NATIONAL QUALITY FORUM

Moderator: Sheila Crawford
March 18, 2014
2:00 p.m. ET

Operator: Welcome to the conference, please note today's call is being recorded. Please standby.

Jesse Pines: Hi everyone, I'd like to welcome you today to the Patient Safety Workgroup 3 Subcommittee meeting with the National Quality Forum. This is Jesse Pines on a faculty at GW and a consultant here at the NQF. We're going to I think wait for probably one more minute before we get started. There are a few folks still calling in.

Hi, operator, has everyone called in? Who's been trying to call in? Operator, do we still have you on the line?

Operator: Yes, you have been joined to the main conference.

Jesse Pines: And is there anybody still trying to dial in?

Operator: I do see two lines. I'll step away and get those. Thank you.

Jesse Pines: OK, thank you, operator.

OK, why don't we go ahead and get started. We actually have a packed agenda today. Again, this is the Workgroup 3 call for the Patient Safety Committee of the National Quality Forum. I'm Jesse Pines. I'm a physician and faculty member at GW and also a consultant at NQF and let's – introductions here of NQF staff. Andrew?

Andrew Lyzenga: Sure. This is Andrew Lyzenga, I'm a senior project manager here at NQF. I've worked on a number of safety projects before. I know a number of you (inaudible).

Jesse Pines: OK, Laura.

Laura Ibragimova: Hi, everyone. My name is Laura Ibragimova and I'm a project analyst here at NQF.

Jesse Pines: Great. Suzanne.

Suzanne Theberge: Hi everyone, it's Suzanne Theberge. I'm a project manager here at NQF.

Jesse Pines: Great. So what we're going to be doing next is going to through and what I'd like to do is get a list of folks calling in here. Just give us your name and just a one-liner, a little bit about your background so let's start with Jason.

Suzanne Theberge: Actually, Jesse, can I just interrupt for a moment? We are getting some messages that the line is not working for some folks.

Operator, we are being told that they can hear voices through the computer streaming but not the dial-in. They're only hearing music on the phone. Is everybody connected properly?

Patricia Quigley: Pat Quigley is.

Victoria Rich: Victoria is.

Jason Adelman: Jason Adelman is.

Stephen Lawless: Steve Lawless is.

Operator: Everyone that has dialed in is joined to the conference.

Suzanne Theberge: OK. We have Judy Burleson saying she is not connected to the phone line. Is she in?

Operator: No, she's not in the conference.

Suzanne Theberge: OK, I'll chat up with her and I guess we continue with the roll call. Sorry about that.

Jesse Pines: OK, great. So Jason, go ahead.

Jason Adelman: Hi, I'm Jason Adelman. I'm a physician in internal medicine and I'm the Patient Safety Officer at Montefiore Medical Center which is a big academic medical center in the Bronx, New York.

Jesse Pines: Great. Kim. Is Kim Applegate on the call? OK, Melissa Danforth.

Melissa Danforth: Hi, I'm Missy Danforth. I'm the senior director for hospital ratings at Leapfrog Group.

Jesse Pines: Great. Lillee.

Lillee Gelinas: Hi, good afternoon everyone. I'm Lillee Gelinas. I'm system vice president and chief nursing officer for CHRISTUS Health which is the largest international healthcare system and I co-chair the NQF Nursing-Sensitive Measures Committee with Dr. (Marinell) a few years ago and I've also served on other NQF workgroups.

Jesse Pines: Great. Thank you. Pat Quigley?

Patricia Quigley: Thank you so much. I'm a nurse researcher, a nurse practitioner at the Tampa VA Medical Center, the Veterans Affairs and I'm on this committee with (inaudible) from the American Nurses Association.

Jesse Pines: Great. Victoria Rich.

Victoria Rich: Hi, I'm Victoria Rich and I just (inaudible) recent retiree as a chief nurse for the University of Pennsylvania Health System. I'm currently an associate professor at the School of Nursing at Penn in nursing administration and I'm also a patient safety consultant throughout the world and I was also sponsored by the ANA so I'm happy to be on the committee. This is my first stance on the committee.

Jesse Pines: Great. Michelle Schreiber.

Michelle Schreiber: Thank you, I am the chief quality officer of the Henry Ford Health System which is a multi-hospital system in Southeast Michigan and I'm an internal medicine physician by background.

Jesse Pines: Great. Lisa McGiffert.

Lisa McGiffert: Hi, I'm Lisa McGiffert. I am the director of the Consumers Union Safe Patient Project. We are a project of consumer reports and we're the ones that brought you hospital infection reporting when we started about 10 years ago with a campaign to get reporting laws passed in the state.

I work on many other patient safety issues besides infections including medical errors, safety of medical devices and physician accountability and I serve on the MAP Coordinating Council and I have served on various other NQF committees in the past.

Jesse Pines: Great. Gregg Meyer, has Gregg joined the call?

OK. Susan Moffatt-Bruce. Is Susan on the call? OK and I think Kim Applegate has joined. Kim, can you introduce yourself?

Kimberly Applegate: Yes, sorry about that. I was listening to you but not able to speak to you. Late so just dialed in.

Happy to join you. I am a pediatric radiologist at Emory Healthcare, the radiation safety officer for the health system in the university and a quality officer for the department.

Jesse Pines: So far we're still missing Gregg and Susan. Hopefully, they'll be able to join us today. We have our …

Kimberly Applegate: It looks like Gregg signed in too. He signed in on the web.

Jesse Pines: Yes. OK, I see his name there. Maybe he's …

Gregg Meyer: Yes, I was having trouble. This is Gregg, I was having some trouble getting the audio part to work.

Jesse Pines: Can you do a quick intro, Gregg?

Gregg Meyer: Yes, so I'm Gregg Meyer, chief clinical officer at Partners HealthCare and came back to Partners after a couple of years up in Dartmouth-Hitchcock. My work with the NQF on patient safety began at the very beginning with the founding of NQF and I've been on both the Safe Practices and the Serious Reportable Events committees since their inception and I've chaired both at various times.

Jesse Pines: Great and is Susan Moffatt-Bruce on?

OK, so let's go ahead and then move to our developers. Do we have representatives from the AHRQ that are on?

Stephen Lawless: His, this is Steve Lawless, I'm on also.

Jesse Pines: Hey Steve. Steve, can you do a quick intro?

Stephen Lawless: Yes, I'm Steve Lawless, the vice president for quality and safety at Nemours, a large multihospital, multispecialty pediatric group and I've been working on various committees with NQF (inaudible).

Jesse Pines: Great and are there any other committee members who have not yet introduced themselves?

Leslie Schultz: Hey, this is Leslie Schultz. I'm actually on the line today to listen.

Jesse Pines: Great, thanks Leslie. Can you do a quick intro?

Leslie Schultz: Yes, clinical consultant with the Premier Safety Institute and has served on other NQF committees, a nurse and a health services research by background.

Jesse Pines: Thank you. Any other committee member?

OK, so let's move over and do intros for our developers. Are there developers from AHRQ who are on the call?

Pam Owens: Yes, this is Pam Owens from AHRQ. I'm a scientific lead of the Air Quality Indicators and I believe Patrick Romano if he has not already joined will be joining.

Patrick Romano: And I am here, can you hear me?

Jesse Pines: Yes, we can.

Patrick Romano: Yes, this is Patrick Romano from UC Davis and (inaudible) pediatrician and part of the RQI support team. Thank you.

Jesse Pines: Great. Any other representatives from AHRQ? OK, how about from ACR, American College of Radiology.

Judy Burleson: This is Judy Burleson, (Inaudible) and Colleen Tallant. Dr. David Seidenwurm may be calling in shortly as well.

Jesse Pines: And any representatives from UCSF? OK.

Suzanne Theberge: I see Rebecca is on the web though even though she isn't dialed in.

Jesse Pines: OK. And is there anyone who is dialed into the call who has not yet introduced himself?

OK, so let's go ahead and move award to our agenda for today. So …

Richard Moore: Sorry, this is Richard Moore and I'm a medical physicist at Mayo Clinic and I chair the ACR Dose Index Registry Committee.

Jesse Pines: OK, great. Thanks. So what we're going to be doing today, we've got a lot to do between now and 4:00. We're going to be – we've got six measures to discuss and we're going to be going one by one. We've got about a little less than 100 minutes to get through the six measures so, and we don't want to cut off any of the discussion but we are going to give you just a reminder at 15 minutes in.

In terms of the measure discussions our lead discussants are going to do a short presentation followed by a group discussion on any major issues that

need to be brought up and then for three of the measures we actually have public comments that have come in that we'd like the group to take a look at and we're going to just be going down the list measure by measure. We do have time at the end for member and public comment and then we've got about five minutes left at the end just for next steps.

So let me just go ahead and ask if there are any questions at this time?

Gregg Meyer: This is Gregg. Just one quick comment. I opened up the Excel spreadsheet and I did provide some answers (inaudible) but it does not look like my input is reflected in the spreadsheet yet but that's OK I'll be able to cover my measure without it.

Lisa McGiffert: This is Lisa. The spreadsheet came in from whom?

Andrew Lyzenga: It would have been from patientsafety@qualityforum.org, I believe.

Suzanne Theberge: Yes, just sent a link, an updated link out about 20 minutes before the call.

