NATIONAL QUALITY FORUM

Moderator: Sheila Crawford
February 26, 2014
2:00 p.m. ET

Lauralei Dorian: Good afternoon, everyone. This is Lauralei Dorian from NQF.

(Inaudible) echo, (I would ask that) everyone that if you're not speaking to please mute your phone and also if you're streaming to turn down the volume.

So, thank you very much for calling in today, and thanks to those of you who've submitted the thoughtful evaluation of the measure. Before we get started, I'm going to turn it over to my colleague Zehra to go through a roll call.

Zehra Shahab: Good afternoon. So I'm going to do a quick roll call. Is Don Casey on the line?

And what about Gerri Lamb?

Gerri Lamb: I'm here.

Zehra Shahab: OK. Colby Bearch?

Colby Bearch: I'm here.

Zehra Shahab: Shari Erickson? (Barb Hage)? Marcia James?

Marcia James: I'm here.

Zehra Shahab: James Lee? What about Russ Leftwich? Lorna Lynn? Terrance O'Malley?

Terrance O'Malley: I'm here.

Zehra Shahab: Ellen Schultz?

Ellen Schultz: Here.

Zehra Shahab: Beth Ann Swan?

Beth Ann Swan: Here.

Zehra Shahab: And is anyone else on prompt that was on workgroup 1 on the line?

Dawn Hohl: Yes. This is Dawn Hohl, H-O-H-L.

Zehra Shahab: OK. Thanks, Dawn.

Dawn Hohl: (Inaudible) for the first hour.

Zehra Shahab: OK. Great. Is there anyone else?

OK. Can the measure developers please identify themselves?

Dale Bratzler: This is Dale Bratzler. I'm here with the Oklahoma Foundation for Medical Quality team.

Connie Kuebeck: Yes. Did you want other members of the team to identify themselves from OFMQ?

Lauralei Dorian: That would be great. Thank you.

Connie Kuebeck: OK. I'll just go ahead and tell you who all is here. We have Casey Thompson, Toni Emmons. My name is Connie Kuebeck. We also have Elba Sisco. And then from our analytics team, we have Wato Nsa and (Christopher) – I'm sorry. (Christopher) and Alan Ma.

Lauralei Dorian: Great. Thank you for joining. Do we have any (Wanda) on the call? Or do we have (Jenna) (inaudible).

Gerri Lamb: Lauralei, we're having a hard time hearing you.

Lauralei Dorian: Yes. Operator, were you able to locate that line and mute it, please?

Operator: Yes, ma'am, I did.

Lauralei Dorian: Great. Can you hear me OK now, Gerri?

Gerri Lamb: Yes, thank you.

Lauralei Dorian: OK, great. And everybody's line should be open so if anybody who's called in, other developers, they should have an open line. Do we have Sam from New York City Department of Health?

Sam: That's right.

Lauralei Dorian: Great. And has anybody else dialed in that we haven't called yet?

Russ Leftwich: This is Russ …

Lauralei Dorian: Hi, Russ.

James Lee: Hi. This is James Lee. How are you?

Lauralei Dorian: Hi, James. Thanks.

(Crosstalk)

Lauralei Dorian: Hi, (Lorna).

(Jill Klingner): This is (Jill Klingner).

Lauralei Dorian: Hi, (Jill).

Zehra Shahab: Hi (Jill).

Lauralei Dorian: All right. Well, we might as well go ahead and get started. As I said before, thanks very much for calling in. I wanted to remind you that this call is being recorded and it's open to the public, and a transcript will be available following the call, a transcript of that call recording.

I would like to introduce Wunmi Isijola who is, I'm happy to say, will be starting on this project now as the project manager. We've had to sort of balance some work internally but I'll remain on here as well.

So we Wunmi would you like to say …

Wunmi Isijola: Hi, all. I am looking forward to work with everyone, and like Lauralei said, we will be working closely to make sure that there is a smooth transition. So, thank you.

Lauralei Dorian: The other item that I wanted to discuss, that we wanted to discuss with you, you might have noticed that already, but we actually have very few measures. Unfortunately, we only have one new measure submitted to this project, and this is sort of historic problem with care coordination. In our last phase, we actually had no new measures submitted to this project. And there's very little in the way of really crosscutting measure development that's happening right now.

And so, we determine that it will probably be a better use of resources for the time being to convene, not in person in Washington, D.C., but rather via two webinars. And so, the date of those will remain the same, and that'd be from 2:00 to 4:30 p.m. It will be the same March 18th and 19th.

So it's a little bit disappointing because we did want everybody, you know, we wanted everybody to be able to meet each other and us to meet you because it's definitely a different sense when we have an in-person meeting. But we wanted to assure you that we're going to make clear connections between these links, this work, and some other work that's being run out of another department here at NQF. It's a much more – it's a sort of – it's different plan strategic work. So, we're going to make sure that you are able to work with that group, and we have come up with an exercise that you will present, that we will present during those two calls.

So, I'd like to have the – Gerri, would you like to introduce yourself now?

Gerri Lamb: Sure. This is Gerri Lamb, and I'm co-leading the group with Don Casey. Don is out of the country. Hopefully, he'll be joining us but if not, we will just keep going ahead.

Lauralei, are we ready to launch or did you have other messages to share?

Angela Franklin: Well, thanks, Gerri. This is Angela. I had a quick just preview of what the purpose of today's call is and then we can get started.

So, again, this is Angela Franklin. I'm the Senior Director for the project and welcome to our second workgroup call. I'd like to briefly go over the purpose of today's call which is, of course, to have an in-depth discussion of the measures of the small group ahead of our full steering committee call that's happening in March. And we'd like members on this workgroup to bring to light any specific issues and concerns and questions that they have. This is also an opportunity to post those to the developers that are also on the call and developers have the opportunity to respond to question, and so committee can at this time, also ask for any additional information that they think is needed to continue to evaluate the measures from the developer. And developers can, if they choose, provide that information. This is the opportunity for all of that to play out. I would like to caution committee members, however, that we still must review the measures largely just in the form as they are as they are in front of us so this will not be kind of redoing the measures at this point.

We also use this opportunity for committee members to ask any questions about the process and the criteria. And if there are any questions about how the full steering committee call will be handled, we can take those at the end of this call, too; so this is an opportunity for that. And we're also piloting several new aids for the committee in their evaluation of the measures. And if you have comments about those, the guidebook or the algorithm, please forward those to us as well.

