of this. Do we ask them right after an office visit? Do we ask them six months later? I would have been very pleased with the diagnostic experience with both my husband and my son. But six months later, I learned that they were harmed.

I think this is challenging. But I'm wondering if there's a -- I don't know if this is a measurement either -- but can we recommend that a question be framed for something to go into CAHPS or something to go into another patient satisfaction survey that can capture some of this. I don't know what tool will capture this and when.

CO-CHAIR MCDONALD: David.,

MEMBER NEWMAN-TOKER: I do think that these are measurement concepts in the sense that there's something measured. You could measure patient satisfaction. That's not that far from an actual measure.

But I think that the concern expressed

over whether -- again the same issue I alluded to yesterday -- you don't know whether it's a good thing if the number is up or down. In some sense whether the patient is feeling good about where they're at with their treatment or whatever it is may not be an indication as Hardeep said that they got the right diagnosis.

I think if we specify some of these things more and give more meat on the bone in the questions that we ask the patients maybe there's a better space there, closer to what Helen was getting at which is were the tests adequately explained to you, did you feel as if you understood the diagnostic reasoning that the provider was giving and so on and so forth. That may be better than just a generic were you satisfied with your diagnosis or comfortable with your diagnosis.

CO-CHAIR MCDONALD: Helen and then Helen.

MEMBER HASKELL: I think we're talking about some kind of a survey in that respect. And

I'm also going back to the -- I'm not sure which group it is, but under diagnostic error talking about reporting diagnostic errors. I do think that's critical to have patient-reported outcomes, not necessarily just errors. Because they may not know it's an error.

In fact, I think a large part of the time they don't. But if they're reporting the outcome someone else might recognize that there was an error. So I think to me the challenge is how you would divide these things. We need a lot more patient reporting on all aspects of the diagnostic process. But how exactly you'd structure it and how you divide outcomes reporting from error reporting from a survey on the whole process I think is the issue.

CO-CHAIR MCDONALD: Helen, could you give an example when you say an outcome could be reported that the patient doesn't know was an error, but it might be figured out later? Do you have one?

MEMBER HASKELL: I'll say the ones

that I've seen reported for example reactions to medications which are typically under reported by physicians. And the patient might not necessarily know that it was a reaction to a medication.

And this is a little bit off diagnosis, but not entirely. So they wouldn't know why something happened, but they know that something happened. And in a broader assessment, a clinician might be able to connect the dots.

CO-CHAIR MCDONALD: Okay. Helen.

DR. BURSTIN: I was going to say I don't think every measure can solve all of this. So I think we have to think about how they come together. Patients get results. We just talked about it. Something about patient experience I think is critical here and I really want to make sure something patient-focused/patient-voiced comes out as being at the top.

Just as an example, I think I may have mentioned at the last meeting that there was a tool developed by Glyn Elwyn at Dartmouth called

CollaboRATE. Again, it's three questions. And I think it can be something where you could take something like that and try to adapt it.

I think I read this last time. But just very briefly, these three items are how much effort was made to help you understand your health issues, how much effort was made to listen to the things that mattered to you most about your health issue and how much effort was made to include what matters to you in choosing what to do next.

There's a simplicity to that that I think the way Glyn's been using it is to say thinking about the recent appointment or thinking about your recent process. I mean just something rather than saying you have to start from scratch. There's an amazing body of work now on shared decision making that I think it won't solve everything, but I think it's a real opportunity to think about it.

I know Glyn is going to be submitting it to us as soon as we have an opportunity in the

fall. It has gotten remarkable pickup, and I think I mentioned this to you last time, it was developed with a nine point Likert scale with the intent that it would be done on a cell phone.

When you ask Glyn why it's in one to nine Likert Scale, he said because there are nine buttons on a cell phone. It's time to move to things you can do in real time in real practice and have that dialogue going.

MEMBER SINGH: Helen, exactly this is the point. How long did it take for him to do this work and was it done under some kind of research in Wyoming? Did he actually validate some of these things?

CO-CHAIR MCDONALD: Yes, it takes time to develop.

MEMBER SINGH: Right.

CO-CHAIR MCDONALD: But all of these concepts have some sort of measurement work.

MEMBER SINGH: Yes, that's what my point was.

DR. BURSTIN: I think my point is not

everything has to be a de novo effort of a data collection that will take five years. There is a phenomenal amount of work done to see what's out there that could potentially be used now to try to get some of these issues addressed.

CO-CHAIR MCDONALD: I think on that last package of the patient experience ones it's still worth keeping the four that we have for the rest of the process. And the idea is that -- go all the way down -- they could be fiddled with. But if you look at them as a set, there's something related to the experience of diagnostic care. So problems explained.

Each one of these can be drilled down on, patient reported and understanding the diagnosis. So each of these it's important that those two things they tie into the definition. Patient satisfaction, patient experience of the diagnostic process, this is the idea that there could be some things that relate a little bit more to satisfaction with steps in the diagnostic process. Some steps would be appropriate to ask

about that. Some steps might not be appropriate to ask.

Patient experience with the diagnostic process, actually the idea of how it was experienced. That's a little bit more like the Glyn Elwyn type of measure.

That's the idea here is that there's a lot of room to develop actual measures within each of these. And some may tackle different pieces of the diagnostic process and different failure points differently. Okay, Helen.

MEMBER HASKELL: And one more thing, I do think it's worth having patient reporting on specific parts of the diagnostic process. Care coordination for example which is the bugaboo of course of patients. If that is reported by patients, you might get a very different report from what you're getting from the people on the other side of it who think they're doing great.

CO-CHAIR MCDONALD: Exactly. And so each one of these from the group what people do need to know is the idea is that each one of

these you could go into the nuance of where it's worth trying to measure. That would take a whole process with a group of people, but not such a process that it couldn't happen. It's very doable.

Having looked at everything, were there any gaps that we missed that are vitally important? Okay, Mike. You can say something. Go ahead.

MEMBER DUNNE: On patient engagement number five, you don't specify who will provide the diagnostic information. And we spoke briefly about medical scribes and how they have to be somewhat certified to do this. So who is left with the responsibility? Do you have someone who's trained in this? Or should it be the primary clinician?

But you just can't say someone has to do it because it could end up being a health care provider who's not thoroughly versed in the diagnostic process.

CO-CHAIR MCDONALD: That's a good

point for us to add here in terms of just on the side note part that you need to have not only communicated in an understandable manner but in a manner that has fidelity to the knowledge base that could be transmitted from a knowledgeable person who could transmit that knowledge.

And it's right now put as a structure measure. So the process to assure that the diagnosis and the diagnostic information is communicated in an understandable manner by a person who is appropriately trained to do so.

Great. I think we've covered the gap section.

MS. SKIPPER: I just want to note that our next step would be to move into the measure prioritization criteria. We can either keep moving right into that, or do people want to take a quick ten minute break?

CO-CHAIR MCDONALD: We were going to explain how that process was going to work, and then people were going to individually do it.

MR. LYZENGA: I think that was the

idea that we were going to do another individual exercise. Talking to people yesterday after the meeting, it seemed like people found that valuable to dig into the measures.

CO-CHAIR MCDONALD: I think it would be good to explain what the process is then and then let people take a break and then come back with any questions and then start to do the process that way. If it's clear enough and people want to spend a little bit more and use the break time, they can. But they don't have to.

MR. LYZENGA: And we do also want to talk about the criteria a little bit because we had a few modifications to that and a late addition in that additional criteria.

CO-CHAIR MCDONALD: Are you guys game to go a little bit more so you know what the independent work will be before taking a break? Okay. Good.

MR. LYZENGA: I don't know if you guys have your handouts. There's actually a new

version. Yeah, the handouts from yesterday.