Lillee Gelinas: This is Lillee Gelinas, Dr. Pines, and good to meet you virtually. I want to disclose that Stephen Lawless, Susan Moffatt-Bruce and I, although we are lead discussants, have not collaborated or had conversations around how we would key up the measure and so, I would greatly appreciate just the sense of what the expectation is for the lead discussants and how we might most effectively use our time.

Jesse Pines: Sure. So basically, what the lead discussant typically does is describe the measure and just an overview of the measure and then go through the specific criteria one by one with not too much detail but a basic summary and any issues that you'd want to bring up or you think that would be useful for group discussion. That typically lasts five to eight minutes or so depending upon how much detail folks want to get into or how many issues arise and then we launch to the group discussion after that.

Andrew Lyzenga: And just to remind you, we're not doing formal evaluations at this point. We won't be giving the measures any ratings against the criteria at this time.

Just going to be having some preliminary discussion around each of the criteria on the measures right now, maybe bringing up any issues that you think warrant discussion from the workgroup and maybe bringing to the larger steering committee questions you have for the developers, clarifications on the measure, just anything you think is worthy of a bit of discussion for pointing out as we do our review and again, anything you think you would like to ask of the developers or bring to the full committee.

Again, basically, just a bit of preliminary discussion at this point. We're not doing our formal evaluations but in terms of the lead discussants, maybe just give a quick overview as Jesse said of the measures basically what it's trying to do and a quick snapshot of the numerator and denominator and then as he said walk through and just give your basic impression of how it stacks up against the criteria, each of the criteria, you know, just at a general level, not getting into too much detail.

Does that make sense?

Lillee Gelinas: Yes, the committee guidebook on page 24 has a lead discussant list and it appears to be a little bit more required than what you just described. But that's fine, I don't want to hold us up.

(Crosstalk)

Andrew Lyzenga: And we may get a little bit more into that kind of detail at the in-person meeting. This is just sort of our preliminary discussion right now.

(Crosstalk)

Jesse Pines: So for the lead discussants, just trying to be high level and just trying to keep in mind that we don't have too much time to discuss each measure today because we're going through six measures so any other questions or comments before we get started?

Female: Well, I don't know if it's possible in the future and having served on previous NQF committees that for those measures that are fully vetted and already in legislation through value-based purchasing which means they have been

through a number of filters and forums, it would have been really beneficial if they've had that evidence, the reliability testing and validity testing so we weren't reinventing the wheel.

I don't know if that's possible in the future but I have found that background works to be very helpful in reconsidering measures.

Andrew Lyzenga: Yes, in terms of the maintenance measures, those that are up for endorsement maintenance, re-endorsement, we would expect the developers to come with some if they have been in use in the federal programs which is value based purchasing we would hope that they would bring results testing that they've done during the course of that use for reliability or validity or just to give us a sense of what the performance rates on the measure are, you know, stratified by different sort of populations and so on.

And so that is something that we would like to see in the submission forms so you can, you know, welcome to ask the developers about that and, you know, ask what kind of testing results they've had in the implementation of the measures so far.

Any other questions or comments before we get started?

All right, so the first measure we'll be discussing is 510, Exposure Time Reported for Procedures Using Fluoroscopy. I just wanted to also, yes, as Suzanne and Jesse mentioned, we have – do have some – if you were able to find that e-mail and see the spreadsheet of comments, survey responses from our workgroup members, thank you for filling those out to those who did and if we can – we don't have to base our discussion around those but just know that those are there for your reference.

So again, I think we have as lead discussants on this measure Kimberly Applegate and Michelle Schreiber. You guys can, you know, whoever wants to give a quick overview with the measure, feel free to do that, Kimberly or Michelle.

Kimberly Applegate: Michelle volunteered.

Michelle Schreiber: That's right, I volunteered with Kimberly yesterday and she's going to take over the rest of the discussion since she's a radiologist and clearly has other thoughts. So Measure 510 is part of three radiation safety proposals that actually we're going to be looking at today. This particular measure seeks support for documentation of radiation exposure or the exposure time for all procedures using fluoroscopy. The measure is currently at PQRS measure with sort of limited but certainly improving compliance and of note, it's also a measure that's required outside of the United States.

So I'm going to cover really one, is it important, is there a gap, disparities and priority with a few comments and I was able to see the comments of the rest of the committee but if you sent a link for external comments, I wasn't able to see that one from outside this group. So we know that there's compelling evidence for strong association between radiation and cancer risk and obviously, we're all well aware of the extremely large and growing number of not only all radiology studies but radiology studies using fluoroscopy without benchmark data of those monitoring.

Actually, it was somewhat surprising to me as an intern that this data is not better known and I think it points to an underlying problem in quality that perhaps has not been well addressed. Nonetheless, this measure is just to document the radiation exposure and is the documentation useful as an NQF measure.

I think the argument is that this would improve the ability to benchmark internally and at least start internal conversations about appropriate fluoro times which in theory would lead to setting benchmarks internally and then externally for common studies, standardization of dose information by practice and procedure and participation in the registries so that there would be external comparison data. But those parts aren't included in this study which is just a numerator of recording the fluoroscopy exposure and denominator of all studies.

In reviewing the comments, there were several questions about whether or not just recording the exposure is enough to warrant support as an NQF measure and whether or not, the measure might be better served by including either

recording to a registry, mandating recording or tying this to an outcome measure. It's noted that it's actually hard to find specific outcome measures directly tied to fluoroscopy exposure.

And there was – because of this low confidence that the measure per se would lead to widespread improvement in accountability and change. There was also the comment of in many instances, higher doses or higher times are tied to a higher complexity or difficulty of the case and how would this be accounted for. For example, sites doing more difficult or complex cases might look worse without appropriate differentiation between the difficulty or complexity of the case.

So is there a gap, yes. Will this serve a gap, it will internally. Are there disparities? No, probably not and the priority I think hangs in the balance of whether or not, you believe that this is the first step towards establishing dosing standardization and reporting to a registry in which cases I believe it is worth supporting.

But without that belief and looking just at the face value of the measure itself, some people were asking, is that enough?

I'm going to turn this to Dr. Applegate, again a professor of radiology, for her comments.

Kimberly Applegate: I think Michelle hit all the highlights. This is a complex topic and I want to emphasize that the requirement to report a minimum fluoro time is outside the United States and has been around for a long time, it is the norm in children's hospitals. It is not the norm outside of children's hospitals but we – the National Council for Radiation Protection recommends it and I'm a member of that council and the American College of Radiology recommends it in the guideline.

It is not a requirement. We would just like to reinforce that in another venue. It's a bit of a catch 22 where I believe the ACR is developing a registry to and is piloted from what I understand a registry. I don't – I'm not involved with it. But there is – the goal to have a registry to develop like the CT registry that we will talk about today to set benchmarks for fluoroscopy times. We do

have data and I participated in a couple of projects to establish benchmarks for fluoroscopy times for speech therapists in their guidelines, for common pediatric conditions and now newly last year for myelography and lumbar puncture.

We don't have that same information to tell practitioners what to do for many other examinations but if we have a registry, we would be able to do that and I think that's the goal is to tell people and trainees what they should be shooting for in their day to day practice and then understand when they have outliers to investigate.

So that I think is the power of fluoroscopy times, it's not a perfect measure, it doesn't tell you really (dose) but it does tell you how often how much your using of a device. So I'm in favor of using this given that practitioners today are poorly trained in how use these devices and are not only a radiating patients but they're also radiating the workers in the room.

Andrew Lyzenga: All right, thank you.

Kimberly Applegate: So with that, I'll open it up.

Andrew Lyzenga: So do we have any initial thoughts from the rest of the workgroup? Any particular comments on that, how this measure meets any other criteria?

Gregg Meyer: This is Gregg Meyer. I think that one of the issues, you know, from the start here is this is a process measure, how long the exposure was, the linkage is that to outcomes and just drawing this a lot face validity of this, you know, less has got to be better in keeping track so it's got to make some sense there. But I think on the evidence side, how comfortable are you with to how closely these ties to outcomes so I'm just looking at the criteria that we're supposed to be using.

Kimberly Applegate: So we know that, you know, there is a joint commission sentinel event levels for really high doses for skin burns, so that's a tissue effect and we know that we do track that pretty carefully for the really long, you know, cardiac interventional cases and we really try to avoid those but they do occur and that that's in the document from the 1990 where the FDA put out an alert

and (inaudible) alert level, that's a number of events. But the – but it does happen and we also know that there is a – well, I guess the NIH or the –one of the branches of the NIH is studying what the cataract risk is for workers that are exposed repeatedly to a low-dose radiation and people (inaudible) that is a risk, there is some evidence for that.

So that's not a patient safety issue but it's a worker safety issue. So fluoroscopy time is a very indirect measure of outcome. And if we can add, if we can get feedback to the practitioners, that's the key and in children's hospitals we have that. In other words, I get a monthly record of how I'm doing in my two most common studies that I do relative to my peers.

Jason Adelman: This is Jason Adelman, I have a question as well. So one thing that struck me as a little bit odd about this measure is as far as I can tell is that we're not simultaneously capturing what the procedure is and it seems so much variation if you take, for example, cardiac cath, the difference between putting in one stent, two stents or three stents and the anatomy of the coronaries and the type of lesion might really change the – have a real impact on the duration of fluoroscopy. And without knowing that, how can we really put in context the fluoroscopy numbers?

It's like saying we're tracking surgery duration but we're not tracking if it's a CABG or a simple, you know, tonsillectomy.

Kimberly Applegate: Well, I think we have to know that in order to track – benchmark the type of exam.