Related to those guides and aids to the committee, I'd like to remind everyone that there are several possible outcomes for any measure and including finding exceptions to some of the criteria. The committee does have an opportunity to bring their expertise and knowledge to this evaluation process. And NQF

staff definitely rely on that, and we feel that's the true value add to the committee, discussion and deliberation brings to this whole (step). So, just keep in mind that the criteria are set up as guides and also your values; your judgments and expertise are valued.

So, finally, just a process point for how we're going to through today's call, we're asking that the lead discussant introduce their assigned measure, and preview or quickly summarize their initial comments on the group about the measures, and that the secondary discussant also add their comment and then throw the floor open to the full workgroup for discussion.

Also, I guess, the final thing before we launch is disclosure. All members of the committee did complete a disclosure of interest form prior to being seated on the committee. But in the interest of transparency and to minimize the possibility of conflict, we'd like to quickly ask at this time if anyone has an updates to their disclosure.

Terrance O'Malley: Hi. This is Terry O'Malley.

Angela Franklin: Hi, Terry.

Terrance O'Malley: Hi. I just want to let folks know that Dr. Schnipper, who's the steward for the medication reconciliation measure, the last one, he and I work on a bunch of quality metrics most similar to measures 291 to 297. So, we do work together. I have nothing to do with his medication reconciliation work but I just want to let people know that.

Angela Franklin: OK, thanks. And just to say just and just some response to that, I'm not sure where we'll end up with that, but if NQF feels like there's a conflict, we'd ask you to recuse yourself from discussion of those particular measures where there is a direct conflict.

Terrance O'Malley: Right. I don't think there's a direct conflict …

Angela Franklin: OK, OK.

Terrance O'Malley: … but by all means, whatever is there.

Angela Franklin: Thank you so much. Any other?

Shari Erickson: Hi. I don't have any disclosure. This is Shari Erickson. I'm having issues getting on the SharePoint. Is there a – there's not a webinar link for this, is there? I've been looking for it (inaudible) go through SharePoint to find the materials.

Angela Franklin: Yes, we can most certainly send you that information via email.

Shari Erickson: OK, thank you.

Angela Franklin: Or email (inaudible) …

Shari Erickson: Yes, thanks.

Angela Franklin: Any other questions or concerns before we launch?

Now, I'll hand it over to Gerri.

Gerri Lamb: OK. This is Gerri. Again, welcome to everybody. Shari, if you go into the NQF SharePoint and just go on the agenda, the links to the webinar is there.

Shari Erickson: Yes, that's what I'm having trouble with. I can't get into SharePoint right now.

Gerri Lamb: Oh OK.

Lauralei Dorian: Don't worry. We'll send you an e-mail.

Gerri Lamb: All right. So, if there's info you need, just let us know, OK? But you should have most of everything else in front of you, so hopefully, that won't be too much of an issue.

Again, welcome and we're going to move into the measure review. I'd like to just highlight some of the process again that just went through. We're going to be reviewing the measures really to tee up what are the issues for further discussion at the March 18th and 19th at our full committee meeting online. We will not be voting today. So, this is really an opportunity to talk about the measures, any issues that you'd like to bring.

Lauralei's point about us working with the other committees is really an important one that the measure gaps committee and care coordination is revisiting the domains of care coordination. And so, if those issues related to capturing care coordination come up today, please bring those forward so we can have more dialogues about that. We will have an opportunity, I believe, to really dialogue with the other committee that's looking at that.

In terms of the process, we're going to go through just a couple of things to add to what Angela was saying. We have primary discussants and secondary discussants for each of the five measures we're going to be going through. I would just like to encourage us all to assume that the others have reviewed the measures.

If you would, the primary discussant, if you would start by giving us the overview of the measure, typically, that's going to be the front of the worksheet in terms of the measure title, the description, and the type of measure it is and when the most recent endorsement date has been. We can kind of review that just so that we orient everybody where we're at, then go through your review of the criteria, the important scientific acceptability, feasibility, usability, so forth, any conclusions. The secondary review or any addition to that, you don't need to repeat everything that the first person has said so really highlight anything addition that you'd like, anything additionally that you'd like to bring to the discussion, then we'll open it up.

In terms of the dialogue with the measure developers, it's really essential that the measure developers are here with us. I would suggest that we go through the review, reserve questions directly to the measure developers until the end of our discussion unless there are some pressing issues that we need to answer right away or to clarify misinformation as well because we've only got, what, about an hour and 45 minutes to get through the (slide), and hopefully, we'll have a very rich discussion about that.

And then the last comment is that – and I think Angela said this as well – is that this is an opportunity for us to kind of look at big picture as well as the specific criteria. And for those of you who have been involved in the measure review before, you know, that there's always this balancing act of trying to

look at where the measures fit in the domain of care coordination as well as the specific review criteria. And I think what we're all moving towards here in this committee, the previous one, is really to have a robust care coordination measurements set. So, our deliberation and discussion, I think, really contribute to that and are very important in terms of capturing some of the key structure, process and outcome of care coordination as well as what Lauralei was saying before, you know, the movement towards filling some of the significant measurement gaps.

So, with that, before we go into 487 and with Ellen being our primary discussant, anybody have any questions about process that we're going to go through for the next hour and half or so?

OK. Good. If you have questions as we go along, please just let as know and we'll keep moving ahead.

So, Ellen is our primary discussant for measure 487, and then Colby – Colby, do you go by Colby. You have an R in front of your name.

Colby Bearch: I do. Yes, I go by my middle name, yes.

Gerri Lamb: OK. So, Ellen is going to give us the lead-in and then Colby will be our secondary discussant on the measure. Ellen, are you with us?

Ellen Schultz: I am.

Gerri Lamb: All right. So, 487, if you launch us, please, that would be great.

Ellen Schultz: OK, great. Thank you. So, this is measure 487. It's EHR with EDI prescribing used in encounters where a prescribing event occurred. The description is actually very similar. So, all (patient) encounters in the past months that used an electronic health record that had electronic data interchange capability where prescribing event occurred, how many of those would the electronic prescribing actually use?

So, the measure steward is the City of New York Department of Health and Mental Hygiene. It’s a process measure which was originally endorsed in

2008. And the denominator is all patient encounters where prescribing event occurred, and numerator is in those encounters where electronic prescribing was used. This is applicable to ambulatory care setting in all patients and the data source-based electronic health record.