We talked about this a few times including on our most recent conference call. And then we had some further discussions about it internally. We talked a little bit about it with our Co-Chairs. We've landed on the two main overarching criteria of importance and feasibility, those being the two main axises behind which we would look at these measures.

I think somebody suggested which I think would be helpful is that we could end up being able to have a 4X4 box or something like that where we could see where the most important and highest feasibility measures are and those that are high importance but low feasibility maybe, or high feasibility and low importance.

Depending on what your interest is as a measure developer or a policy maker, you could look to see which are the higher feasibility ones and which are the most important ones and areas where those line up. Depending on what you're looking for you can find that.

We did add up another third criterion which is cost. Maybe I should first say that there's a lot of stuff jammed in here in importance. This is intending to keep it at least relatively simple for us in terms and low burden in terms of doing our actual ratings. Didn't want to have to go through each of these subcriteria you could call them.

But these are really in that second column considerations to take in mind, questions you might want to ask yourself when you're thinking about whether a given concept is important. How relevant is it to diagnostic quality or safety? Is this more of a general quality or safety area? is this really specific in capturing something about diagnosis?

There's this element about high priority. This is actually drawn from some work that we're doing around prioritizing measures generally at NQF. And these are three goals for measurement that we're trying to focus on as part of that process.

We have a focus on outcomes. Outcome measures are obviously higher priority. They're meaningful to the patient and that they support a systemic or integrated view of care. Just some things to keep in mind as you're considering whether this is an important measure concept.

Impact is another aspect of that. Basically, is this going to have an impact on patient health? If a provider is performing well on this, is it likely that they're going to have better outcomes essentially if it's a process or structure measure. And as part of that, you could think of it through Hardeep's lens, and to what extent is it reducing diagnostic error specifically.

And then finally actionability or the likelihood that measuring this issue will drive changes in organizational behavior and the results of the measure will drive improvement. Any questions on that? Does it make sense? Go ahead, David.

MEMBER NEWMAN-TOKER: So it's golf,

right? The lowest score is the best.

MR. LYZENGA: Yes. Actually, we're thinking that is what's here. But now I'm reconsidering that.

MEMBER NEWMAN-TOKER: Yes.

MR. LYZENGA: As I'm thinking about how we're going to tally these up, maybe we should flip that and say three is the highest importance and one is the lowest importance. We're going to have to tally these and get a --

MEMBER NEWMAN-TOKER: The only other thing is, could you just clarify for us on the time period of what's short-term? Can you scroll down a little bit?

MR. LYZENGA: Yes, the feasibility. We haven't scrolled down to that yet.

MEMBER NEWMAN-TOKER: Is short-term like a few months? A year? Two years?

MR. LYZENGA: I don't know if anybody has any input or thoughts on that. We could leave it to your discretion.

MEMBER NEWMAN-TOKER: Could we just

pick something arbitrary within here?

MR. LYZENGA: Anybody have any thoughts on what short-term, medium-term, long-term means to them? We have some of those fully developed measures that we think could pretty feasibly be implemented immediately or in the very short term. And I don't know.

To some extent you can use your judgment and discretion in saying this is a much more aspirational goal. I'm going to give this a three. The two, maybe there's a lot of wiggle room in there. But I'd welcome any thoughts from the Committee on how we might further define those.

MR. HENRIKSEN: Wouldn't it depend on the concept under review and the disease entity in terms of what's long-term or short-term?

MR. LYZENGA: Probably would.

MEMBER NEWMAN-TOKER: Yes, that's a good point, too. So the measure concept may already have some measures that are well established. But others that need a longer term

development, so that does even make it messier. Thanks for pointing that out, Kerm.

CO-CHAIR MCDONALD: Preshant.

MEMBER MAHAJAN: The way I would look at it is: How impactful are we anticipating the report from NQF to be? Say, for instance, we anticipate to help move the policy in this direction. Then the measure in the OPPE to me is a broad reaching measure which could impact to say it should happen, something like that, that we could prioritize. That would be one way to approach it.

Or someone did mention about oncology which is like the most commonly missed illness than some measures related to transmission of results specifically for oncology should then be more sharp than what you're able to measure comparative. That is one way to look at it.

CO-CHAIR MCDONALD: Martha.

MEMBER RADFORD: To me this question is all about data sources. Are they available now? Could they be developed quickly? Will it

take a relatively longer time to develop the data sources for this measure? For example, we do have some that are available now. And that's how I'm going to rate it.

CO-CHAIR MCDONALD: The other factor I think of, too, is just measure development. I mean to develop a decent measure does take some time. One way to think about this would be to think in terms of as you're looking at the group for each tab and you're having to indicate the feasibility. You could do it comparatively, right, of which ones among these have the data. So now it's really more of measurement time cycle.

Or they already have a measurement. So if there's data or there's already a measure that's already decent, then obviously it's short-term. But if there's not, then it's relatively medium-term. And if there's a bunch of other barriers in the way, then it's long-term.

MEMBER NEWMAN-TOKER: And I think Martha's point is probably spot-on that the

availability of data is what's going to determine the immediacy with which people can start to take action towards the measure and the fewer political hurdles there are.

We did our stroke measure at Kaiser in three weeks. They just -- they have all the data.

CO-CHAIR MCDONALD: But they might still have to say -- The other piece of this is whether something is possible to do at the national level. Being able to proof of concept of the measure within the Kaiser system or at NYU is different than being able to have a measure that could go through an endorsement process here.

Are we thinking about just feasibility of being able to get to the point of having a measure somewhere of this? Or are we thinking of a measure that could go through an NQF process?

DR. BURSTIN: I'll just make a brief comment. I think we are increasingly trying to think about these graduated approaches to

measurement. I think part of what we've seen is when measures go directly straight into accountability with any effort in the field to use them, see them, test them is not ideal.

So I would just put forward whatever the best measure is and if it's something that's developed and initially used for improvement and benchmarking and eventually tested and eventually brought into NQF, that's fine. But I don't want to hold up the field while waiting for the perfect measure when in fact we're just kind of nowhere on this field yet in terms of building it into anything we do in terms of clinical practice.

CO-CHAIR MCDONALD: Lavinia.

MEMBER MIDDLETON: Another way to look at the data is to look at the existing data on diagnostic errors that we have. I remind the group to the first meeting again that we had together with the continuum of where the diagnostic errors occur. And maybe we should develop the priority of the measures to make sure

that at each step of the continuum where we know that errors occur that we at least propose a measure that impacts that area.

CO-CHAIR MCDONALD: David, were you wanting to add?

PARTICIPANT: No, it was said a lot better than I could have.

CO-CHAIR MCDONALD: Okay. Martha.

MEMBER RADFORD: Just about this issue of data availability, we have a number of suggestions for structure measures which are basically by attestation which is actually a pretty fast way to get data in theory. And just to make that point.

CO-CHAIR MCDONALD: And now on our list I saw it looks like process and structure and our immediate process and all of that are right on our concept list.

MR. LYZENGA: It should hopefully be correct now, yeah.

CO-CHAIR MCDONALD: It's a good point. Structure ones are faster by definition usually.

MR. LYZENGA: And then on the other side of the coin --

MEMBER RADFORD: They are reasonably good first steps toward getting more nuanced measures in the future.

MR. LYZENGA: And then on the flip side you might expect outcome measures and intermediate outcomes to be higher importance generally. I was just saying in general we might expect also as a corollary of that outcome measures to be higher importance. I don't know what the implications are for feasibility, but it's just something to keep in mind.

MEMBER NEWMAN-TOKER: But when we're talking about feasibility, we're talking about the feasibility of moving the measurement ball forward for a particular measure concept in the immediate short term as opposed to the political correctness of including it in a pay-for-performance framework.

CO-CHAIR MCDONALD: Yes, I'd like to underscore. I mean, Helen, when you said

graduated and quality improvement given where the stage of this area is I think it's really important to take into consideration.