(Crosstalk)

Michelle Schreiber: Yes, it's really your – one (inaudible) to put the fluoro time on the actual procedure itself so, you know, maybe the summary of the case itself so you tell all the descriptors of exactly what the case is but it does get to the point of this is kind of internal data, you would have what the cases, you would have whether or not you report it. The metric is whether or not you have actually put in the dose or the exposure time. I was assuming you would know the case because it's actually on the case report itself.

Kimberly Applegate: Right, so for PQRS reporting, it's linked with the radiology report, it has to be linked with it.

Jason Adelman: So shouldn't that requirement be part of the measure, that's my point. But it's sort of assumed that it would be but maybe it should be prescribed.

Kimberly Applegate: It's a really good point.

Michelle Schreiber: I think if one comes to reporting it in a registry, you would absolutely have to have, you know, pretty good descriptor of what the case is because it gets to the point we made before the complexity of the case. In some cases it's going to dictate how long your fluoro time is and you don't want institutions doing more complicated cases to look kind of "worse," that they have longer fluoro times.

Kimberly Applegate: True.

Stephen Lawless: This is Steve. I also have a question about the, is it clear from the measure the way it's written about how they're going to be documented in the chart? Exposure time or time, I mean they would know how much time like my foot is on the pedal. There's a quick read out that they know how much actual on fluoroscopy time versus case time.

Kimberly Applegate: That's automatically recorded in the machine.

Stephen Lawless: OK. And then the other question is I didn't see the driver in this, that you're recording it on the charts. I don't think the intent is absolutely I'm here with it, I think it's good. Is it inherent within this that just by recording it on the sheet of paper that or the report that they're proposing as this will decrease variability and stuff. There's an inherent trust that people will look at like you said you get your report. Is there enough in this measure that that inherent driver is obvious and will be done by people or is it a step towards eventually on registry?

Kimberly Applegate: That's my understanding but let me go back to this. There is now a software that is auto populating just like one of the questions I had for the (FTT) measure, they use that (FTT) measure. I think it's out of date because

now it auto populate the – from the machine to the (inaudible) it's available. Now not everyone has it. You have to buy it but you don't have that people error and also trust issue that people are actually, you know, have to write the correct information in so this information can now be auto transferred.

Michelle Schreiber: But I agree with your point, I think the measure might be strengthened by either mandating reporting to a registry or having some kind of other steps that sort of make one look at the comparison data. The fact of just putting fluoroscopy times on the report, you honestly have to have a belief that it will lead to either better internal conversations or that this is a first step for registry recording.

Lisa McGiffert: This is Lisa McGiffert. It seems also that the usability would be difficult if they were – if every – it appeared that every facility was reporting according to their own standards and so there really isn't any comparing, it's just to say they do it or they don't do it, right?

Michelle Schreiber: That's what the measure is, yes.

Lisa McGiffert: OK. Thank you.

Jesse Pines: Also, just to let the group know. We are at the 15-minute point for this measure for the discussion but you can certainly continue.

Patricia Quigley: This is Pat Quigley and my comments are, I'm sorry, my notes I don't see them in the Excel spreadsheet. I know Suzanne checked on it for me yesterday but my question was related to the feasibility of the measure because in the report, I certainly hear the urgency of it is that out of 5 million Medicare patients, only 1 million of them had this measure reported for them, so I think, you know, you don't have so many that are not being reported, does that also affects the usability of it.

And my other question that may be for the larger group is if this was one of those measures that maybe would be harmonized with the other measure that we were dealing with, with CT exposure because they're both dealing with radiation dose exposure so I just didn't know if that was, you know, part of future discussion. But my concern was how on the report this is and I know a

lot of places don't have an electronic medical records but that was my major concern.

Michelle Schreiber: Well, I thought it was largely under recorded because it's actually an – I don't know if it's an optional measure but …

Kimberly Applegate: It is optional.

Michelle Schreiber: And so there's a lot of people who are just frankly choosing not to. This is, you know, a measure in the step of getting everybody to report. So I don't think it's not feasible because you can't do it, I think that people are just choosing not to do it.

Lisa McGiffert: This is Lisa McGiffert.

(Rebecca Smith-Simon): This is (Rebecca Smith-Simon). I'm actually, I was muted, I apologize. I actually am a measure developer for the CT radiation dose metric from UCSF and that has two parts. One part is recording the radiation dose in the medical record very similar to a fluoroscopy metric that we're discussing. One of the differences is it's more explicitly defined what you put in the medical record and just by slight analogy, California passed a law a little over a year ago that required a similar process where you record the radiation dose metrics for CT in the medical record.

And so as an experience in California it's been an interesting first year, people are reporting very differently and the data are not currently usable to create benchmarks (inaudible) consistently but it certainly raise awareness of these metrics. So certainly in terms of raising awareness for the time to fluoroscopy, it's certainly a first step to raising acknowledgment that you can measure and those measures have some albeit somewhat weak association to the trend of potential detriment.

So at least in raising awareness, I think it's potentially a good measure and I would certainly be happy to work with the ACR to harmonize our piece of art radiation dose for CT in this measure.

Kimberly Applegate: Thank you.

(Crosstalk)

Melissa Danforth: Sorry. This is Missy from the Leapfrog Group. My only quick comment is that, to me these seem like a measure to urge providers to do more and better documentation and I think we've seen other examples whereby making the measure more outcome oriented. It has the added impact of changing the behavior and improving the documentation. So to me, my only concern was that, you know, the – for – you know, providers can potentially do very well on this particular measure. But that to trickle down to the patient actually not receiving too much fluoroscopy, a lot needs to happen in between so all the other doctors need to be able to read the fluoroscopy times. There needs to be a general understanding of what's too much and what's, you know, where there's still room to do more depending on the type of case.

For the patient to actually see the benefit of this work, to me, it seemed like a little bit further down the road. And so, I didn't know if there's an opportunity to have a little bit of a measure that's a little bit closer to – you know, for the patient to understand what this means for me and why this might be important for me that would in turn drive better documentation practices.

An example that's not necessarily radiology related would be early elective deliveries where having a measure looking at rates of early elective deliveries has really done a lot to improve physician and hospital documenting, for example, medical exclusion for early elective deliveries, gestational age, things like that. So, that was my only comment and again, I'm not sure if there is an opportunity in this measure but I thought I would mention it.

Kimberly Applegate: Yes. If we have benchmark data to tell people to be transparent about what should a fluoroscopy time be for X procedure or Y procedure, that would be great, the problem is we don't. I mean, except for the publications that I've put out I'm not aware of, and the speech therapist, I'm not aware of any.

Gregg Meyer: This is Gregg, one to point that I don't see this is appropriate for this call. I think this will be something for us to talk about at the in-patient meeting is or the – sorry about that, the face to face meeting is what – and the issue is I fully agree that this as an awareness intervention, this has a great deal of merit and I

think the question will become and there are several in the set where whether or not that, you know, the value of us in awareness intervention fully justifies endorsement by the NQF. And I think that's a broader issue than this particular measure but I was hoping that we can put that on the agenda for upcoming discussion.

Victoria Rich: This is Victoria Rich. You know, I've been listening to everyone. This is my first stance with this. And as I read at the first three that we talked about with radiology and I always have the patient hat on. And I think it's important for us to have a process metric for us to be aware of it. But it seems too premature to be an NQF metric at this time.

I really like the idea of having the ACR registry where you could really start then to look at what this means and start to on hospitals PI and quality thing to start to realize what you're doing with fluoroscopy but more importantly as a patient and being in a family, if I'm going to be a patient I'd like to know not only would I be exposed to a fluoroscopy but what (inaudible) with CT and on and on. So it almost seems to me that this is great but we probably need to have a registry first to start to really look at what we are doing and to set up standardized doses.

Jesse Pines: Great. And just to keep everyone on track here, we're a little over 20 minutes into the discussion. We still five measures to go though. Are there any other final comments on this measure before we move on?

Lisa McGiffert: This is Lisa McGiffert. I was noting that the review said that there was no measurement time period provided. So, is it collected once a year or?

Kimberly Applegate: It's per procedure, so that's probably why that wasn't put in there. I can tell you that, you know, the PQRS is annual.

Lisa McGiffert: OK, thank you.

Jesse Pines: Great. We also do have one comment that we wanted to take a look at. So, I think it should be on everyone's screen. This one came in from (Kelly Robertson) from Providence Health Services. And I think, is there anyone

who's not able to read it in the screen? OK. So any comments on this comment that the committee needs to address?

All right, hearing none, any – just to check any questions for the developers at this point from the committee? Anything you'd like them to sort of take a look and come back for you at the in-person meeting?

Female: Well, I think the question for the developers is whether or not they wanted to have a Part D sort of like the CT did and enhance this by reporting through a registry. But my other question is having discussed this as the next step that we just discussed this in person and in rank. I'm new to the committee, maybe you can help us what kind of next steps, what follows?

Jesse Pines: So, we will – this discussion will feed in to do a brief summary of your discussion here and we'll provide that to the full steering committee. And we will consider that during our full discussion at the in person meeting. And at that point is where we will do the steering committee votes and actual official rating of the measure against each of the criteria, formally.

We will, I believe, do a preliminary survey before the in person meeting as well to get your sort initial, you know, vote on the measure. Is that true, Suzanne, are we still going to do a preliminary vote?

Suzanne Theberge: No, I think we actually stopped doing that.