The level of analysis that's reported in the information sheet is individual clinician. However, there was some other documentation in that sheet I found in other sections that suggest that it's also used for the group of practice level. So, (that was a question that I had) for the measure developers maybe we can address a little later on in the discussion. There is no risk adjustment.

So, going into the criteria for the – first criteria of evidence following the evidence, algorithm that results in a low level of evidence. There is no systematic review provided and no evaluation of quantity, quality or consistency of evidence. There were a couple of studies in a (RAND) report site that provides some amount of evidence. But comments from the committee members that were submitted felt that that little evidence was fairly weak and questioned whether it clearly showed a link between the measure of number electronic prescribing events and other health outcomes.

And the committee – or excuse me, the (inaudible) staff noted that this measure would probably require an exception to the evidence criterion in order to endorse. So, maybe that's something we can discuss as a group.

None of the committee comments noted any additional relevant evidence, and there was no (voice) support for finding an exception to the evidence criterion. One comment noted that there is a similar measure included under meaningful use. That measure, I believe, sets an objective of 40 percent of prescribing events using electronic prescribing.

So, going into our criterion 2, looking at reliability testing. The testing was for reported for reliability as a measure score as a whole rather than the data elements, which is fine. That's one of the two ways of submitting reliability testing. Results were aggregated at the level of practices and for the region overall in New York City but no reliability testing was provided at the individual clinician level (scores). So, again, I think we'll need to resolve the

question of what's the actual level of analysis for this measure to see if the reliability testing matches. Following the algorithm assuming clinician level of analysis then the result is insufficient reliability testing evidence.

There also were barely minimal details on the measure specifications. There were several exclusions listed but not (inaudible) on how they would be documented or whether they occur with sufficient (groups) to impact the measure score. The comments for this aspect of the measure generally were that the exclusion should be reevaluated, more detail specifications are needed and the date on reliability on reliability are insufficient as currently reported, and there are questions about whether the testing sample was adequate for widespread implementation of this measure and whether this measure could be implemented consistently.

For validity testing, the validity of data elements was evaluated by comparing what the electronic transmissions to manual chart review. But there wasn't a lot of detail given about what the outcome in that comparison was. It was often noted that the observed electronic prescribing rates were a pretty good match to publish rates. But, again, we didn't the detail of what the published or observed rates were so it's kind of hard for the committee to judge how strong that match is. No information was provided by the ability to detect meaningful differences.

Also, (there were) some issues of missing data that were brought up, so it does look like there are some evidence with systematic (bias) (in the resubmitting) data particularly for practices that are later adopters of electronic prescribing. And that would interact with the numerator because there might be some cases that did use electronic prescribing but the transmission data is missing so it would look like they had a lower score. And the rate of missing data was fairly substantial at 20 to 30 percent of cases from what was evaluated.

So, the possibility of missing data from the denominator was not addressed. So, committee comments noted that validity testing was insufficient to really determine validity of the measure and that more robust testing is needed rather than relying on spot checks. It wasn't really clear that this a true indicator of

quality and that the issues around missing data need to be addressed to minimize the risk of bias.

For feasibility and use, the measure steward reports that this is used for public reporting by about 2,000 providers in New York City area. There was a link for that but a couple of the comments noted that we couldn't get any details about that public reporting, for example, could not find the actual public reported scores. It's also noted to be used for internal quality improvement, but again, there wasn't a lot of detail available.

So, a couple committee comments raised questions about the feasibility and use. For example, could this measure be used with standalone electronic prescribing systems that are not integrated with an HER? Is the data collection process seamless? And with changes to reimbursement structures, how would this measure be modified and relevant in the structure with reimbursement really downside risk and is it relevant in that context?

Let's see. Other comments about usability where that it's unclear if this is currently used for accountability. It's not totally clear if this measure actually could be used for the goal of high quality efficient healthcare. So, the benefits in use were kind of questioned in some of the comments, but no one noted any unintended or negative consequences that seem likely from measure.

So I think I'll pause there and turn it over to Colby.

Colby Bearch: I appreciate that. It's very comprehensive. I really (missing my concerns around that), again, the validity in the – in reading it just in general, just for the group to think about, it seems to me in healthcare, we solved this problem once before. It really were measuring how we're documenting prescriptions from prescribing clinicians and that, I mean, really to me is fairly obvious that if you use an electronic source to transmit that information that it is going to be more just reliable in itself, it's going to be more correct, and the human error piece from legibility perspective, you know, is kind of wiped away.

So, it just seems to me from a measure perspective for us to discuss as well, it just – it seems as if that it's been solved to me.

Gerri Lamb: Colby, is there anything else you wanted to add to your comments that Ellen didn't bring up in addition to the validity question?

Colby Bearch: No. In the spirit of efficiency, I've kind of gone down my checklist and she absolutely hit them all. The inclusion and the, the sample for looking at the (spot thesis), the word spot check in there kind of raised the red flag. And also looking at the risk of bias, yes, I think she absolutely covered it all.

Gerri Lamb: Great. Thank you. Thanks for such a comprehensive review, both of you.

Let's open it up for discussion of the rest of the committee and let's keep track of the questions for the measure developers so that we can get back to that as well.

Other comments about 487?

Russ Leftwich: This is Russ Leftwich. Are telephone encounters included in encounters?

Gerri Lamb: Russ, is that a question for the measure developer?

Russ Leftwich: Yes, I guess so.

Gerri Lamb: OK. Do we have someone from the City of New York?

Sam Amirfar: Yes. Hi, Sam Amirfar.

Gerri Lamb: Great. Thank you. Tell us your name again.

Sam Amirfar: Sam Amirfar.

Gerri Lamb: Hi, Sam. Sam, can you answer that question related to telephone?

Sam Amirfar: Yes. The telephone encounters were not included. It was meant primarily as measuring a viable face-to-face encounter between the clinician and the patient, and we wanted the opportunity for the physician to have all the electronic health resources available to actually use EDI. And we are afraid that if there was a telephone encounter, it might have been a verbal cue, and

we didn't want workflow issues of the clinic to interrupt or kind of complicate the measurement parts of this measure.

Russ Leftwich: Well, thank you for that clarification because I think that is important to know whether or not. It seems like it almost should be listed as an exclusion. But in practice, I submitted many more prescriptions a day by telephone encounters than by actual face-to-face encounters. So I – it just …

Sam Amirfar: Yes. We were just concerned that if the provider is doing over a telephone encounter and they did not have the electronic health records software available to them, then I guess it would have been, you know, easier for the provider just to use the usual workflow method rather than having opportunity to use electronic prescribing in a system. So, in order to make it a fair and even field, we thought it'd be fair if we counted electronic prescribing rates and it's kind of fair field where you're physically in front of a patient understanding that, you know, every three months, periodic amount of time, a patient would have to come and see the provider again before a provider (inaudible).