MEMBER NEWMAN-TOKER: I mean it's a really important point. We could be thinking about feasibility like what will our leadership accept tomorrow as a we have to do this kind of thing. And that's not what we're after. What we're after is are these measure idea things that can be moved forward quickly or not.

CO-CHAIR MCDONALD: Right. Is it possible to measure something reasonable in this space in the short-term, medium-term or long-term by some organization in some way?

MEMBER NEWMAN-TOKER: Right.

MEMBER SINGH: And I don't think we should be really focusing on either pay-for-performance, penalties or public reporting when we think through some of these things. We're just not ready. The only thing that gets close to getting ready for anything accountability-wise is something on test results follow-up.

CO-CHAIR MCDONALD: Sue.

MEMBER SHERIDAN: Something, the third criterion in cost savings concerns me that this committee would potentially rank important quality measures on what we think is cost savings. And there might be cost savings to the health care system, but there's also human costs.

So I recommend we not touch this one. I think it could be limiting without us having the expertise to really know what are we saving. At this level, it concerns me when we're trying to say we're going to recommend or rank things like cost savings

CO-CHAIR MCDONALD: Other comments on cost?

MEMBER RADFORD: Yes, I just want to agree with you completely there. None of this is supposed to result a direct cost savings. There may be indirect cost savings related to fewer diagnostic errors. But I think that's really minor, a minor consideration, for this family of measures.

CO-CHAIR MCDONALD: Kerm.

MR. HENRIKSEN: Yes, the other concern is: How many health economists are in the group here?

CO-CHAIR MCDONALD: I know I definitely had a concern with it just being health care costs because societal costs would be an appropriate thing to think about. But then you need to the expertise to do the accounting and that gets pretty tough.

One approach might be to think about something related to burden. And I don't know if there's a concept related to burden that members of this group would have a way of tapping in to.

MEMBER SINGH: Unintended consequences of some kind must be considered as we go forward. And they couldn't be built into the additional two things. But rather than cost, I would rather weigh in on unintended consequences.

CO-CHAIR GRABER: Just to defend costs, we understand that it's difficult to know what the cost savings might be and all the

comments are very relevant. But at the moment health care organizations don't see diagnostic error as a priority at all.

If there were a couple concept areas that had prominent cost saving potential, I think it would be important to know what those were as a way to get some of these adopted more quickly by health care organizations. At the moment, there are no other leaders.

The Joint Commission has no performance expectations in regards to diagnosis by in large. So I think we do need to consider costs if we want to be effective.

MEMBER NEWMAN-TOKER: I agree with Mark that the lever to push for leadership is around cost. And I don't think we all necessarily have to be health economists to be able to estimate whether a measure has anything or not to do with the likelihood of driving, over testing or under testing or whatever.

But I am worried like the rest of the group. We know that some of these things are

more or less cost agnostic, probably a lot of them, other than the cost of implementing the measure in terms of how they change people's behavior. It's hard to know whether it's going to cause people to test more or not if you educate them better or there's more communication or whatever.

So I don't think that this is necessarily something we should spend a ton of time on. Maybe people could star the ones that they think could be used as better levers for leaders.

MR. LYZENGA: I think that's maybe a good approach instead of rating each one. We can indicate by star some of the ones that are maybe more likely to contribute to cost savings or something like that.

CO-CHAIR GRABER: The key thing, Andrew, is did you want us to reverse this scale or use the one that's printed here.

MR. LYZENGA: Reverse.

CO-CHAIR GRABER: Okay. So everybody

get that.

MR. LYZENGA: We'll print a new copy of this.

CO-CHAIR GRABER: We're going to print new pages and it's going to be the reverse of what's here. And if we get it backwards, it's going to totally screw up the tally. So everybody's got to use the same ranking criteria.

MEMBER SINGH: I think we should get a verbal acknowledgment from everybody in the room that they understood that three is the highest and one is the lowest.

MR. LYZENGA: Three is better.

MEMBER SINGH: Please, everybody, raise your hand if you understand that three is the way to go if you really, really like that measure.

MR. LYZENGA: Thank you.

CO-CHAIR MCDONALD: Three is the highest. Mike, did you want to comment?

MEMBER DUNNE: Can I recommend that for cost we simply say potential for cost savings

and put yes or no? I mean how are you going to grade this?

MEMBER SINGH: The cost proponent, I'm very practical in case you haven't noticed. Can you give me examples of some measures that are cost savings or potentially cost savings?

CO-CHAIR MCDONALD: Actually, you have to give an example of a measure concept.

MEMBER SINGH: Okay. That's what I meant.

CO-CHAIR MCDONALD: That's the challenge.

MEMBER SINGH: Fine. Measure concept. Please just tell me. Just try to convince us.

MEMBER NEWMAN-TOKER: The test appropriateness measure.

MEMBER SINGH: Okay. It may be one.

MEMBER DUNNE: Turnaround time to critical results. It's going to reduce overall hospitalization in certain circumstances.

MEMBER SINGH: So the inappropriate test is if we know that the tests are

inappropriate, right. We'll have to figure that out somehow. And how are we going to do that when? Wouldn't that involve costs to figure out which tests were inappropriate?

MEMBER NEWMAN-TOKER: Not really if you do it in a symptom disease frame. It's actually pretty easy with administrative data. Somebody gets diagnosed with BPPV with an ICD 9 code of 386.

MEMBER SINGH: Outside of neurology, please give us an example.

MEMBER NEWMAN-TOKER: The fact that I've done that in neurology doesn't mean that it can't be done in other places.

MEMBER SINGH: Okay, one example. I'm just thinking of how we're going to use this framework. Other people can give examples.

CO-CHAIR MCDONALD: How about this because I bet we will debate this for a long, long time? How about just going with if you have some information that you think is relevant to thinking about costs or adverse effects of this

concept that you think are really important like write a little note about that?

I think this is something that leaders will pay attention to because it could seem like it could end up being net cost saving. Or I think nobody will pay attention to this because it will cost them way, way too much to pay attention to this.

You could write notes about that and you could write notes about whether you think this concept is really prone to severe adverse harm. Those are two things. We talked about that.

MEMBER SINGH: So use only two and then write notes on either cost or unintended consequences.

CO-CHAIR MCDONALD: Right, for the third. Don't try to do a rating. Then we can bring that back for discussion. It's going to depend on how different people are thinking about it to end up with a ranking.

MEMBER NEWMAN-TOKER: And one other

critical thing on the issue of cost, I think we should move away from the cost piece and just talk about if we think it's something that's going to drive down unnecessary or inappropriate utilization kind of behaviors or reduce length of stay.

In different payment model systems, those things mean different things. In this world, doing more testing that's unnecessary is to your financial advantage. But in a global budget system like Maryland, it's now to your disadvantage.

CO-CHAIR MCDONALD: David Hunt.

MEMBER HUNT: Yes, I just want to say that it sounds like what we really want to know is whether or not the measure is effective and improving diagnostic accuracy. In a global sense, if we get diagnoses better, then that will have an effect on the overall cost to the system. I think there's a general agreement to that.

I have a few issues with the cost thing, too. Cost is really tough to calculate

even if you have the CMS actuaries here.

CO-CHAIR MCDONALD: You're absolutely right on that side of it, too. If we're going for diagnostic accuracy, then you're going for optimal efficiency ultimately if that were gained.

One thing though I think in terms of your discussion, Mark, is this idea of organizational attention. When making a note about whether the measure concept does something that grabs organizational attention, we can just note that.

Feel free to note anything in terms of -- I think, Andrew, it would be cost burdens --

MR. LYZENGA: Cost burdens or other issues.

CO-CHAIR MCDONALD: Yes, other issues just based on this discussion you think are important in terms of conceptualizing this.

MR. LYZENGA: From a logistical standpoint, we did give you these printouts. Some of you seemed to think it was easier to go

manually in which case we'll have to enter these in and take a little bit of time to put them and tally them up.