Jesse Pines: OK. No problem, I think that's probably for the best anyway. So, the next steps really will be to take another closer look at the measure and be ready to discuss it at the in-person meeting and give your votes on it.

Female: Great, thank you.

Lillee Gelinas: Jesse, just one – this is Lillee. Just one clarification as I look at the measure comment on the screen. It would be helpful around all the measures to understand the data collection burden. And so, that we are informed around that because when the comment talks about the fluoroscopy time as noted in the procedure log, that's one of the big issues that we have in the United States without (inoperable) medical records where we have to manually look at data.

And just in our system alone, there are 50 to 60 RN, FTEs who do nothing but extract data from paper sources because they are not electronic. So, I would really love to know this for data collection burden when we're talking about even considering so many new measures.

Jesse Pines: And, also just to clarify. So, I think that in the feasibility section I think the measure developer does address that. So, it should be some information in the submission form on that.

Well, in the interest of time, let's move on to the next measure. We – and again, we'll have plenty more discussion of this at the in-person meeting. So …

Gregg Meyer: Jesse.

Jesse Pines: Go ahead.

Gregg Meyer: So, this is number 531 which is Patient Safety for Selected Indicators. Gregg Meyer, I'll be happy to lead off the discussion on this one if that's OK.

Jesse Pines: Great, no problem.

Suzanne Theberge: That will be great, Gregg, thanks.

Gregg Meyer: Yes. So, let me start by – with a disclosure. And disclosure is I was part of the original team that developed the Patient Safety Indicators. So, that's just to have that out there. I think that my views on that were generally pretty well known.

Let me just start by saying that what this measure is that it's a composite of indicators. It's actually an average using observed to expected. So, the norm here would be 1.0, that observed – what observed, what was expected across right now eight of the patient safety indicators and those include pressure ulceration, iatrogenic pneumothorax (inaudible), venous capillary and bloodstream infection rate, postoperative hip fracture rate, perioperative pulmonary embolism or DVT rate, postoperative sepsis rate, postoperative wound dehiscence rate and accidental puncture or laceration rate.

In addition to that, in addition to those eight, there are three additional perioperative hemorrhage or hematoma rate, postoperative or physiologic metabolic derangement rate and postoperative respiratory failure rates which right now are coded as zero but they could be added in. The measure as this was previously reviewed by the NQF included just those eight but the additional three could be added into this. So, to kind of run through the framework here, the first in terms of the evidence to support the measure focus on.

This is a composite measure and I use measure in quotes and I'm sorry that you don't have access to my stuff on the Excel spread sheet. But it will be – it is in the system now. But it is a composite measure (in quotes) derived from administrative data-based indicators. And I think that's an important point here. And I think the AHRQ team who put together the measure information that's available on the website is incredibly thorough, has a very rich appendix to it.

What is not discussed there though is the fact that these are really intended to be indicators in that measure. So essentially, what's done here is the composite was constructed to increase the statistical precision because you effectively increase the sample size by lumping together these eight. And in some ways, that will smooth out some of the variation here. What that doesn't do is it does not get rid of the underlying construct which is that these are built off of administrative data derived measure. And so, that's an important caveat as to the evident.

The actual validation of the indicators, there are a number of studies that was cited, Patrick Romano has done much of that work, actually did it during my time in (inaudible).

The true measures though they would require clinical data validation. And so, the composite has not been validated in that way. The second issue is there, you know, is there a performance gap and there clearly is a gap among the providers. And I think that, you know, this does create some spread around that 1.0 and in the materials provided by AHRQ they gave a review of around 5,500 or so hospitals involving about population risk of around 18 million.

And they can show some variation there around performance ranging from, you know, the 5th percentile of 0.25 to the 95th percentile of 2.1. Again, this is normalized at 1.0.

I think the more nuanced where I think the much more important question around the performance gap isn't whether or not it can be demonstrated statically but it's what does it really need because this could be a gap in true performance but it equally could be a gap in coding and capture. And because these are again are derived from administrative data and all the validation was done using the age gap database (inaudible).

The question whether or not this is a high priory, there are a number of high risk outcomes here included. I think some of you commented that you didn't think they were all here. That you wished some others were here and I think when you get to the – we went from eight to 11, it would perhaps increase that and there is no doubt that these are important things that this, you know, nationally ought to be focused on.

I think the only question is whether or not this is the right way to do it. You know, the next question is the composite performance, the quality construct of the composite and it is very rigorous – their specifications are very rigorous. I mean I think in many ways the rigor of the analysis here is in some ways could mask the flaws of then underlying administrative data.

And so, in some ways, it could create this false sense of precision but there is a lot of rigor to the measure in here. The question is about specification, these relative specifications for the indicators are very well defined and have been used over time. In terms of reliability test, these individual indicators happen but again, I would point to the report from Patrick Romano on that.

The reliability of the resulting composite would need to be studied further. In terms of validity testing, clinical validity composite really hasn't been rigorously tested, although again, I think many of you have said it's got a good face validity. These ought to be captured something important. The threats to validity is really the question whether improving performance in the composite has the same clinical meaning is focusing on individual

components. If you could, for example, spend your time focusing on pressure ulcers, that may have a relatively small contribution to the composite and maybe very, very important in terms of (inaudible).

The composite performs as it's specified and that means it does create a broad picture of performance to create spread just by statistical methodology. One of the things that it raises is the weighting among the indicators because, you know, there is some judgment there as to which of these are more important than others and would be difficult for me to tease that out.

Feasibility, I think this is both the blessing and the burden of the PSI and its composite that's derived from them, is that they're relatively easy to capture from a measurement data. That these don't involve extra cost generally for being able to collect the data. They do generate a fair amount of cost in terms of the issue I talked about earlier which is focusing on coding and capture rather than improving performances.

And finally, in terms of usability (inaudible), I would say that these are not as actionable as other metric since the data comes from administrative sources with a fair amount of lag through them. It's easy to use for accountability purposes and in many ways, you know, when you – the zone of preamble here was that it was – the methodology was to be applied at the national, regional, state provider level. The national, regional and state level for accountability purposes is probably not quite as important that deal with the time lagging effect in retrospective.

The difficulty as I said, using them in real time is much more of a challenge. And again, I would point out and my disdain for my own work here shines through that the patient safety indicators were originally designed to be hypothesis generators. They were not designed to be this positive or to be considered a true measure of performance. They're actually to point the way towards digging into the data deeper. I'm not convinced that creating a composite of them that improves the fiscal precision but not really the clinical validity has a lot of merit but I, as always, I have a strong feelings on the subject and I'll stop there.

Jesse Pines: Great thank you, Gregg. Any other comments from either our lead discussants or any other workgroup member?

Jason Adelman: This is Jason Adelman, so there are three measures that we'll be discussing more on DC that I guess most people or many people on the call realized are part of this HAC provision in the Affordable Care Act were 1 percent of Medicare reimbursement will be tied to performance on these three measures. One of the three measures is this AHRQ PSI 90 so as Dr. Meyer just suggested when he was at AHRQ and they originally developed this metric it was meant to be hypothesis generating and now it's hospitals across the country are going to be held accountable to the performance.

And that being said I was on this committee once before and I remember discussion on the committee about, you know, should we consider how these measures will be used when reviewing metrics in general. Although there was never an example like this where in the Affordable Care Act coming in 2015 we know that I think there's something like a billion dollars tied to three metrics, the Quality, NHSN, CLABSI and this one. This one represents 35 percent of that program, so hundreds of millions of dollars.

So with all that being said I think validity is incredibly important here. Like what validity testing, the developers, AHRQ, Dr. Romano, did for each of the individual PSIs that underlie the composite. And so, you know, it's a tremendous amount of work to go through given that there is some, I don't know, roughly 20 or so measures, give or take, to look at the validity of each one of them yet I think it should be done. So I didn't and also there's one thing that Dr. Meyer said about the face validity of this metric.

I personally feel that the face validity is not apparent meaning like claims based measures or fraught with problems and so I know Dr. Romano and AHRQ have bent over backwards to try to even change ICD-9 codes to make these work. But it's just too hard to go through just the application itself and understand all the validity testing that were done at each of the metric yet I think that we should try to do that and I personally will try to do that and bring it to the committee when we meet in DC. And so with that being said I would love AHRQ, Dr. Romano, whomever can help by providing some of the

actual articles, you know, the peer reviewed that describe the methodologies, research that shows the positive predictive values and sensitivity (inaudible) or whatever it is that shows the performance of each of this metrics.

Perhaps they can be sent to NQF and shared on the SharePoint drive and, you know, it can save a lot of time if the articles are there then for me and others to go and dig for it. But I think we have to be very careful specifically on the validity of the underlying metrics that make up this composite.

Gregg Meyer: This is Gregg. I don't disagree with that at all. I do think that the NQF there should be able to pull up the report that Patrick did looking at the underlying indicators. I do think that, you know, my take on that review was that they showed that these were problematic.

Melissa Danforth: So this is Missy Danforth from Leapfrog. I'm just sort of thinking back to a few minutes ago when one of the committee members urged this and I'm in complete agreement that we want the information, we want to be sensible about the burden of collecting these measures and, you know, certainly as an organization that collects data from 1,400 hospitals, I'm extremely sensitive to the burden of data collection.

So I always find it difficult in these conversations about, you know, can we talk about measures that use administrative data because they're so easy for hospitals to participate in those reporting program and I think they do give us really important and good information. I mean the numerators for these are actually incidences of these things happening, the numerator for these measures.