Gerri Lamb: OK. I think everybody was able to hear that, Sam. We have – one thing for everybody to remember is please don't put us on hold because we typically get music.

OK. Other comments about the measure before we go back to Sam and ask some additional questions?

OK. Now, here, I'd like to throw one thing out. Colby, this may sit with what you are requesting about the validity, one of the questions that I would like to raise about this that 487 is just to fit with the domains of care coordination, and this is probably something we'll need to have more discussion of when we all talk in March which is while EHR, using EHR for prescriptions if you have it may well be an antecedent to care for patients. Is it really, you know, a core component of it? And I'd like to have more of that conceptual definition that relates to face validity and then ultimately to I think the questions that Ellen and Colby were really relating about, you know, more of the construct validity.

Colby Bearch: Yes. This is Colby. And equally, exactly, from a care coordination because I still am kind of slicing the definition that we're using here for care coordination, but from an acute care or any type of prescribing clinician facility like the acute care or a clinic scenario where you're discharging that person from, that individual, patient from your care. Oftentimes, the biggest concern is, did they fill the prescription because, you know, that truly imply – it's not so much about the legibility and the coordination perspective but what after care have they (sought to me), just for the bigger, from the quality perspective.

And so, that measure may better fit my head in care coordination. I share the same challenge. I'm not sure how this really fits from a measure quality.

Gerri Lamb: And you know, for me, it's kind of what we were saying before, it's the big picture which there is face validity is how does this fit into the whole measurements of care coordination in addition to the specific item review.

Other comments about the measure that we'd like to bring to the March discussion?

Ellen Schultz: This is Ellen Schultz. So, I think speaking to the issue of, you know, set within care coordination, you know, I like your point, Gerri, that maybe this is more of an antecedent. You know, clearly, this is a form of information exchange, and having those electronic records of prescribing, you know, really can probably help with later processes such as medication reconciliation, you know, have a better written records documented of prescriptions in a way that's connected because in a reconciliation process, you know, I don't know how feasible it is to be able to start calling also to pharmacies and find out what really was prescribed or is there a record then that is an electronic record that could be, for example, connected to a hospital during an admission in a more direct way than a handwritten prescription.

But I agree that it seems like, you know, this given the high rates of EHR adaption that are going on right now, given that meaningful use is really pushing use of electronic prescribing. I do kind of question whether this measure is still needed now and certainly in another few years. As we move

in to the later stages of meaningful use, you know, this will be expected as standard practice and increasingly becoming so now.

But I do question some of that future utility, but in terms of looking at it today, for my view, I think it's clearly information transfer. It's maybe not – it doesn't get into some of the other juicier parts of care coordination that I think are especially challenging but it's sort of baby steps within that realm, at least as I see it.

Gerri Lamb: Well put, that's great. Any other comments?

Russ Leftwich: This is …

Gerri Lamb: Go ahead.

Russ Leftwich: This is Russ Leftwich. You know, to look at it a different way with respect to care coordination, it is a prescribing provider communicating and exchanging information with a pharmacist provider. And in terms of coordination, not really mentioned, I don't think the major discussion itself here, but in many analyses of e-prescribing, there is evidence that the number of callbacks from the pharmacy decreases substantially with e-prescribing and that, you know, reflects the fact that it may really be looked upon as care coordination where there is actually bi-directional communication about a patient's care treatment.

Gerri Lamb: That, great, Rich. So, conceptually, in terms of the current domains of care coordination, this would fit within communication and IT and it's that handoff of information. That's …

Russ Leftwich: Exactly.

Gerri Lamb: I think, you know, between what you and Ellen were saying is very helpful in kind of putting it in that bigger picture.

I'm going to move in to – I had a couple things down for the measure developer. Sam, are you still with us?

Sam Amirfar: Hi, we're right here.

Gerri Lamb: OK. Here's – let me tell you the questions that I heard and please feel free to respond to anything else that you wanted us to know about the instrument. One was Ellen's question about the level of the measure. It's listed as an individual measure, but it's also reported as a provider group, so if you could clarify that. The other is how missing data are handled. And the third one that I had is, is it possible to use this in a stand alone EHR system?

Sam Amirfar: Sure. So, the first one, the practice at the provider level. So, the 400, you know, line of data we have are actually from that practice level. You know, we work primarily with small (doc) (one Z) practices, so 80 percent of them are small (doc). So, sometimes, we use it interchangeably because there's only one provider particular practice. But if you have practices up to five, 10 and a couple of community health (benefits) which have 30 providers per practice, but the data came back to us on a practice level.

Number two, about missing data. So, it's – we have a couple other measures that are transmitted to us from an electronic health record. And if a practice is providing zero data across the board, we've assumed, and we got maybe almost 100 different, you know, process, and clinical measures back. If we got zero across the board, we assume that there was something wrong with the system, either a transmission issue, they're not using the system, or some kind of major problem that needs to be addressed. However, if the other clinical and process measures were transmitting valid data and e-prescribing was zero, then we did count that as a zero.

And number three, stand-alone process. So, we actually try to pilot project in the Bronx, using an Allscripts brand of just standalone e-prescribing and having providers use that. And what we found out is surprisingly, just having a standalone e-prescribing tool where there's a (Palm Pilot), I think we used a (Palm Pilot) at that point, or the mother tool is almost as hard as implementing electronic health record. And the only problem is at a certain point when you start adopting electronic health record, having a standalone system and electronic health record was a little bit confusing for the providers.

Gerri Lamb: Thank you. Thank you. Any other comments or questions for the measure developer?

Ellen Schultz: This is Ellen. I do have a request that maybe something to come back to us as an additional information. I did feel like throughout the reliability and validity section that there were hints of information there but really like a lot of level of detail was missing, you know, maybe this is me putting on my own measure developer hat, but I would want to see things like what is the variation of rates across practices and what was the variation like over time within individual practices? You know, what do the comparisons look like when you did a chart review or some other investigation to find out really whether electronic prescribing was used and then you compare to that to what was actually reported (on the measure)? So, I think, you know, presumably that's the information that you have and just it didn't include a lot of detail in a form, so they have something that I would be interested to see to really be able to get a better sense of reliability and validity of the measure.