If you'd like to do it electronically, we can send you a spreadsheet that you can use. It may be a little unwieldy. Excel can be a little buggy and things like that. If you would prefer that, we can send around the spreadsheet to have you enter in your rating in that document.

Would anybody like to do it electronically? Or would you all prefer to do it on paper? All right. We're doing paper.

CO-CHAIR MCDONALD: Feel free to take a break and do the ratings.

(Whereupon, the above-entitled matter went off the record at 10:51 a.m. and resumed at 11:45 a.m.)

MS. SKIPPER: All right. We'll move to our member and public comment period. Operator, could you please open the lines to hear if there are any comments.

OPERATOR: Okay, at this time if you would like to make a public comment, please press star and then the number one. And there are no public comments at this time.

MS. SKIPPER: Thank you. Are there comments in the room?

Hearing none, you all please continue working. Lunch should be set up soon. And I'm going to estimate that we can come back around 12:30 p.m., 12:40 p.m. or so. About 50 minutes. Thank you.

(Whereupon, the above-entitled matter went off the record at 11:46 a.m. and resumed at 12:58 p.m.)

MR. LYZENGA: Maybe we can reopen and maybe first just get some reflections on that exercise, whether you, I don't know, were finding it difficult or easy to do the ratings, if you were having any difficulty or confusion applying the criteria, and just sort of thoughts or reflections on how that went and we'll have the results, again, momentarily. But, initially, any

thoughts?

MEMBER NEWMAN-TOKER: I'd have preferred a five-point scale rather than --

MR. LYZENGA: Yes.

CO-CHAIR MCDONALD: Kerm?

MR. HENRIKSEN: Yes. We were just discussing how useful the averaged information that you get will be, because we're going to regress towards the mean here and so, you're going to have a lot of averages that perhaps certain around two, at least, I think, for feasibility.

There's probably a predominance of importance here that's almost hard to say something is not a three and so, you might get finer degradations for importance. But I'll be interested in what the averages look like and if those are useful for making fine distinctions, but maybe at the end of the day, you're not interested in maybe fine discriminations among items.

We're still sort of measuring with --

we're going to implementing with a hatchet and so, you don't really need a micrometer for measurement. And so, I was just wondering how satisfied --

MR. LYZENGA: Yes, well --

MR. HENRIKSEN: -- you will be with it.

MR. LYZENGA: -- we'll have to see. It did seem like there was a lot of threes in that importance column, from the ones that I was helping to enter in. So, I do have a sense that there's going to be not a whole lot of discrimination in that category, but we'll see how it comes out to be useful.

MEMBER NEWMAN-TOKER: So, I have to say, I tried to be ruthless in the sense that -- because clearly it's all important, right? No one's arguing that -- I mean, we've come to this point, we've all argued a lot that all of this stuff that we've put out there is important.

So, I guess what I tried to use, for me personally anyway, was this question of, if I had the measure, how likely would it be that

having the measure alone would move the needle on the diagnostic error problem?

And I do think if you frame it that way, instead of saying, is this an important idea, like, that we should all do, maybe there's more room to differentiate between threes and non-threes.

CO-CHAIR MCDONALD: I think there's also, I mean, because we went through -- part of this is process, right? So, we've gone through a process where we've been in subgroups, presented, but the others haven't gotten to say, I agree with you Subgroup 1 or Working Group 1.

So, I think it just allows us to be sure that the entire Committee does actually feel that the set are an important set and nobody says whether a set should be ten or 20 or 50 or 1,000 concepts. I mean, there's not a number that we have to shoot for.

MR. LYZENGA: And we've also got, again, a large amount of sort of rich discussion around this that's going to help us add context

to whatever the ratings are for each of the concepts.

CO-CHAIR MCDONALD: David?

MEMBER HUNT: I was going to say, actually, as someone who has actually used one of the products from our last work on Health IT Safety Measures that we did with NQF, as someone who has actually used this end product report, is the discussion that really it adds so much flavor and texture to it.

So, the numbers are important and they're good, but it's the context of the discussion that really produces a lot of value for us. And the -- I will say, also, one thing that we found is that the feasibility numbers are important, but as, I think it was Martha who said, feasibility in many ways is just a synonym for data availability.

And when we see something that may have been, say, a low score in feasibility, but then we realize, we actually can get that data, it makes all the difference in the world to know

what the importance of this then was for the entire group. So, it's of tremendous value and discriminating between those two makes a big, big difference.

MEMBER SINGH: So, I was just thinking, in the next couple of hours that are left, we're going to have a discussion on these, just to sort of go through and see? Because I think the main message here should be what kind of an actionable report we can produce so that people like you would get the most bang for their buck and it doesn't look like it's just another ivory tower report that nobody's going to do anything about. So, I think that sort of should be our --

CO-CHAIR MCDONALD: I think that's the idea, right? We're going to look at what these scores were, not so much discriminating, but just it will allow us to sort of see them in some order, and find out if there were some that got sifted out or some that just looks like they're nailing it and then, talk about the usefulness of how to present the information and what else is

needed. Yes.

MR. LYZENGA: Something else we kind of wanted to get input from you on is, it would be helpful for us, again, as we're talking about writing the report, are just kind of overarching issues. Any issues that sort of cut across these different concepts and measurement areas that you think would be important for us to highlight in the report, to discuss some of those.

Again, the things that have come up are burden, and we've had some discussions around that and we'll try to flesh out a little bit of the Committee's discussion. If you have anything to add on that issue, we'd welcome any thoughts.

Anything else that's emerged from this, sort of these exercises in the last couple of days that you think are worth talking about. I know I'm going to pull some things out of the previous discussion and I think we've got some sort of themes that are emerging here that we're going to talk about. But anything that you would -- that came into your mind that should be

highlighted as part of this.

One thing I also kind of wanted to get just an initial thought on, and I think maybe we may seek some further input from you after this call is, is the question of accountability, whether we've seen any concepts or areas of measurement that may be suitable for things like public reporting and even maybe payment and whether -- how far we are away from that.

If you don't see any at this moment, but if you do see any areas that might be promising and what those might be, or just any thoughts in general on that issue that we can, again, kind of work through in the report a little bit and have some discussion about. So, I'll just sort of open it up, does anybody have any thoughts?

MEMBER SINGH: And are we thinking maybe, like, what would be just one or two things that the organizations could relate to and say, this is so important we need to work on it, and it might get close to some kind of an

accountability measure and we prioritize just that one thing and that could come out in the report? Is that what you're trying to get to?

MR. LYZENGA: It doesn't have to be, like, something that we identify as, this is the Committee's highest priority, but we may want to sort of highlight a few issues or we may want to have, say, a section on accountability and say, the Committee identified some areas of measurement that may be promising for public reporting or payment in the future.

As these measures get developed, these would be something that we could -- as the field tries to move towards accountability, as measures are developed, where should we be looking? Does that make sense?

CO-CHAIR MCDONALD: Yes. I mean, I think that's a discussion topic for sure.

MR. LYZENGA: Yes.

CO-CHAIR MCDONALD: So, we tagging it as a discussion topic, do we want to try to discuss that now --

MR. LYZENGA: Yes.

CO-CHAIR MCDONALD: -- while we're still -- and, Martha, you had something to say, I don't know if it was on this.

MEMBER RADFORD: It was kind of, I just wanted to say that, first of all, I gave a fair number of ones, I don't know, mainly because I felt like they were a little bit tangential to diagnostic quality.

Second thing was, I really look forward to what I hope is the next stage of this, which is a draft of the document, because I think that the discussion has been wonderful and very rich and will provide, if you can summarize sort of the take-homes from that in the document around -- I think that will provide a tremendous amount of context. Thanks.