So I just want to mention that that I think, you know, we need to be thinking both ways, both of the burden and if we don't want all this extra burden but we still want the information, what's a way to accomplish that. I mean (inaudible) I think really a good job providing us with hospital information about how they accomplish getting this information and the way that keeps the burden fairy low for hospitals.

If you think about the amount of information we're getting just from this one composite, the amount of time it probably takes to calculate these measures, I think that, you know, the payoff is substantial.

The only thing I would add is I think that this I want to earlier point because this measure is now going to be used in that HAC reductions program. We use this measure in one of our programs throughout Leapfrog. I know hospitals do have some questions about exactly how the measure is calculated because it's taken off their claims and they're now active in the data collection and the algorithm is completed and it is a composite. I think that there could be some additional education to hospitals about what the measure is capturing and how those are calculated that could lead to the composite core to be more something that hospitals can act on and use as an improvement tool. I think hospitals in ability to (inaudible) for an improvement tool is they don't understand how the measure is being calculated themselves. So I think there's part of an education that could happen, it's not necessarily the (inaudible) measure that the hospitals aren't doing more with it.

Gregg Meyer: This is Gregg. Just one point of correction in that statement and that is these are not – the numerators, these are not actually instances of these things happening, you know, just trying to be true to the folks who originally worked on these metrics. They are – the numerators are – basically they are codes. Those codes sometimes do reflect things that actually happen but there are many the coding is not perfect. Because of the imperfection of coding they don't always reflect and so we did some work looking at one of them in particular that has to do with birth trauma which was not included here. But it did show that in fact that statement that incidence of the numerator actually are things that actually happened. It's not true.

Lisa McGiffert: This is Lisa McGiffert, I'd like to ask a question following that. I noticed that there were quite a few exclusions, code exclusions in this measure and so I'm assuming that those were – those exclusions are an attempt to weed out those false positives or the false – the things that aren't – didn't really happen. And I'm wondering, you know, if it doesn't indicate all the time that happened, does it – you know, how is it? Is it rarely indicating this when it happened or is it almost always indicating that something happened?

And rarely – there's a rare outlier where it didn't happen because that's a really important piece of information. If most of the measures in the numerator, most of the codes indicate that these events happened then I'd like to know what the statistics are on that.

Gregg Meyer: I think actually the – I think some of the reviews can probably do that. It can give us some ideas for the individual indicators. Most of the time there are. These are pretty good. They're just not – they weren't designed for accountability purposes and for payment purposes. More often than not, they're going to be reasonable but they're not perfect.

Jesse Pines: We are at 15 minutes.

Lillee Gelinas: Hi Gregg, this is Lillee and you can probably just help me with this small piece that the administrative data and (inaudible) billing and coding data. And when we are moving to ICD-10 by mandate of October 1st I am assuming although I don't know that ICD-10 will provide much more coding clarity. When we have our (face to face) meeting it would be helpful with that with all the efforts that's going on right now with the provider to move into ICD-10. But once (inaudible) and just affirm this for me, administrative data are unlikely to capture pieces of complication, right, regardless of if they were preventable or not. It is just, you know.

Gregg Meyer: To tell a quick response say no, they want to move on. I think, you know, what will be the impact of ICD-10, it has to make it better. I don't know how much it will make it better but it has to because of the, you know, the increasing number of codes there and the greater specificity to them.

You know in terms of the second point, you know, the – in the end I think that these are – we'll have to have a broader discussion about this administrative data and I probably again will cut across several of these metrics and we probably should leave that for the face to face meeting.

Lillee Gelinas: There's measurement and there's (quality) improvement, two different things.

Gregg Meyer: Yes. Yes, and again these were meant to be hypothesis generators that you would then look deeper on the clinical data and other collaborating evidence and improve from it. As a performance improvement tool I think they still can be quite useful.

Lilllee Gelinas: Thank you.

Jesse Pines: Just again in the interest of time maybe we could move on to the next one and I suspect that there will be many similar issues. The next measure is the pediatric composite and may have the very similar discussion here. Do we have any lead discussants who would like to give a quick overview?

Stephen Lawless: This is Steve Lawless, I can take it if you want.

Andrew Lyzenga: Great, thanks Steve.

Stephen Lawless: (inaudible) and everything was just said about the development of the measures, mark, but (inaudible) the word pediatric versus adult and changes from eight measures to six measures. This is very similar by our need of six measures. Actually, on a puncture, a laceration, pressure ulcers, postoperative sepsis, pneumothorax and wound dehiscence rate or wound (inaudible) infection.

So this is a weighted average of six indicators. Methodology looks like exactly the same as with the prior discussion was on the adult that's been out there and well-known in the pediatric circles and we've been getting these reports – most hospital have been getting – children's hospitals have been getting these reports for a few years with the same – I believe it's probably the same intent out there.

This is a subset of a larger group of the (inaudible) quality measures that have been collected by AHRQ and I think a few of them were eliminated because of volume variation to the site. So with that as a backdrop, does anybody have any questions before we go into my take on the measure itself? I have some critique that may be just a little bit different than the prior discussion.

Female: Steve, this is (inaudible). The one thing again that goes to what I asked at the top of the call because so much good work has already occurred and I can't see us as a committee recreating that work but the TDI composite measure workgroup final report, that would be really helpful in teasing out what some of the considerations were and that's one thing I didn't have is that within the materials, I apologize, but I didn't have that.

Stephen Lawless: So a lot of these is that there was lot of reference materials in the report itself that we're not – they say refer to this website, refer to this website. When you have an 85-page document and a lot of it is refer to this website for other details, that's making it a little bit hard to look and actually you just said that. You want to see a little bit more the critique here and the evidence of that.

Maybe for the developers that they have for the bigger meeting just kind of more of a composite of that rather than just the website link. That would be you know (inaudible). I knew that's an indicator because I've seen the reports from the indicator over the years and there's variation there.

I think that there's one subtlety in it in that I find it interesting and that is – my impression with pediatric hospitals rather than adult hospitals, a lot more of the pediatric hospitals see the sicker kids in community hospitals and so community pediatric based hospitals and so this is one where the risk adjustment a little bit in terms of the severity of illness. Maybe I would like the developers in the meeting to kind of address that piece.

You can conceivably see that community hospitals that don't see a lot of children with complex surgical condition. They're doing ASA 1s and 2s. Less likely are they going to have postoperative sepsis versus a lot of academic centers or pediatric hospitals which are going to see the more complicated kids, they may have septic kid or the bad appendicitis or not. So I think just addressing of how you deal with what maybe a bias against actually larger children hospitals who get referred these cases versus community hospitals who see some of these kids.

That was just one concern I have there. The other and I think the reviewers talked a little bit about they're modeling and smoothing in that. I think maybe

a better discussion about what that smoothing does to eliminate that potential discrepancy. And then the other thing I just had a comment with this, again, is about the coding but this is more present on admission coding. The developers mentioned present on admission but they don't mention when in terms of the references that they took that into consideration or not.

I think it's just a little bit more clarification of how they exclude present on admission diagnosis or make sure that (inaudible) aren't there. Again, this is a little bit different than just administrative coding would be beneficial. That's all I have.

Jesse Pines: Great. Thanks Steve. Any other comments from the workgroup?

Lisa McGiffert: This is Lisa. Can the developer speak to the issue of how present on admission plays into this measure? It seems like it is a component of all the administrative data measures. Any …

Jesse Pines: Pam or Patrick, are you still on?

Pam Owens: Patrick, I'll let you answer first.

Patrick Romano: Yes, we can speak to that. The present on admission information is an intrinsic part of the specification of each measure in both its composite as well as the PSI composite. I think what the reviewer may be referring to is in the reference population or the benchmark population that is used for calibration of the risk adjustment model, that some of the state that are submitting data into that reference database don't have POA information, at least for all of their records.

So there are – over time there's been an increasing percentage of cases in the reference data set that have POA information. But it is an intrinsic part of the specification of the measure. So certainly, for accountability applications it would be expected that it would be used.

Male: All right. To that to you (inaudible) from page six of the document where that reference that. So I think that, again that clarification how to correct going forward would be beneficial.

Pam Owens: Right. And just to give you some information, in 2011 data, 36 dates are preventing complete POA data. And we just recently completed that study under the (HCAC) project. So that more – you can see more and more in an all payer database are actually providing POA.

Male: All right, so you would maybe consider it or maybe discuss whether it should be part of the specifications or not?

Patrick Romano: Well as I said it is part of the specifications for each and every indicator.

Male: OK. OK, thank you.

Jesse Pines: All right, thanks. Any other comments on this measure? We do have one public comment I believe on this measure and we'll pull that up on the screen for you again.

Female: This is …

Stephen Lawless: One other – This is Steve. One other comment just in terms of the scoring of it and I just – and the developers maybe think about this. It's very – you're all presented as individuals, I believe, right, where you have the composite score and you can bring it down. The composite score is the presentation or is the measure also the breakdown of each six of how you perform this six or is it one global score?

Jesse Pines: Pam or Pat, can you address that?

Pam Owens: So the composite measure is one global score.

Stephen Lawless: One global score. Got you, OK.

Patrick Romano: So it will be a normalized global score.

Stephen Lawless: Got you.

Lisa McGiffert: And this is Lisa McGiffert. I have one other comment. I just think that it's pretty – I think it's very important for consumers to – for us to start doing

these composite scores and just because people do want to see a broader assessment of how (inaudible) the hospital is or, you know, rather than individual measures and trying to piece those together.