Sam Amirfar: OK. (With a series of) (inaudible) over time in that period, would that be useful to figure out the distribution of the rate among the different practices over time?

Ellen Schultz: Yes. I mean, I think, you know, so one of my concerns was that, you know, some of what was put out there sounded like, you know, kind of describing the use of the measure, for example, showing oh, you know, the rates of electronic prescribing increased over time. Well, that's fine. So, that kind of assumes that it's already a valid measure. I'm interested in, you know, does the data show that the measure is actually reliable? Or are you seeing that things that are jumping all over the place over time with a single within a single practice? Are you seeing – you know, what's the variation you're seeing across? Does that pattern suggest that these are two differences or that they're similar?

Those questions are not always really clear-cut, which is why when we develop measures, we always really want to look at that data entity. So, yes, something like that or some tables that show, you know, rates, just that aren't rolled up to the same degree, I think that would be really helpful for my self to get a better sense of what's really going on with the measure.

Sam Amirfar: OK.

Gerri Lamb: Angela, Lauralei, is that something that we can rely on you to follow up with and see what we can get?

Angela Franklin: Sure, yes.

Lauralei Dorian: Certainly.

Gerri Lamb: OK, great.

Angela Franklin: And we can (be open to that), the form as long as, Sam, you can provide that information relatively quickly so we could send it to the committee before their call on March 18th and 19th.

Gerri Lamb: OK. Sam, thanks so much for being with us. We're going to move on now to measure 495. And, Lorna Lynn, you're the primary discussant; Shari, secondary.

Lorna Lynn, you want to start us off?

Lorna Lynn: Sure. And I'll mention – this is the first of three related measures. This measure is the median time from ED arrival to ED departure for patients who are admitted from the ED to the facility. The measure steward is CMS. And it's – the title kind of says it all, it's the median time from the time the patient arrives at the ED to the time that the patient is admitted to the facility.

The developers cite a well-recognized problem of overcrowding in emergency departments and states at reducing that time that patients are in the ED could reduce this overcrowding issue. They also say that it can increase the quality of care.

This is a – my (inaudible) in looking at this measure was that I honestly had a fair bit ambivalence about it if the question is going to be, does this represent care coordination? I think there are a lot of aspects of care coordination that are in here both within the ED in terms of assuring that the right people are seeing the patients expeditiously and that also links (inaudible) with people outside the ED who need to be involved in seeing the patients and making

judgments and assessments. What that optimal result, though, was pretty unclear to me and where there were several comments related to this in terms of the link between this measure and outcomes of care other than reducing ED overcrowding, which could be an outcome worthy of looking at in and of itself in terms of the measure relating to improvement and other areas, I think it's less clear.

The measures described differently throughout the document as a process measure and an efficiency measure, so I think I'll think of it as a process measure focused on efficiency. It was unclear whether or not turning to the importance of the measure. It's unclear whether the measure had lead to improvement in its use over the pass few years. The scientific evidence behind it (that cited) two systematic reviews that are of moderate quality, so I think in terms of the importance, my rating it right now would be low to moderate, but I'm interested in hearing comments from others and from the developer and just mentioning that comments from other sort of comments that were entered earlier as we're able to look at earlier this week, there's a lot of variability in the opinions from people who had filled out the worksheets.

The feasibility of this measure seems to be very good. It's been widely used by thousands of hospitals. So, it seems to be part of the workflow now. The scientific testing on it, there is good testing in terms of validity of the measure, I thought that the developers had done a very nice job with that.

With regards to usability and use, again, it's in very widespread use. There was a notation that there have been no unintended consequences noted from this measure, but I was actually curious about that. If there could be a rush to get people out of the ER and that does some care that might have been developed, delivered in the ED, would not be delivered in the patient could be transferred to a nursing unit that perhaps was not quite ready to deliver the care.

I think one example that was given was that with the lower time in the ER starting antibiotics for pneumonia could happen more quickly. I thought when you know that antibiotics could be started in the ED and perhaps there would be a further delay if that didn't happen.

Again, in terms of usability and use, it's just not clear whether or not this measure has led to improved outcomes of health. And once again, I would say that it's not clear to me what the right median time would be.

So, I think I'll stop there and see if the secondary reviewer has some additional comments.

Gerri Lamb: Shari?

Shari Erickson: Thanks, Lorna. Yes. Hi, this is Shari and I think you've covered it very well. I mean I would like to echo some of the things that you said. You know, I had some questions and seem like they were echoed in others as well as to whether this measure was, you know, really a care coordination measure. You know, it kind of focuses on (one ask one side of care). And, you know – and largely, you know, a lot – and I studied this issue. It's been several years ago now but they're – you know, a lot of the issues in terms of long wait times within EDs is related to the flow into the inpatient side of the hospital and the really the structures and processes that are in place from the inpatient side of the hospital rather what's in the ED.

And so I, you know, I guess that aspect could get it to a care coordination measure. But I guess I wonder, you know, a lot of what could impact long, you know, times from ED arrival to departure, actually it has to do with activities that are not within their immediate control within the ED.

Let's see, you know, I do think – you know, I guess in terms of feasibility, I agree, it's quite feasible to measure. It's the question of whether it really does get to improvement and health for the individuals going through. And, you know, maybe perhaps wonder if there, you know was some means of looking at it, you know, more specific to conditions or use of the patients, et cetera, because as the initial reviewer mentioned the, you know, I do think there – there is potential actually which didn't occur to me actually when I was first reviewing these of some (inaudible) if you're trying to push patients, you know, into the inpatient setting, you know, if it's not appropriate or if it's, you know, in order to improve on this measure.

So, let's see for anything else to add. I think those are the main elements that I wanted to highlight sort of base on my review of it and reading through the other reviews as well.

Gerri Lamb: OK. Thanks, Lorna; thanks Shari. Let's open it up for discussion. Are there comments about 495?

Shari Erickson: I'm sorry, this is Shari. The other thing I wanted to bring up real quick which I forgot till just now is the issue of the observation status and how that's being handled. I think that is an issue that needs to be looked at for this measure, you know, as it moves forward.

Gerri Lamb: Great.

Russ Leftwich: This is Russ Leftwich. Two quick comments – one, with respect to an unintended consequences remark. I think rushing to get patients out might be a consequence of overcrowding, but I don't think we can say it's the consequence of the measure. Second, I, too, was bothered about how this was care coordination and then in thinking about it occurred to me that one of the effects impacts negative consequences of long time between admission and discharge would be the longer patient individual is in the ED, the more likely they are to be in a change of shift hand-off, and that is care coordination.