CO-CHAIR MCDONALD: So, I think --

MR. LYZENGA: And we hope you are not going to -- this isn't going to be the last time you provide input on that. We'll try to give you drafts and what have you.

MEMBER RADFORD: No, I'm sure it won't be. But do we have another in-person?

MR. LYZENGA: Another in-person --

MEMBER RADFORD: No? Okay.

MR. LYZENGA: -- we do have another, I think we have two more calls and we'll be actually talking about that in a moment. But hopefully we can also just do some work online via SharePoint or email and just get input, share -- circulate some ideas and drafts and things like that.

CO-CHAIR MCDONALD: Kerm?

MR. HENRIKSEN: Yes. Going back to Andrew's earlier question about accountability, there are -- and I imagine there's not much that is being recommended for those purposes from comments from Hardeep and David, but there are other uses for the information that's being gathered and it all depends on how you package the materials.

So, there may not be measures for public reporting, but there certainly are for

quality improvement projects that anybody can undertake. And there's also other measures that are more relevant to research.

And so, in terms of how the report is sort of divided into modules or chapters or sections, do you sort of help the broadest audience possible that might be interested in diagnostic safety to find their place in the report so that they can -- so that there's something in there for them --

MR. LYZENGA: Right.

MR. HENRIKSEN: -- sort of action items or things that are within recent --

MR. LYZENGA: That's definitely a helpful suggestion and we can try to, yes, incorporate some concept like that, a way of indicating for different audiences who are looking for different things --

MR. HENRIKSEN: Yes.

MR. LYZENGA: -- where they can find what they're looking for.

CO-CHAIR MCDONALD: Yes, that's a good

idea. Would it be worth articulating what maybe the top three or four or five audiences are?

MR. HENRIKSEN: Yes. Well, I would think researchers, for one.

CO-CHAIR MCDONALD: Researchers.

MR. HENRIKSEN: And then, the other is closer to the clinical practice, in terms of things that they can be doing. And, as David mentioned, this can help start a conversation for people that are really looking -- it's not so much the value of its importance or feasibility, but it helps to start a conversation.

CO-CHAIR MCDONALD: So, maybe it's researchers, it's the delivery system, it's those who are involved at policy level.

MR. HENRIKSEN: Yes, policy level.

CO-CHAIR MCDONALD: Okay.

MR. HENRIKSEN: So, those are the major groups.

MEMBER HASKELL: And patient advocacy.

MR. HENRIKSEN: Yes, patient advocates. Yes, definitely. Patients themselves or patient

advocates --

MEMBER HASKELL: Yes.

MR. HENRIKSEN: -- definitely.

MEMBER SINGH: Yes. In this day and age, when there's so much lack of emphasis, I would say, on scientific momentum to push this forward from many stakeholders who should be involved, could we actually just outline some agencies who would take this work forward and especially say what's in it for them.

Why would CMS want to do anything with this report? Do we have a good case? Can we build it? Can we build a paragraph which is, this is why CMS should be interested in this report? And I'm going to pick on my friend David here as well, this is why ONC should be interested in this report.

And I think, that is a type of an accountability that I think we need to be giving. It's a shared responsibility, just like we say for Health IT, improving diagnosis is also a shared responsibility and I think we need to try

to think about, who are the stakeholders who are going to take the brunt of moving this forward?

CO-CHAIR MCDONALD: This is -- okay, so that was a little bit on making this useful. If there's more ideas, let's get through those. Yes, Martha?

MEMBER RADFORD: We've already mentioned specialty societies as being an important contributor and they've got skin in the game, potentially.

CO-CHAIR MCDONALD: And they're professional societies, since Marilyn's left.

MEMBER SEIDENWURM: I don't think we have to do a whole ton of work getting CMS interested in this, because it's in the QPP and it's a major priority and they've actually emphasized that now for at least one or two or three rulemaking cycles. And I don't know which -- they overlap. So, I think CMS gets that this is important. Now --

MEMBER SINGH: They are, I was just making a point. Yes, it's actually in their 2016

report --

MEMBER SEIDENWURM: Right. Which we --

MEMBER SINGH: -- for diagnostic measures, yes.

MEMBER SEIDENWURM: Right. Which we helped write.

DR. BURSTIN: It couldn't hurt to just be very clear about what the actions are. And actually, I was just saying to Andrew earlier, I think a lot of what we've talked about has blended, has sort of sailed between measure concepts and, really, recommendations for the field broadly. And so, I think we can try to tease that out.

And I was actually suggesting to Andrew, maybe we could actually send it around to -- the report, one of the reports our committees did a couple years ago on rural health and specifically how to handle low volume providers, it's maybe a little off topic, but the way they organized it in terms of a clear set of recommendations of what needs to be done, beyond

even just the measure framework and the concepts, I think is an important element.

Because so much of what has been said here fit into that almost policy recommendations, but not yet really a measure. And I don't want us to feel like you're only in the box of everything, because it's NQF, being a measure.

And I'm glad David's nodding his head, because I do think some of this is about standards for EHRs, some of this is about policy direction, some of this is about measurement. And, again, CMS funded this, so they fully recognize they want measures in this space.

CO-CHAIR MCDONALD: Good.

MEMBER HUNT: And I'll just say, particularly when you're working in policy, having a good idea of what is an area for measurement or is not saves a tremendous amount of time.

It allows you to triage things, to, okay, well, this needs to head over to this group, rather than this group and we can use this

as a good solid meal for starting some measurement development.

Having done the work of discriminating what is in and what's out of different spheres, what is good for quality improvement activity, but perhaps not suitable for reporting purposes on a national basis.

These are the questions that are always brought up when we talk about actually starting a new initiative or starting a measurement initiative. So, all of this is incredibly helpful.

CO-CHAIR MCDONALD: And, actually, as a segue from that, so you -- we're not quite ready yet with -- are we ready for that? Because, I mean, I think -- so, we could look at that.

Maybe as we look at that, one of the lenses we should look at it is through that lens, Andrew, you're asking for, which is, are some of these areas more important for pushing towards some sort of accountability?

I mean, my own personal view is that the measurement field in this space is very far from being able to be effectively used for accountability in a way that we can feel confident would actually improve diagnosis. That's my own opinion writ large.

But as we look at this at a more granular level, there could well be some areas in here that would be ones where you would think, hey, it would be really good to have measures in here that make the system accountable. So, we can try to tease that part out and have comments on that too, so that NQF has that as they're working on the report.

MR. LYZENGA: Actually, just before we see this, as a caveat, we entered this all pretty quickly, I haven't seen the results and we may have had some data entry issues. I can't be confident yet, so this is sort of a preliminary look. We'll go back and kind of clean up and reconcile everything and get you final numbers, but we can at least get the first look here.

MEMBER HUNT: Can I just go on record that, after today, that someone owes Vanessa a lunch?

(Laughter.)

CO-CHAIR MCDONALD: Yes.

MEMBER NEWMAN-TOKER: Just to follow up briefly on David's excellent point about the helpfulness of knowing what's not a good measure. I think, in terms of getting, not just sort of, these are areas we think are promising and these others here should go to policy things, but even a little richer.

Like the context of, we discussed this idea, right, that came to the table and the reason why this wasn't a good measure was X, Y, and Z, so that people don't think -- because there is a tendency for people to think that we're just sort of thoughtless idiots when we leave stuff out, as opposed to -- at least me anyway.

And so, the -- I think, really, bringing that in in a richer way somehow and kind

of explaining the logic of why something wasn't included, will stave off a certain amount of criticism from all the people who have their own passionate interest in a particular measure.

MR. LYZENGA: Sorry, it turns out we're not quite ready yet.

CO-CHAIR MCDONALD: I kind of object to summing between importance and -- yes. Yes. How about doing -- how about just leaving those two as they are and just sorting it first by importance, letting us look at it that way, and then sorting it, after we've looked at it by importance, on feasibility.