And I think that composite measures are going to be really important in the future and I think these two are pretty important now.

Male: Yes, and of course the questions would be is with, you know, the composite measure of six, what you have to consider is variability in the six measures, I mean you can score perfect in one and really be bad in another and the two of them to get your average.

So is there a way of looking at variability even within your variability, your performance within the six as a subscore or some score as you can say, "Right, you're really consistently low or consistently high" versus "Don't go there for a pneumothorax." You get your pneumothorax there but you're not going to get the pressure ulcer.

Female: So you're asking a really good question. And part of the reason that people have argued to have within hospital, compare for instance that not only are you reporting with composite, you know, the actual composite score, but you're actually also reporting on each of the individual indicators.

Female: Yes, and I think that's a pretty universal position to try merge for that breakdown as well as the composite.

Melissa Danforth: Great. This is Missy. I mean I'm not as close – I'm not familiar as much with the use of the pediatric, but I know for example that sort of the measure itself and the opportunities to present individual breaks within the composite plus the rolled up composite and lost of other details. And then there's the way that a national organization like CMS might implement the measure on its own.

So I think that's the distinction, I think, you know, this is the ability to do to, you know, to show different – to show the rates by individual PSI and then to show the rolled up version and then to provide some other information to hospital to help them understand how it was weighted and how one might have influenced the other. And where, you know, you might not have had

enough cases to even get a score on that particular PSI and then where did the weight go.

I think some of the frustration I hear from hospitals is it's not the measure itself, it's the way that it's being implemented by CMS and what information they do or don't make available to hospitals. And I think those things are distinct. So I just wanted to say that. The other thing as I, you know, Steve, I think you brought up a really interesting point about how the measure will play out or how the composite plays out in say a community hospital that treats pediatric patients versus a pediatric hospital.

And I think that, you know, that's one of the notes I had as well, just that there seem to be some performance differences and there was, you know, a small pediatric population versus a large pediatric population and could that be the difference between like a community hospital and a pediatric center.

From a consumer – so and I put that note just out to the – in my comments, but from a purely like a consumer or purchaser point of view is, people expect whether they should or not, exceptional specialized care when they go to a pediatric facility. Like the idea is that if your child is really sick, you would travel to the end of the earth to find the best hospital to care for them.

And I think some pediatric hospitals do market themselves that way to consumers. So there has to be some way to balance the safety outcome the patient can expect when they go to one of these sought after pediatric facilities. In consideration of, you know, some of the points you made like our patients are sicker and this and that, there still has to be some expectation of not having these patient safety errors happen, even though your kid is every sick, because you specifically sought out this facility for that.

Kind of a side comment, but I just wanted to add it to the discussion.

Jesse Pines: OK. Thank you.

Male: Just a note have been …

Jesse Pines: Go ahead.

Patrick Romano: I was going to say this is Dr. Romano. I just wanted to make one other comment just in general response. I think that for these composite measures, the clinical validity of the composite does rest on the clinical validity of the component, because there's really no way to evaluate the clinical validity of the composite as a single measure. It really rests on the component. So I think that we do hear the message about really conveying that information and certainly all of the peer reviewed papers and reports that provide that background information about the individual indicators can be offered to the committee.

And also I would just point out that most of the components of both of these composites have been individually endorsed by the NQF. In some cases the individual measures are on a separate timeline for endorsements, so some of them have been endorsed more recently than the composite measure and NQF staff, you know, would be able to clarify those issues.

But there's information from those individual endorsement processes that can be brought into the discussion as well.

Jason Adelman: This is Jason Adelman. I appreciate what Dr. Romano said and the fact that many of this metrics have already been endorsed. But I am interested in does he offered to see the underlying studies and evidence. And so perhaps at the end of the call we can discuss with NQF leaders how to best go about Dr. Romano's sharing those studies.

Jesse Pines: OK. Yes, we can discuss that.

Great. Just again to note the public comment that did come in on this measure, this was from the Armstrong Institute for Patient Safety at Johns Hopkins. Sort of echoing some of the points we've discussed already, saying that they do not support the measure that the indicators are – the true preventability of this PDIs is relatively low and suggesting that the indicators are not useful for a public hospital comparison on that basis.

So just to note that comment. Any reactions or questions or comments on that public comment?

All right hearing none, let's go ahead and move on to the next measure. The next one we have is again related to radiation safety. We've already heard from (Rebecca Smith-Simon) a little earlier on this call from UCSF. Let's see. We have Kimberly Applegate and Victoria Rich on this one. Would one of you like to introduce the measure? (inaudible).

Kimberly Applegate: We didn't actually coordinate ahead time. I don't know if Victoria, you want to introduce it.

I try to answer the questions when there were comments from the staff that was very helpful. In this case there weren't and since this is the first time I've been on this, I'm happy to take a crack at it, Kimberly. Go back to be – I'm plopping to the measure here. This is the two-part measure. And it looks to me and I and please correct me if I'm wrong on this, I'm glad (Rebecca) is on the call. This is a two-part measure, part A is an outcome measure, part B as the process measure. Part A is for to be amount of radiation doses associated with CT examinations up the head, the neck, the chest, the abdomen and pelvis combined, and the lumbar spine in both adults and children.

They have different numbers, for adult it's 100 per body part and 50 for children because there is less CT performed in children. I wondered if that's really an outcome measure. It seems more like a process measure. But we can come back to that. Part B is the proportion of CT exam where the measure of doses included in the final medical report and I assume by that they mean the radiology report.

And the measure of dose is (inaudible) and this is where I have a little bit of a worry is the measures of the CT output from the machine, the (inaudible) or the (DLC) or the effective dose. And I think that there is no doubt that there is – if you look at the rational report I would put this as high priority. In terms of the evidence I didn't see a complete literature review. But I think that there is a lot of literature about the increasing use of CT.

If you look at my comments in the Excel spreadsheet there are over 300 million CTs performed and over a quarter of them are performed in the US. We're one of the highest users of CT. And it's not the dose per CT that people

are worried about, it's the number of CT that are performed in the population and the growth rate around the world is the same, it turns out. It has leveled off somewhat because of public concern and copays.

As (Rebecca) mentioned, California is the only state that passed a law requiring CT dose record, inpatient and outpatient, after two variants in the cases that has CTs, one in a child in Northern California and the other at one institution in Southern California that caused some very high doses in potential stroke patient and the problem is that CT has been associated with potential doses of cancer link but the data are mixed and I want to emphasis that. I know that maybe (Rebecca) may feel differently about this.

But if you look carefully at the data through the National Council for Radiation Protection, a data that there has been a review of the couple of papers that made a lot of press in the British Medical Journal and Lancet that she had referenced and there is the risk-causation in those papers. There is bias in those papers and we don't know why the children had the CTs that they got which had shown a small increased risk of brain cancer and leukemia.

So it's not really clear that we can link directly CT and later cancers. But we certainly are concerned that there is a link and so we very much are interested in monitoring those and reducing those. And there has been evidence that with working with vendors that we have lower dose with a number of papers. One of the things that I noticed for the documents and testing updated and the papers of letters of support, one letter of support that I found from (Jerry Bush Burg) is from 2010, that's four year ago.

So that the information is not quite up to date and that the organizations like the (NCRP) have produced some very good document about those reference levels by body part and include adult and children doses for CT and for other radiation studies which have been incorporated in the American College of Radiology Guideline. There is the AAPM, the American Association of Physicists in Medicine which have developed very detailed CT protocol to address optimization which is how best to perform the study to use the appropriate dose for the clinical indication.

And in light of the manufacturers (inaudible) equipment have partnered with the FDA, another medical organization to provide some dose check in the appointment so that the events that happen in California are less likely to have and, you know, never say never. People can always do things that aren't good. But they prevent some dose checks that make it harder for technologist to deposit really high doses in people.

So that some of these things have been addressed and I want to make sure people are aware of that. The other thing I want to say in term of the dose registry, I think it's really important that we do this, that this is something that has been done in other countries, not in the US and I think there's room for both the ACR dose index registry and the UCSF registry.

I will caution that I think it's important not to have an or for choices for people but to have an and if we're going to have (CTTI) volume and the (DLC) dose metric that's supposed to be out for some of the machine. I think it's important for people that use the machine to understand both of those metrics from CT. Whether we have an effective dose or not is irrelevant. I think it's not a measure of quality assurance. It's a measure of risk for patient.

That's fine if people want to have that. But it's not the critical element in understanding those output from the CT machine and to understand how to optimize the CT machine. It can be helpful in looking at patient doses across population. But for the purpose of understanding how to optimize the CT machine equipment and to look at those references level that's not a dose metric that we use.

So I just want to make that point so it shouldn't be an or between the two choices. I think it should be an and. And then I think that it's also important to know that we now have software probably because of the California legislation that makes it automated and not a manual process available to people and that decreases human error and more confidence and compliance with people using this process that the information will be automatically transferred in almost all equipment today. So that it can be – you can have more assurance that the information is accurate.

So I hope that's a – I think there's compelling information that we want to help people understand what they're doing and I think there's a compiling argument that we want to have these registries and I just think that there may be a couple of concerns that I've raised in my comments in the Excel sheet.

The other is the size of the patient. You know, I think it's OK to have categories, I think it's more or less categories for age, it's OK, I think before we have the size-corrective dose metrics. It's just less accurate because patient weight is highly variable, especially in children. I had a 400-pound 4 year old that I took take of at one point so I think that unfortunately age is not always very – a very accurate reflection of the size that you're going to get.