Shari Erickson: And, you know, that brings up, this is Shari again, another thing that I thought about it. What I struggle with, with this measure was, you know, thinking about using it improve quality in terms of its impact on influencing decision-making at the hospital level. How do they know what changes to make to impact the measure? I guess you test different approaches, look at it more specifically over hand-off times, et cetera, but I'm not exactly (sure) how this measure by itself could really get to translate into meaningful specific approaches to QI to improve on the measure.

Terrance O'Malley: Yes, this Terry O'Malley. And just a take off on that comment, this is a measure that waiting times in ED, which is really a system measure rather than a site measure, is if you think of the roles that the ED plays in the healthcare system, it's really the triage point where you can get rapid evaluation and rapid stabilization. And then often, you're dealt off to some

other (plate). And now with the waiver on the three-day SNP overnight requirement, they are going to be direct admissions our of the emergency room to skilled nursing facilities as well as to home care as well as back to the office.

Because the ED itself is playing this really complex triaging role and it inputs sort of where to get the patients from and it outputs where you can send them to are really variable. And the (poor) ED is kind of right in the middle, and I'm not sure it has a whole lot of control over either what's coming in or where is it going out to. That's sort of one general comment. I think that – I would much rather see this as a system measure than an ED performance measure.

And the second is someone commented on sort of specific conditions, and certainly the extraordinarily long waits in the ED for patients with behavioral health issue, particularly who are awaiting transfer to some other behavioral health facilities, you know, it borders on being shocking, I mean it's terrible. And that's again sort of an indication of the access that the ED has these other sites of care. And the constriction of the ED flow is not with the ED, it's with the ultimate destination of the patient. Thank you.

Ellen Schultz: This is Ellen Schultz. I absolutely want to echo that point that I think if you're looking at this as a measure of care coordination, I think it does really reflect on the system and not so much on the performance of the ED. I mean certainly workflow processes within the ED will impact, to some degree, the wait times. But I think that, that question of the inputs and the outputs is probably far more important. And, you know, some other works that I've done, we talked about the ED as a window (into community) health because it does capture this cross section of what's going on and it's so dependent on, you know, what is the health of the population in access to care that determines of lot of the input? And then where can they send these patients going out? What resource is available, you know, how long do they have to wait to admit a patient on the output side?

You know, the other thing I'll add is that, you know, the one place where it seemed to me like it could be a little better measure of care coordination as

performed by the ED itself is in the context of an ACO. So, if you have an ED that's within an ACO, then, you know, a part of the task of that ACO is to really make connections across the different sites of care within that ACO network and that then, you know, the onus is on the ACO maybe not the individual ED to really help to make to those connections and make sure that their system has enough capacity to meet the needs of their patients so addressing the input side that maybe some patients would be better served to get care elsewhere if you can make that care really available and then on the output side to make sure that you have enough capacity to send your patients to whatever level of care they need once they leave the ED.

And so, in that context, it seems to me, to be a better reflection of care coordination, but again care coordination of the system, the ACO system not necessarily the individual ED itself. I think the a lot of the (inaudible).

Gerri Lamb: Hopefully we'll get that echoing under control there. Other comments about this measure?

OK, I would like to just add one thing about the overall discussion in terms of the fit with care coordination. I think this is very rich, and I think we're going to have to identify how in the March meeting that we come back to this and continue this dialogue. It strikes me that what we're raising here is that the causal chain that's being proposed here about reducing wait times it's linked to evidence based treatments, access to medications is that in fact is a proxy for the organization of care and ER and much more broadly in the system. And we may want to come back to looking at better specification of how some of these measures act as care coordination processes.

We did have a couple of questions for the measure developers. Do we have some measure developers with us related to 495?

Dale Bratzler: Yes. This is Dale Bratzler and also the other OFMQ team were all here.

Gerri Lamb: Great. If we could go back, if you could give us some insight, one, and – and Lorna, Shari, help me make sure we hit these. One was if you could comment on the level of evidence of this measure 2 outcome, if you want to comment on the comment about unintended consequences of moving people out of the

ER. And I think there was a question about where things are at with re-looking at the observation status and how that fits in that, that was suggested in the newest feedback. So, those three things is level of evidence, unintended consequences, and observational status.

Dale Bratzler: (So) …

Lorna Lynn: And I'd also be interested in – and this is Lorna Lynn. Do you feel like there is, you know, is (lower) always better? Or is there some ideal that you envision in the measure?

Dale Bratzler: Yes. So, that – let me take your last question first and then I'll address the others. So, when we initially developed this performance measure, we did so in very close coordination with the American College of Emergency Physician, the Emergency Department Benchmarking Alliance, and then have representation from other group such as the American Hospital Association and others, and we had a long conversation about, is shorter always better? And clearly the answer is no.

So, the great example – and one of the reasons that this is reported as a median time measure is we kicked around the idea of doing something that was done in the U.K. which was to set a set cutoff point and report this measure as a proportion; so, as an example, the proportion of patients who had ED throughput time of less than four hours.

We elected not do that because we knew a numerous examples of patients where at times it made sense to hang on to the patients longer to decide whether they really needed to be admitted to the hospital or not. So, a great example would be 2 in the morning, a patient comes in with chest pain. They don't – they have a negative EKG but you're concerned about symptoms so maybe you hang on to until 7 a.m. when a cardiologist can come in and do a treadmill exam. You may end up admitting them after that but you also may dismiss to home. So, that's why this measure is explicitly reported as a median time measure to avoid that, our concern about the unintended consequence of if you set a cutoff time, then people wouldn't be pressured necessarily to perhaps move them through faster.

To the other questions about the quality of the evidence, I can't give you a (great) level for the number of studies that have looked at outcomes related to emergency department throughput time but there have been a number of studies that have clearly showed that hospitals are on divert more commonly when they have crowded emergency departments, that there is some data that mortality rates are increased particularly for those conditions that are time-sensitive. Clearly, there are delayed care issues for conditions such as heart attacks, stroke, and others that are quite time-sensitive.