So, let's look at importance first. So, they've got a summary for each concept. They do, they have a summary for importance and they have a summary for feasibility.

MR. LYZENGA: So, here's our highest rated concepts. I'm not sure how that happened. Yes, we've got some higher than three ratings, I'm not sure how that happened. All right, we'll go with what we -- again, we'll clean this up.

I'm not sure how that could have happened, but --

CO-CHAIR MCDONALD: They're probably relatively still about right, so --

MR. LYZENGA: I would guess, yes. It's divided by the -- so, maybe we got that number wrong. We were going to go in and reconcile that too, so that -- it's the number of raters, yes, so we may have missed one.

CO-CHAIR MCDONALD: Yes, the importance is there and it looks like it's still ordered the same, at least as far down as we can see. So, the top one has 44 as the sum.

MR. LYZENGA: Right. There were also some blanks that we needed to go in and --

CO-CHAIR MCDONALD: Yes.

MR. LYZENGA: -- clean up, so that may have thrown off our -- yes. I did take your names out of there before I did that. So, that's -- we're still basically in the same place, but these are the -- those are the totals.

CO-CHAIR MCDONALD: That's good. Can you make it just a tiny bit bigger?

MR. LYZENGA: So, we'll clean up those averages, figure out --

CO-CHAIR MCDONALD: Yes, that's nice.

MR. LYZENGA: -- what went wrong there.

CO-CHAIR MCDONALD: Okay, yes. This way we can look at about five at a time, which is --

MR. LYZENGA: A little bit more.

CO-CHAIR MCDONALD: There we go. Yes, we've got five there, that's good. So, the first observation, we have measures from five different tabs in the top five. It's kind of nice. We want these, yes. Okay, there we go. Yes, now we've got next set of four to five.

MR. LYZENGA: We'll take a look at that too, see if we can get some ranges and --

CO-CHAIR MCDONALD: You can go to the next, yes. So, a little bit more. One more. There we go. It took a while to get to diagnostic error as a subdomain. There's not much difference here, we're still at 41, these are all tied, these are all still tied.

Okay, show us some more. Yes, this is Kerm's hypothesis, yes. Oh, look at that. Okay, yes. Actually, go down to the bottom. Go down to the bottom and let's move up from the bottom now. Okay. So, starting right there. Yes, the bottom one is not a measure, so some people --

MR. LYZENGA: Oh, right, yes.

CO-CHAIR MCDONALD: -- realized that, probably. But there, the composite, that was a measure. So, the composite was the lowest of our group. Make it bigger? There.

MR. LYZENGA: Yes.

CO-CHAIR MCDONALD: And now you can start going back up. So, just -- okay, yes. That will -- no, like, yes, we'll just -- yes, go up just the next chunk, sort of chunk of four or five to chunk of four or five.

(Laughter.)

MEMBER SINGH: I lost count, are we kind of in the middle or are we --

CO-CHAIR MCDONALD: We have that too, it's a mess. We're close to the bottom --

MR. LYZENGA: Still close to the bottom.

CO-CHAIR MCDONALD: -- right now. So, we're doing, like, the utility analysis where you go high, low, high, low.

(Laughter.)

CO-CHAIR MCDONALD: Don't get anchored. The next five is still near the bottom. Okay, yes. You can go another five up. Okay, go ahead up again. Special one.

MR. LYZENGA: I can't remember why that was --

CO-CHAIR MCDONALD: It got de-escalated.

(Laughter.)

CO-CHAIR MCDONALD: Yes. There was another patient experience that was lower than that one, yes. Okay, go on up. Okay. Yes, exactly, we're in the middle of the pack now. You can go again.

MR. LYZENGA: You can scroll up one.

CO-CHAIR MCDONALD: Yes. Pop up one

more. I mean, another chunk, yes. There we go. Okay, again. Yes, now we're --

MEMBER SINGH: This is by importance, right?

CO-CHAIR MCDONALD: Yes.

MR. LYZENGA: Yes, this is by importance, we haven't --

CO-CHAIR MCDONALD: And by feasibility. And I'm also thinking, maybe what would be nice when we look at the next category is to actually go ahead and categorize by subdomain first and then, like, feasibility second, just to see within our domains, because we might have an easier time seeing it that way.

But go up, we're almost -- we have to get back to 40. Yes, but we're still looking at importance, we haven't quite hit 40 yet. I mean, now we're in the top half. Okay. Okay. Okay. There's another closed loop one that's really high that was in the first five, it was the second, I think. Yes. Yes. Yes, the other external environment one, it was low on

feasibility too. Yes.

Okay, go on. Okay, we're back at 40, so there may have been a whole batch of forties. Okay, go up again. Okay, go up again. Up a little bit more. We haven't seen these yet, so we're still moving back to that place where we'll bump up -- go up. Yes, I think we -- yes, we're now back in the top range there. Okay.

Everybody want to see the top again, the top five? And then, you can resort. See, there it was, closed loop was the second one. Yes. You just want to see the feasibility list through 40 of them or do you want to look by subdomain at feasibility? You want to see all? Okay. Feasibility for everything, yes. Yes, okay. Yes.

MEMBER SINGH: You just sorted it by feasibility?

CO-CHAIR MCDONALD: Yes. Though this next one is an outcome -- that's -- yes. Scroll down, there we go. Yes. Okay, go ahead, go down another set. Want to go down another set? Yes.

MEMBER HUNT: Remember that analysis by committee is like art by committee.

CO-CHAIR MCDONALD: Yes, we can't -- yes, exactly. And I'm thinking one of the ways to have the conversation is to be looking at maybe the top ten, just to sort of say, okay, here's the importance rating, here's the feasibility rating, kind of, what do people think of that? What's sort of the -- yes.

DR. BERNOT: And, also, just to even validate that we look at the top ten and see that that is actually the ones that we think should be there?

MEMBER SINGH: So, will we do the top ten for importance and then the top ten for feasibility? Like that? And that way, there could be a little separate discussion. And then, some will be in top ten for both, so you'll only end up doing, like, 14 or 15, maybe, max.

CO-CHAIR MCDONALD: Use your microphone, Martha. Yes.

MEMBER RADFORD: I'm not sure why we

need to discuss it, because, I mean, we have discussed these measures pretty much ad nauseam.

(Laughter.)

MEMBER RADFORD: And I think that, since there's some consensus about what's important, that's something that we can celebrate, frankly. And there's enough votes on all of these to say we have some consensus here. I don't -- yes.

I mean, we're not going to -- this is not going to be the final ranking, by any means, and maybe we won't even have a final ranking. I, personally, feel like we're kind of done, in the sense that the next thing I want to see is a draft.

CO-CHAIR MCDONALD: That will be over the next lunch hour.

MEMBER RADFORD: With a bunch of tables, yes.

(Laughter.)

MEMBER RADFORD: No. We're not -- we're going to let you have a couple days off

there.

DR. BURSTIN: I'm not sure if summing it makes sense, I think they're very different concepts. I mean, traditionally, what we've tended to do is, it's really a four-quadrant map, right? You have importance on one axis, you have feasibility on the other axis.

The measures that are ready for prime time, David could pass off, we could pass off and try to get measured, are going to be in the high importance/high feasibility quadrant. We can separate them all out that way.

There are going to be some that are going to be high importance/really low feasibility and those are going to be important and those are probably on a slower track until we figure out how to do them.

So, there's nothing about this, I'm with Martha, there's nothing about this exercise that's going to get us any closer to, at this moment, doing that, other than you clearly -- I think the face validity of, did the top ten on

importance look like they were important? And I saw a fair number of heads nodding. And I think we can then go back and do the analysis and then share back with you sort of the quadrant analysis.

CO-CHAIR MCDONALD: Yes, that's what I --

DR. BURSTIN: Or even lay out for you the importance measures across the range of feasibility. And if you could even help us identify, so, based on this, which are the ones we should really push on trying to get developed in the next one to two years? Which are the ones that are really important, but are probably going to take three to five years, given data sources?