So that we have a way to correct by the girth of the patient which is a document that allows us to measure on the (scout) image of the CT and really get a better understanding of what dose is to the patient based on their size.

So I'll let Victoria comment now.

Suzanne Theberge: Victoria actually sent an e-mail that she was disconnected from the call and is traveling and can't reconnect so she left the call.

Jesse Pines: Other comment from the workgroup on this measure?

Lisa McGiffert: So this is Lisa McGiffert. I want to be sure that I'm clearly understanding because this measure requires – does it require reporting to a registry because I don't …

Kimberly Applegate: I guess that wasn't clear to me.

Lisa McGiffert: Yes, it's unclear.

(Rebecca Smith-Simon): So there's nothing – this is Rebecca Smith and I developed the measure. The measure does not have anything to do with a registry. What the measure does is provide a simple way for facilities to summarize the doses they've used in their population and compare it to other populations.

And it provides sort of a very simple framework for facility to answer the question as, am I doing a good job in keeping dose as well compared to other

facility? So there were two articles in this month's Journal of the American College of Radiology. One of them uses this measure as a way to assess the facility performance over two years and using this measure, you could see the facility dose went to 40 percent over those two years.

It also shows the results for these four metrics that are provided in these measures, CT, (DI), (DLC) effectiveness and also add (SSDE), another metric and shows that all of those measures, I think it's slightly different information but they all move directly in proportion to one another. So if you look at any of them, you'll get a lot of insight although I agree with the concern that maybe you should just report all of them but I think having any of them.

So number one, it has nothing to do with the dose registry. It rather becomes a simple way to know how you're doing. The second paper in the Journal of the American College of Radiology was a randomized trial, the intervention to lower doses at the technologist level and again, the doses were set in the framework of using the NQF measure and the randomized trial was successfully intervened, lowered outlier radiation doses among technologists.

And so the purpose of the measure is simply to provide a tool to lower doses. And the only other point that I'd like to make since I have the floor right now is to say that, there are two concerns of the population's exposure to radion from CT. One is, that there's a lot of CT and therefore, there's of dose but second that the dose was recurrently profoundly variable between institutions. So if you go to one institution and get a CT, the dose might be 20, 30, 40, 50 times higher than if you'd gone to an institution across the street.

And that has to do with the fact that the institutions are not standardized in what they're doing and this is one way to allow facilities to begin to look at their practice. The radiation overdoses that Kimberly pointed out, it wasn't a patient in Southern California. It was over a thousand patients across four US states that had doses for CT that were comparable to radiation brain cancer treatments. They got doses that were two, four, six and 10 (inaudible). These were doses that were indescribably high and there is continued concern that doses could be high. And this one concrete way and the only way currently

for facilities to really get very easy information at how they're doing very quickly, very inexpensively.

Lisa McGiffert: OK, this is Lisa McGiffert, I wanted to follow up because – so it's not standardized but they're able to – and who do they report it to and who provide the comparison? As a consumer, I look at things like what's the consumer going to see? Are we going to see the variation among the facilities? Are we going to be able to – because what I like about this measure is that it said, yes you recorded and here's what the dose is but it's unclear to me how when it's publicly reported are we going to be able to identify which facilities are able to provide the most optimal doses?

(Rebecca Smith-Simon): So the metrics itself is a way to summarize the dose but in terms of setting standards of what's allowable, this measure does not address that. So there are – we have several publications in press that are setting out standards that I believe are good standards. There are currently no national standards like around this issue. I agree with you, those would be nice but I think the public reporting is a first step (inaudible) to have those data to assemble.

(Crosstalk)

(Rebecca Smith-Simon): Yes, there are two national standards. There's the national guideline and then there is the National Council for Radiation Protection Report, you know, that provide both the dose reference level and the achievable level which I refer everyone to at the next document. It's not a static number, set of numbers though because these are based on survey of national performance and having a registry is also a moving target. And you'll see in the – I think in the discussion about it says, those numbers move and hopefully move downward with time just like the European data moved downward within dose reference levels overtime so that, you know, we track them over time and see where we should be moving toward within the collaboration to keep these doses going in the right direction.

So I think it's important to understand that that we do have some national norm and we benchmark to that. Now in 2010, we have less information.

Jesse Pines: All right. Well, thanks (Rebecca). I think maybe we should move on to the next measure at this point unless we have or – well, I do want to bring your attention to one comment that we did get in on this measure. It's a little long so I won't read trough the whole thing but again, just bring your attention to that and ask you to take a look for the steering committee members when you have a chance and give it some thought. There are some suggestions for the developers as well if you'd like to take a look at that comment. I believe we'd provided it.

So let's go ahead and move on to the next measure which is 740, Participation in a Systematic National Dose Index Registry. I think Victoria and Lisa, we had you as the lead discussants. Does one of you want to give a quick intro? We'll have o be pretty brief, a couple of measures here.

And also so we just have, again, we've had about 25 minutes left so let's see if we can limit the discussion the to about 10 minutes here.

Lisa McGiffert: OK. Did somebody say Victoria had to get off the call?

(Crosstalk)

Lisa McGiffert: OK. So this is very similar to a couple of other measures that we've seen and so I'm going to try not to repeat. But this is a, what they call, a structural measure. It is a yes or no response and that is all you get. Yes, they participate in a registry like in the national dose index registry or no, they, do not. And it was previously endorsed in 2011. They had fewer facilities participating at that time, like about 104 and then in December 2013, they have 510 participants. And there was no indication that I could see that there was – that the measure has been changed in any way but there was an analysis that was provided based on the data that's been collected since 2011.

So, since it's been around 2011, it seems like we would have more evidence but it really didn't seem to be very much evidence provided. General summaries and manuscripts that are in various stages of publication and there was this analysis that showed a decrease of dose levels by the facilities that participated in the registry. However, they did not provide any kind of comparison with facilities that did not participate in the registry and the

America College of Radiology indicated that they didn't have any information about those facilities.

So the assumption is that facilities participate in the registry are better than ones that don't because there is more awareness. And we've all heard about those discussions. The goal is to gather enough information to establish a benchmark and to eventually assure that patients receive the lowest possible dose that's appropriate for safety reason.

And there's indicate – the ACR said that they intend to submit future measure to NQF that would more specifically gauge a site's adherence to the benchmarks. And my opinion is ideally these two measures should be submitted together so that we would measure participation in the registry and adherence to the optimal doses of radiation.

It seems to me that there's already significant data in the registry now and I'm wondering, you know, they've got 6.5 million records and whether that might be enough to establish a benchmark that could be adjusted over the years as they get more and more participation. Similar to the questions on the previous measure I think that this measure might be worthy of supporting if it leads to a mandate reporting in the registry but I have some pretty – pretty low confidence that it would lead to widespread improvement without such a mandate.

There – this is an e-measure and so the data, it looks like the burden of collecting the data pretty – pretty low, it's captured mostly electronically. They're currently on a small number of facilities participating. There were 8,000, 8,500 facility performing city in United States and only 658 report to this registry.

So it's not clear whether all of these facilities have the digital equipment available to be able to use this automatically collect the data but something else I read in one of the other measures, seemed to indicate that that was pretty – pretty common that people had – facilities had this digital equipment whatever you call it. It is definitely a high priority like people said and I was also pretty shocked to see that all this document said there were no

benchmarks that existed even though we do have some information. And that that seem to be surprising. And seemed to be the most compiling reasons for indorsing this measure. But I am hesitant just because it seems like an almost, almost a quality measure but not quite.

And then I would really like to see it paired with some kind of outcome measure. The – let's see – my concern with the Yes or No are issues like how do they define participation. If a facility participated with 10 percent of their cases would they be termed as participating? And also that a simple Yes or No measure is not an indicator of quality, you might have a facility reporting really dangerous dose levels. And they would get favorable qualities or regardless of what they were reporting.

So it seems like I had one other thing – oh the feasibility seems good, the registry seems to have a pretty good system for converting and standardizing data from lots of different electronic sources and a broad array of scanners and models. And so they make, make a good case for making it easier on the facilities to report, report these measures. As far as the use – there also is a fee attached to it. So as the facility that participated in, this measure would have to pay a fee to the registry of 500 to 1,000 dollars.

So that would be and that might an issue for participants. Obviously this ACR feels like – and people have talked about this with the other measures that this is a really important use of – it has important use for providers. It gives them a feedback loop. But there really isn't any clear evidence presenting dose levels of nonparticipating facilities. It's higher than those currently participating. One concern that I had with public reporting is that it's – it's unclear as to how this would translate into accountability without some kind of outcome measure attached to it. And I think that's probably enough.

Male: All right, thanks Lisa and we …

Gregg Meyer: This is Gregg, I just have one quick comment or two (inaudible). The first one is, is in some way this strikes you more of a state practice than a metric, I mean, this is something that we recommend folks to do. And I just want to know whether or not it should be considered under that rubric rather than here.

The second point you know, I didn't understand about the cost of the, of the registry participation. That we should have to be careful about – in some ways this kind of feels to meet leverage or forcing people into a registry. And whether or not, whether we endorse this specific registry or not, we had to be very, very careful about that.