And then finally, it's actually pretty well-known that patient satisfaction goes down with long emergency wait times. Again, I don't know that I think someone – one of the reviewers at the first commented that moderate quality evidence – and I'm not going to argue that, there are really high quality studies that have currently linked patient outcomes to the length of time that they're laying in the emergency room because as we all know that may vary at the individual patient level. But clearly, patient satisfaction is not as good, and there is empiric evidence that when you have delays in the emergency department, hospitals do go on divert so your trauma center – I am in a trauma center here in Oklahoma and if we go on divert, there is no other level one trauma center in the state.

The question, I think we've addressed the intended consequences, and that's – I think I highlighted that. That's explicitly why we decided to use the median time. The other change I'll let you know about is our technical expert panel met recently and the hospital compare (display) of the data is going to start taking into account volumes (straight) for the number of the visits that an emergency department has annually because that clearly is known to have substantial impact on total throughput times.

The other question that I heard was about the observation data element. You don't have it in the data and we can update the forms for you. But we just – I think in the material we submitted, we highlighted that we had made a number of changes to the observation data element to improve its reliability. We just got back the reliability data yesterday. And the cap for 2013 data on the observation data element for the inpatient measure was 0.73. So, dramatically

improved from the cap that we had for 2012 data but that was based on substantial revisions to the data element specifications.

I think the final thing that I'll highlight is that we completely agree with all of your comments that we knew from work that was done at Georgetown like Urgent Matters and the Emergency Department Benchmarking Alliance that many of the interventions that have to happen to shorten throughput time don't happen in the emergency department. They are related to coordination of care for the patients that are upstairs in the hospital getting patients moved out of intensive care unit beds, timely discharges and other things. Clearly, the interventions oftentimes occur outside of the emergency department but this metric serves as a marker of those delays.

Gerri Lamb: So, would you say in regard to the conversation about this being a care coordination sort of signal, proxy, what's your thinking about its link to care coordination?

Dale Bratzler: We have usually, you know, in the framework of the institute of medicine, others have frequently thought of this as an efficiency measure, the efficiency of moving patients through your system and again most of – you know, I think this is, you know, you want to call it the proxy in the system. The efficiency, I think, that's very reasonable because from, from previous work that we've done around emergency department measures before this measure was ever rolled out, we knew that many interventions had to work outside of the emergency department. But this was a measure of how long it took patients to get into the system when they were sick and needed to be admitted.

Gerri Lamb: Great. Thank you. And thanks for your complete and efficient answers, that's very appreciated. Any other comments or questions for the measure developers?

OK, now going back then to what Lorna introduced when she mentioned of measure 495. This is a set of three related measures. We're going into the second one now 496. And I think in the interest of time. If there are similar critique, comments, discussion points, just highlight them quickly but move to anything that has not been mentioned before in 495.

Now, (Barb), are you with us?

(Barbara Hage): I am. Can you hear me now?

Gerri Lamb: Can everybody hear (Barb)? I can.

Female: Yes.

Female: We can.

Gerri Lamb: OK, great, all right. So, (Barb), you're primary, if you would launch us into 496.

(Barbara Hage): I'd be happy to. As mentioned, this is the part of three measures, and the difference between the one we just discussed and this one is that this is for the time between the ED arrival and departure for non-admitted cases. So, these are the discharge ED patients.

Again, it's y it's a CMS measure. Many of the details are the same. There is one point that I would add, I think, to that care coordination discussion, and I think one of the important roles of this measure team is in looking across the parts of the systems. So, I know there were comments about whether this should be more measure of a system versus measure of an ED.

In order to improve the quality of the services provided, I think that the components of a system need to do their part to coordinate with the other settings. So, as we think about accountability and we see the same issues with all the readmission measures, you know, who's really responsible for that readmission? If you want to give the incentive for all the parties to be working together so that the patient isn't sitting in the ED room for an extraordinarily long time, it does seem like a care coordination measure that you're giving the message that they should be coordinating with the upstairs floor that if the upstairs floor is having trouble discharging that last patient, maybe they should move on that so the ED doesn't get dinged for not moving them out. And similarly, if they're having problems discharging them in the community, then maybe they should be working on the relationships with the

other providers in the community. And I say that with some hesitation because I know those site populations are hard to transfer.

So, maybe doing some subsequent analysis stratified by the condition group would be important for the use of these measures. But these measures are currently being used in – it was noted in the QR program, the public reporting, the payment, the accreditation, the benchmarking, and they do seem like good measures of coordination in the sense that they really have to work with the inputting group and the outputting group in order to reduce their times in ED.

So, I won't repeat the rest of the – I did have one question. I noticed that the target population and all of these measures with the children's health and the senior care, why is the measure not applied to the general population?

Dale Bratzler: I'm not sure I understand the question.

(Barbara Hage): For the measure, the target population is listed as children's health or senior care.

Dale Bratzler: I think it is all age groups. Correct me if I'm wrong, Connie, but it's all age groups, is it not?

Connie Kuebeck: Yes, it is all age groups. And if those two boxes were checked on the submission that was sent, we need to state that when we went through and wrote those submission forms, I believe all of the boxes were checked. So, pardon us, we don't know how just those two came up, our apologies.

Gerri Lamb: OK great.

Female: (Barbara), are you finished?

(Barbara Hage): Yes, thank you.

Gerri Lamb: OK, Marcia?

Marcia James: For being the primary discussant on that. My comments on this go to the measure itself because I think everybody has kind of already commented on relevance with care coordination. But the question that I have is regarding the

exclusions on the measure. And the denominator exclusions currently say patients who expired in the emergency department. Is there an exclusion for patients who leave or, you know, leave the emergency department without, you know, against medical advice? You know, they didn't do ED – they're treated but they don't remain for the entire time. Is it assumed that they are not discharged? I mean how is that accounted for?

Dale Bratzler: So, Connie, again, correct me if I'm wrong, but for those patients, it's the time the patient physically departs the emergency department. So, I don't believe they're excluded, but there are times it would be shorter because, you know, the discharge time (tells that) to be the time that they physically depart.

Marcia James: OK. So, I guess what you're saying is that those patients then are not, they are included in the denominator. OK, thank you. I mean that's the only other relevant comments that I had on this.

(Barbara Hage): Marcia, that was a good point. This is (Barbara). I had one more question, and that was we're excluding patients who died. And I wonder if that isn't a group that should be looked at as well because they may have died from dehydration or some other unmet need in the ED room.

Toni Emmons: And this is Toni with OFMQ. Dr. Bratzler, they go into the non-reporting group.