Which are the ones that are just so important, but you know what, they're probably never going to be a measure, but maybe they should be a standard for the Joint Commission, maybe they should be a standard for ONC, maybe they should be a policy direction for CMS. And so, I think that's where there's a lot of work to

tease this out and to do that, you've got to really sit and be with these data.

MEMBER SINGH: And you're almost suggesting there should be another meeting face-to-face.

(Laughter.)

DR. BURSTIN: I don't know that we have the budget for that, but we will certainly do all the pre-work so that that next meeting, whether it's virtual or whatever, I think we can get the work done. Yes.

MR. LYZENGA: And we'll, again, we'll clean this up, we'll send it back to you so you guys can play with it and kind of sort and cut it up how you'd like.

And then, we'd like to get some discussion going, again, via email or maybe a discussion board or something like that and try to work through some of these issues that Helen was just talking about and, again, crosscutting things or where we want to -- what we want to emphasize as part of the report and how. And

guidance on all of that would be very much welcome as we get into that.

CO-CHAIR MCDONALD: Good. Is there -- oh, go ahead.

MEMBER HUNT: That way we'll get a chance to hear the voices of those who had to leave early and -- yes, that'll be important.

CO-CHAIR MCDONALD: Okay. So, we've gotten to look at this. Now, the question is, what else do you need from us to help do the work you're going to have to do when we're not sitting here? And there's still enough of us here that --

MR. LYZENGA: Yes. I mean, again --

CO-CHAIR MCDONALD: -- use us how you need us.

MR. LYZENGA: The most, from my perspective, the most important thing we need, again, is the discussion, sort of broadly construed what do we want to address through this report? How do we want to present this? What are the large issues we want to highlight?

Again, we'd like to kind of continue that discussion as we move forward, but just any thoughts that you have.

Again, reflections on the last couple of days, things that you think that we ought to cover or anything you think would be helpful for us as we begin to craft this report. Just talk at us, we'll get the transcript, we're going to go through it and we'll pull things out, so just give us material.

MEMBER NEWMAN-TOKER: Well, let me start by saying that you guys have done an amazing job. A round of applause.

(Applause.)

MEMBER NEWMAN-TOKER: Obviously, dealing with a complicated subject, an unruly crowd with at least one loud mouth, and so, we really appreciate all that you've done, for making this happen.

I think, for me, the same issue that's come up repeatedly, this issue of granularity, is I think kind of central to the way that this gets

framed into a report. Because, clearly, not all -- some of these were more like measurement themes or ideas and others were more like measure concepts that were one rung down in the food chain.

And I think evening that out in some way with one or two people at the helm of kind of maintaining that consistency as you guys write the report will be a critical place for you guys to massage what we've done here into something others can digest.

And I would advocate for making sure you try to maintain the level in between the subdomain and the measure concept, because I think, what I found for myself at least during this process, was that when I got to write things down, I just found myself jotting down whether we had covered stuff.

I had errors, harms, accuracy, timeliness, reliability, and appropriateness. Just, like, if I had those as kind of the sub-headers within some of these fields, the

subdomains, at least I kind of made sure that there was something there substantive that I could latch onto, like I'd covered this space.

And I think for people who are coming to this for the first time, if they have that sort of intermediate step, they'll kind of get it, that we kind of covered each of the subdomains a little bit more thoroughly. So, I think that will be the most important thing for you guys to do to help us.

MEMBER HUNT: And one thing I should say, again, as someone who has used one of the products of this type of process, you can rest assured that, particularly because of the strength of the NQF process in general in terms of when a measure actually gets up for endorsement, you can have faith that a measure developer will look, and if something's a concept, even if we think that it's a pretty good measure, and they start to say, well, what's the numerator, what's the denominator, what are the exclusions, if it doesn't shake out, they'll work

with it, tease it, and try to make it into something.

Because one of the beauties of having NQF is that the process now for a measure that actually is endorsed is pretty solid, that there's a rigor and a structure to it that you can be sure that at the end of the tunnel that it will come out with some validity and strength.

MEMBER RADFORD: I say this with some trepidation, because what I'm really doing is volunteering to help you, as opposed to you volunteering to help us, which is, it's possible that some of these concepts might benefit from drafting specifications. Which is kind of what I heard in your comment as well.

And as someone who basically does this on a daily basis, I'm willing to do that for a few and maybe other people who do this also might be willing as well. And I have a very, very sort of compulsive way that I write measure specs and I'm happy to share that with you. Just, I mean, I can't do it for all of them, but I can do it

for a few. So, pick your poison.

MR. LYZENGA: That's not a trivial offer, that's tremendous. That's -- wow.

MEMBER RADFORD: Well, this -- I know it's not going to be the final --

MR. LYZENGA: Yes, but --

MEMBER RADFORD: -- but at least it gives somebody an idea of how someone might measure it, yes.

CO-CHAIR MCDONALD: Yes, I think that's actually an incredible offer and it would behoove us if you see any concept on here where you can do that, as just sort of -- I mean, they're almost like -- yes, examples. Yes.

MEMBER MIDDLETON: And along that vein, I'm happy to add content to the FPPE, OPPE, and examples for diagnostic accuracy, because that's kind of where I live right now. I'm happy to do that.

CO-CHAIR MCDONALD: Actually, examples, I'm thinking examples tagged a little bit to that context from which you sit would be helpful too,

because many times a measure can be made because you have access to particular data or a particular frame of reference and it shows that, here's an example of how it would be done in an environment like mine. So, I think we would want to include that too.

MEMBER SINGH: I was just going to add, as we think about this, should we also sort of think about who the stakeholder would be whose measure this would most apply to and who could make things like this happen?

Again, I'm just picking on this one since it's the chief complaint from the EMR, let's say, if we go forward with this, I mean, this may be something that the institutions could also do, but this is something that your shop, David, could do very well. And so, I'm just thinking, do we need to sort of also think about sort of --

MR. LYZENGA: Yes, that would be really helpful --

MEMBER SINGH: Yes.

MR. LYZENGA: -- I think, if we could think about --

MEMBER SINGH: And that would make it more operational. I like Martha's idea, I'm not sure how you would want to write that, because I was thinking, okay, how she's going to write that?

But when you start thinking, some of these things might sort of drop off as recommendations, not specifically measures. I'm not saying this one will, but I just think that might help us determine, what's the actionability that's going to come out of it when somebody looks at this from, let's say, a healthcare system perspective.

And they're going to say, well, I can't do anything, this is under ONC's realm. And, therefore, we need to be thinking proactively that this should be moved somewhere else.

MR. LYZENGA: Yes, I think that would be really helpful, if you could think about, yes,

for each of these who might be the accountable entity or accountable entities. If you had a measure of this, that might help think through how you would then further specify the measure and kind of flesh it out.

CO-CHAIR MCDONALD: Also, where the data might be coming from. Any other comments? We're in free range comment zone.

(Laughter.)

CO-CHAIR MCDONALD: Okay. I'm hearing that we --

MR. LYZENGA: Should we do --

CO-CHAIR MCDONALD: -- can probably do public comment.

MR. LYZENGA: -- public comment?

CO-CHAIR MCDONALD: Yes.

MR. LYZENGA: Operator, could you see if anybody on the line has public comment?

OPERATOR: Yes, sir. At this time, if you would like to make a comment, please press star, then the number 1. And there are no public comments at this time.

MR. LYZENGA: Do we still have Paul in the room? No comments? You did have something you wanted to maybe mention and you can if you want, or we can --

MR. EPNER: Oh, okay

MR. LYZENGA: -- so, we --

CO-CHAIR MCDONALD: Okay. We have Kerm with his card up, too.

MR. LYZENGA: Oh, okay.