Lisa McGiffert: Just fact quickly and the comment that I would like to make very briefly is that you know, their own prior committee that we did look at structure process and outcomes. For example, fall prevention, pressure ulcer prevention, you know, did they have assessment in them and upon admission. Did they have a plan of care? You know those kind of structures that I am not in favor of NQF endorsing participation in the registry. And you know, I could see the exposure risk was 0510 and 0739 being further linked to an outcome. But just to have a registry that's not a quality indicator that's a reporting indicator.

So I just want to say that you know, as we've talked about in relationship to other outcomes but it was more structure that would drive practice. This is just documenting with something done yes or no on a registry.

Gregg Meyer: Thank you.

Female: OK.

Andrew Lyzenga: I think our developer, Judy Burleson wanted to make a comment on this measure. And Judy are you online?

Judy Burleson: Yes, but actually I was asking to make a comment on 0510 the Fluoro measure earlier.

Andrew Lyzenga: Oh I see. You want to get that in now?

Judy Burleson: Well I don't want to derail this conversation. So …

Andrew Lyzenga: OK, maybe we can wait for a moment, Yes OK. Any other comments or questions on measure 740?

Richard Morin: This is Richard Morin, the director of registry of the registry. I would like to clarify just a couple of things. On number one from its inception, it is

completely automated. When a facility signs up and configures their scanner. Every scan they ever do on that – on that scanner is reported to the registry. So we do have a real time shot of what's actually going on in practice. And the reason why you can't compare with data in the registry versus people that do not participate in the registry, if there's no way to build the data base of what their – of what they're – how they're doing things. We have never known what everybody is doing outside of a registry.

So this is actually the first time that we are getting real data about the state of practice factors with regard to those indices. And I think that that's important. This is a major step in allowing facilities to be able to see how they are doing and – and optimize their protocols. We have had reports at two of our meetings, or summit meeting, in which users are now beginning to discuss and publish how they are going about using that. And there's no question that when facilities see as you may have in your handout there see that there are over the 75th percentile they now work very hard to say why is that and, and then change their protocols.

So behavior is changing. So that these are becoming now higher quality events. I think it's important to realize that in the discipline of CT imaging, quality is also measured by the image quality as well as the radiation dose because the amount of those can affect image quality. And that affects the ability for a confident interpretation of whether a lesion is present or is not present.

Also, in terms of the availability of people being able to go to registry this is pretty much the only one around at least in the United States. It is the largest of the registries and now are then counted as over 10 million. So I think that – I just wanted to clarify that – that one of the very important things here is that every scan that's done on that scanner goes to the registry – there's no manual intervention. And they can't selectively send some and not the others –that's not the way it is set up.

The other thing that's extremely important in this area that just provide some clarification for the work group, just as it is in some other settings in medicine. People do not call the same thing the same thing. So one of the big

advances that's occurred her in terms of improving quality is that the facilities that join, use a standardized lexicon to determine what their exam names are so that you can compare apples with apples. If you don't do that it is very difficult for one facility to know how they stack up compared to other one.

So the, the mapping of the names and the fact that we do measure now in all the bodies, a size specific dose index, dose estimate has – has allowed us to hone in on what the tools that facilities can utilize to optimize their CT protocols.

Andrew Lyzenga: All right great. Thank you. And unless we have any specific comments or questions from the committee on this measure. Maybe we should move forward and just have a very quick review of the last measure on our list. Does anybody object to doing that? So we'll come back to you at the end of the call.

Judy Burleson: OK.

Andrew Lyzenga: So I think we had Pat Quigley and Lisa McGiffert on 2426. Elder Maltreatment Screening and Follow-Up Plan, if you wanted to just do us a few brief thoughts on that?

Lisa McGiffert: Pat, why don't you go ahead since I just did it.

Patricia Quigley: Oh, absolutely Lisa, thank you so much and yes this is the last measure of our discussion today. This is a process measure and sponsored by CMS and as a process measure, this is going after identifying the people over the age of 65 who have been reported elder abuse. So as a process measure it's – those people who've document at elder maltreatment on the screen using electronic malfunctions screening tool.

So the numerator in all those patients who are over the age of 65 will actually have this recorded in their Elder Maltreatment Screening tool, the denominator in all those patients over 65 and they're – as the process measure, the problem with this measure is that of course there are some people that can be excluded from this or from people who are not reporting.

But the aim of this quality measure if this is eligible providers and identifying suspected cases of elder abuse and we all recognize the importance of this. It does affect large numbers of high resource use. Most of those cases of elder abuse are under reported. I think we will all complement the CMS in identifying the literature and references 2009 and higher. So we have the reporting references to support this.

In terms of the specifications and reliability and the validity though, if there's a subject topic measure area in behavioral health, screening mental health, domestic violence and my question with that was really broad enough. It's really anyone who comes in the care of older people.

We already talked about the denominator being those people who have elder treatment screened on this positive screen but one of the problems is that they have multiple documentation tools. They're able to document this, and this measure as a process measure is only if it's reported once. So there's a period of time in a year from January 1 to December 31 where you would have reporting of this and if there's reporting of this and there should be a referral for treatment. So because a lot of the limitations of this measure is that there's an electronic reporting evidence all reported by paper that they've tried to look at during some reliability and validity testing of this measure.

So with that reliability testing, this was all handled by the quality inside the Pennsylvania and it was very cross-intensive to do this reliability testing. They looked at inter-rater agreement on inter-rater reliability with two RN on whether or not G codes were completed whether it was an indication of elder maltreatment and then whether there was and RN versus not and the one that was – had both RNs had greater inter-rater reliability.

But again, there was a cross factor that was associated with that. So the question of this still is going to be, you know the question of being able to use this as an indicator knowing that it is under reported. There's a lot of eligible patients that are out there that is – this is not occurring whether it's a paper documentation system.

So they have a tendency to further develop this and even with electronic reporting. So I think overall that we are all recognizing importance of this but there is limitations with the reporting and then following up to see if it actually was a treatment of this than having a referral.

Andrew Lyzenga: Great. Thank you so much Pat. Any thoughts or comments from the …

Lisa McGiffert: This is Lisa. I might add a few more things and when I first started digging into some of the reports of things that were – I was being referred to, I thought that maybe – I became very confused because one of the major reports that is cited throughout the discussion have found pretty strong indications that there was insufficient evidence to support this kind of screening as a tool to improve quality. They indicated that there really aren't any valid reliable screening tools and this, you know this measure really looks like the physician have to choose the tool and document the date. They used the tool and then if they find maltreatment, they would have to report it to some authorities.

Most states actually have mandated reporting for these kinds of things and so I'm sure that's under reported but I think that's an important thing to keep in mind. And it just seems that there wasn't – it's not a very feasible measure. It requires physicians to do a lot of work. It's kind of a heavy list for them. There wouldn't be standardized reporting. So I don't know how it'll be used by the public. It doesn't seem to be something that would give the public information about the quality of physician care.

Again this is, you know, one of those things that's more of a practice trying to get physicians to do something that is important and definitely this is an important issue to be addressed. But it seems to me that we need something that's a little bit more outcome-based than what we have right here.

Andrew Lyzenga: Thank you, Lisa.

Patricia Quigley: Thank you, Lisa and again, you know, I think the issue too is if it's really a quality measure just to say whether or not something has been reported once in a year's period of time, it's just documented when someone is seeing one.

Andrew Lyzenga: Any other quick thoughts from the committee or the workgroup? All right hearing none, let's give Judy a quick chance to say her piece on measure 0510. Sorry we didn't give you that opportunity earlier Judy.

Judy Burleson: That's OK thanks. Just real quick I just wanted to add that one piece of the measure when element is that having the exposure information whether it does through time and the follow up report. Also provide that in a readily available place so that the patient, his radiation history for that exam is not in case – it does come to, the patient has a radiation induced effects from the exam.

So that those information is there and makes it potentially easier to discover that that's the cause.

Andrew Lyzenga: All right, thank you, Judy.

Judy Burleson: Sure.

Andrew Lyzenga: And lets, operator at this time can we open it up for public comments if there's anybody on the line any members of the public would like to make a comment or question?

Operator: Thank you. At this time, if you have a question or a comment please press star then the number one on your telephone keypad.

And there are no public comment or questions at this time.

Andrew Lyzenga: Great, thank you. I think Suzanne have a few things to say about the next steps, but do we have any last thoughts from our committee or workgroup members? Any issues that you'd particularly like us to note for the in person meeting or the full steering committee?

Male: None (inaudible) it was already said.

Andrew Lyzenga: OK.

Suzanne Theberge: All right. Well thank you everyone for your time today. We really appreciate it. In the interest of time, I'm going to be real quick about the next steps. If you have not already done so our committee members, you should

begin reviewing the remaining measures in the project. The other nine measures that are going to be discussed at the in person meeting. We'll be following up as Andrew mentioned with summaries of these workgroup calls. So you can see what the other workgroups discussed on their earlier calls. But we do also want to review the remaining measures individually.

You should have received an email last week from our meetings department with your travel and hotel information, information on how to book that. If you didn't get that email, please let us know so we can follow up. We had a couple of people who's emails got lost in spam filters and whatnot. So just let us know if that didn't come through.

Other than that, well I look forward to meeting you all in person in about a month and you'll be hearing more from us with additional information about the meeting over the next few weeks.

Thanks very much for your time.

Female: Thank you.

Male: Thank you.

Male: Thanks, everyone.

Male: Thank you.

Female: Thank you.

Male: Thank you.

Operator: Ladies and gentlemen, that does conclude today's conference call. You may now disconnect.

END