CO-CHAIR MCDONALD: So, Kerm, go ahead.

MR. HENRIKSEN: Yes. Just going back to the last comment. One way of sort of pointing out the relevance of different stakeholder groups for a particular measurement concept is through a bubble chart. And you can do that in various ways.

You can do it for each individual concept, just little circles with the initials of the stakeholder group, and you would have, maybe, it might be relevant to three or four different stakeholder groups, but at least they know they have a place at the table and that they've been

thought about in terms of the measures or the concepts that are being evaluated.

CO-CHAIR MCDONALD: Human factors visualization? Paul?

MR. LYZENGA: To introduce, Paul has been doing some work with the Society to improve diagnosis in medicine and I think -- and among the other things he'll talk about, located some examples of measures related to diagnostic quality that might be helpful as we try to spec these out, I think they may relate. But I'll let you kind of talk through what --

MR. EPNER: This will be quick. So, there's something called the Coalition to Improve Diagnosis. Several of your organizations are already in it, 31 organizations are currently a member. We have three collective actions.

One of them is to do a, try to do a census, literally, of any intervention that anyone is trying anywhere to improve diagnosis. Just, what's going on in the field? And one of the questions in this very short survey is, do

you measure diagnostic quality? And, if so, what is your measure?

And so, as of this -- this will be open all April, this survey. As of today, we only have 186 respondents. But in just scanning the responses an hour or two ago, and I showed them to Helen, there was -- I saw -- first thing, it's amazing, 70 percent are saying they measure diagnostic quality.

And you look through their answers -- they say they do, but they don't. So, yes. So, when you look at the answers, they just lump it in patient safety or lump it in quality or something like that.

But one of the first ten that showed up was, there was this site that was measuring day one, day three, and final day diagnosis discordance. So, that's actually being used somewhere, which is kind of surprising to me.

So, Helen suggested we send each of you the link to the survey, not the results yet, it'll be open all the month of April. To the

degree that you can further disseminate it, AHRQ was kind enough to send it to 55,000 people, they told us to expect maybe 100 respondents, but we're asking every one of the 31 organizations to disseminate it.

It's open all of April, it's a very short survey, you can take it yourself. And, again, it's nothing statistical, we're collecting minimal demographics. High response rate from risk managers, high response rate so far, relatively speaking, from academic medical center, lower community, lower primary care, et cetera. But we're hoping we'll get hundreds of responses by the end of April. Any question about that? So, you'll see it from Andrew. Thank you.

MEMBER SINGH: Yes. I just wanted to add something. This is a very common response, I know some of you were surprised. Many organizations will say, yes, we address diagnosis all the time, because we're sort of teaching people, residents, the clinicians are taking CME,

we're improving our diagnostic skills.

There's a lot of difference, and I know this came up in the first report and it didn't come up in this -- not the first report, the first session, but it didn't come up in this one, I think there's a difference between just trying to do stuff to improve diagnosis and trying to do actual stuff to reduce diagnostic errors. And I think there's still a lot of confusion in the field, oh, yes, we're improving diagnosis, so, therefore, we are solving the problem. It is actually not true.

And I think as you couch the report, also, we need to be careful, and that's why the preventable diagnostic harm comments that I made before, we need to be able to make sure that whatever measures or measurement concepts that we're coming up with directly address the safety component and the preventable diagnostic harm, rather than couching it as, oh, just general improving diagnostic quality, because we're not getting anywhere with that.

We haven't got anywhere with that for the last several decades. And unless we move the needle on sort of safety and preventable diagnostic harm, it's going to be hard. And so, I think we need to sort of weave that into the report. At least, that's what I would recommend.

MS. SKIPPER: Kerm, did you have a comment? No? Okay.

MR. LYZENGA: David?

MEMBER HUNT: Yes, I would just agree with what Hardeep said. That we've got to make sure that people understand that, just because you're a healthcare organization and you do diagnosis, that you're not necessarily improving on it and you don't necessarily do it with any rigor or any method, necessarily. And part of the reason for this whole effort is that we need to take it from an art to a science, maybe.

MS. SKIPPER: David?

MEMBER NEWMAN-TOKER: Yes. I mean, just to flip -- to say it in a slightly different way, but echo the comments that were just made,

in a systems engineering framework, every system produces exactly the number of errors that it was designed to produce.

And that's what we have now. Like, the old way of doing things, whatever it is that we think we're doing to improve diagnosis produces exactly the number of diagnostic errors we get, which is way too many. And the question is, what can we measure to improve beyond where we are today, rather than what can we say, we patted ourselves on the back, that we've done in the past?

CO-CHAIR MCDONALD: Okay. I think we're at the point where we've given a lot of comments, had a great -- a lot of wonderful inputs, some volunteering for the future. What do you all want to do to sort of set us on our way?

MS. SKIPPER: I have a couple of announcements. Yes. So, first, thank you, Kathy, for stepping up to help co-chair the meeting the past couple of days, you did a great job.

(Applause.)

MS. SKIPPER: And thank you to all of you in the room for your help and hard work over the past couple of days and for looking at these measures over and over. And I think we're in a much better place.

So, following this meeting, we do have a webinar, originally scheduled for Tuesday. That's a little bit too close for us to be able to sort of clean up some of the work that we've done over these two days, so we want to reschedule it.

I'll be sending out a Doodle poll for dates, either the 25th or the 26th. So, I'm hoping that will work for most of you, but I'll be sending that -- no, Martha's shaking her head no. Well, we're going to do our best to try to schedule a follow-up webinar that works best for everyone.

And then, you'll see a couple other emails, hopefully one other email, from me containing the information that Paul shared. And

there was also information that Sue shared and mentioned about a CMS special session, QIO 12th Scope of Work, she wanted to share the information with you all, so I'll also be forwarding that link. And then, the discussion guide or the slide sets that you all put together of your report-backs, in case that will be helpful in any way for you.

So, those are my announcements. We will, once we have that follow-up webinar, we will be working very hard to write the report and we will be posting it for our second public comment period on May 16.

So, that's right around the corner, so we have a lot going on over the next couple of weeks. And then, we'll reconvene you all again to sort of respond to any comments received on the draft framework, the report, and your concepts.

MEMBER SINGH: So, you will be sending the report to us before it goes for public comment or not?

MR. LYZENGA: Yes.

MS. SKIPPER: Yes. I was --

MEMBER HUNT: If I could just say, you mentioned the CMS and the information on the 12th Scope of Work, and that's where for the QIO program, I hope I'm preaching to the choir, that's for the QIO 12th Scope of Work, correct?

MS. SKIPPER: Yes.

MEMBER HUNT: Those of you who don't know, the QIO program is basically the CMS army in the field for quality improvement. And there's a tremendous amount of power that can be leveraged by getting things into the statement of work or the scope of work that these organizations across the country do.

And so, many times, you hear statements like, wouldn't it be good if everyone could start working on X, Y, or Z, and this is one way, this is -- CMS has gone through a number of iterations of developing how they decide what these organizations do.

So, if you have any interest in a

national effect on, not just this, but virtually anything in quality improvement, I really encourage you to take a look and get involved as much as possible.

Before, it was a very cloistered, very insular process of deciding what it was that the QIOs would do and now, this is opening up tremendously and I encourage everyone to try to support that if you can.

MR. LYZENGA: Okay. Well, yes, thank you all so much. It's been fantastic.

CO-CHAIR MCDONALD: Thank you.

MR. LYZENGA: We really appreciate all of you.

CO-CHAIR MCDONALD: Thank you to all of you. You guys have just done an absolutely tremendous job. I mean, we've been saying it, but we really mean it.

MR. LYZENGA: Same to you. And safe travels home.

(Whereupon, the above-entitled matter went off the record at 1:58 p.m.)

T

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: 04-13-17

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R Gross

Court Reporter