Dale Bratzler: OK, so thanks, Toni. So, that brings up an issue. A couple of things that I've heard, one was the psychiatry, patients that have psychiatric diagnoses is also a group of patients that are profiled but they are not a part of the publicly reported metrics. So, a hospital (contract) their performance on a patient that has a psychiatric diagnosis or these other non-reporting populations that Toni just mentioned like the patients that die. But they don't go into the actual calculation of the measure that CMS post to hospital compare. We did that explicitly, particularly with the psychiatric diagnosis population because we know how challenging it is for many hospitals to be able to place that. So, they get the data and they can use it for improvement, but it's not used for public accountability.

Gerri Lamb: Other comments? OK, I'm not hearing any then.

Let's move on to the third in the measure set, and that is 497. And, James, you're primary. Again, this is set of three related measures, so if you could highlight if there are commonalities but emphasize what might be unique.

James Lee: Certainly. Well, first I appreciate developers' discussion around the framework of this system measure. And – but physically on 497, if we first do admit decision time to ED departure time for (admit a patient). And I think that's a very important part of the efficiency in terms of coordinating care and going from one care setting to another. So, I appreciate this particular measure.

I think questions for the community is that, one, decision to admit is a judgment and mental process, and there (variables) as to when a clinician decides to admit and then also (system-wide), when was it registered on EHR. So, maybe (the developer) can help us understand some of those issues. And then the second, is this a measurement (inaudible) system-wide? Will the accountability in the emergency room (end) versus the hospital (end) in terms of being ready to accept patients or having transport issues?

So, those are my questions, and I don't know if Beth Ann or others have other concerns.

Gerri Lamb: OK, thank you. Beth Ann, why don't you go and then we will ask the measure developers to join us? Do you have additional comments?

Beth Ann Swan: My only one comment tagging on to James and everyone else for 495 and 496 is that this issue of, is this a care coordination measure? And looking it at those, a micro level within the ED and as well the macro level in looking at it from a systems hospital perspective. Because in looking at this for all conditions versus specific conditions at the macro-level this is care coordination in the ED of many services for certain conditions but not all. And at a macro-level, it looks differently and I just wanted to bring out that nuance. But I agree with the other comments that have gone before.

Gerri Lamb: Great. Thank you. Other comments?

Beth Ann Swan: And can I just echo – this is Beth Ann again – I want to echo one point that was made, I believe, during the 495 conversation. Is this measure dates back to 2008? And now with the huge growth and observation units, what's the impact? And looking ahead of this measure, is there some way that influences what this measure is really telling us?

Gerri Lamb: Maybe we can invite the measure developers to talk about the two things that have come up so far which is how you think of observation, the change and the use of observations measures will influence this. And I think there was in James's comments a feasibility issue about data capture is, are you experiencing – or what are the issues related to capturing decision to admit?

Dale Bratzler: Yes. So, this is Dale again. Those are great questions. So, let me just start by saying that, again, for this particular metric, we did make the substantial change. The change to the definition of the observation term also applied to the outpatient measure and the inpatient measure. And so, again, for the admission measure of this admit decision time to departure time, the reliability that observation data limit is much better.

And I don't have data back to 2008. I can only look at it on my desktop right now. I'm looking at data for the past year so – and we haven't seen a lot of change yet in the performance on this particular measure. So, I am not sure I can predict how the increased use of observation beds may impact this particular measure.

A couple of points about the measure, there was enormous port to develop this particular performance measure by the American College of Emergency Physicians. Some of you may know this particular measure, this admit decision time to departure time as the so-called boarding measure. These are the patients for whom the emergency physician has evaluated the patient, made a disposition of the patient and now they wait. They wait whether it's transporter for an open bed, for all the myriad of issues that keep them waiting lying in an emergency department bed.

And I was just meeting with a number of the ACEP representatives here in the past week or so and they highlighted that the level of care – they have real

concerns about the level of care for these patients that are now, you know, officially assigned to be in-patients and yet they lay in a bed in the emergency department and whether they get the same level of care that they would in the inpatient bed.

So, that's particularly why we have the measure. And now, you're going to have to remind me of the other question beyond the observation.

Gerri Lamb: That was actually a really helpful lead into it. It was the question of feasibility. What's the ease of capture reliable information?

Dale Bratzler: Yes. So, fabulous question. So, you may have noticed from the reliability analysis, the kappa statistics for this particular data element on admit decision time actually, the major specifications that the hospitals have to use to capture the data have been extensively revised over time. And the reliability to data element as it exists today is pretty good.

That said, we are currently in the process of working with folks from the Emergency Department Benchmarking Alliance about whether we need to redefine the data element that starts the clock on this particular metric, this admit decision time. So, the discussion was lengthy, and we actually, the Joint Commission, OFMQ, and the Emergency Department Benchmarking Alliance Subcommittee on time stamps have a call next week to actually talk about this.

But, you know, as any good emergency department physician will tell you, when the patient hits the door, a lot of times, they know their decision is made as the, you know, the stretchers are rolling through the door that they're going to admit the patient. But they may not be documenting a "decision time" at that point in time until they have time to stabilize the patient, you know, work the patient up initially before they do start the formal admission process.

So, we are working on that particular data element. There is a desire by the Emergency Department Benchmarking Alliance to standardize that time stamp and it would be defined as something like, you know, the time that the actual hospital admission process is initiated which might be the request for the bed, it might be an actual order by the emergency department physician or

it might be the order by the admitting service that is required to come down to the emergency room to admit a patient. Many emergency room doctors don't have admitting privileges and can't write admitting order.

So, we are working on that data element with the Emergency Department Benchmarking Alliance to decide, you know, can we more standardize, find a better standardized time stamp for that recognizing that it's a little bit challenging because every emergency room has a little bit of a different process in place about how they initiate the admission sequence? But as it stands today, the reliability in the data element has been pretty good because it's so well-defined in the specifications manual.

Gerri Lamb: OK. Did I understand there is going to be a continuing discussion about that before we get back together again in mid March?

Dale Bratzler: Yes. So, there is an ongoing discussion. I don't know that we'll be able to rule out specification changes, you know. In fact, probably the best we could do with respect to the actual specifications in our hospital manual would be somewhere mid 2015. But those conversations are ongoing now in Joint Commission and OFMQ, CMS representation will all be participating in that process. The goal is to standardize both time stamps that are built in to an emergency department electronic medical records that standardize this time stamp of when this admission processes is initiated across all settings, all measurement platform.

Gerri Lamb: Great, great. Thank you. Thanks for that additional information. That was very helpful. Any other comments before we move on to the last measure?

OK. We're moving now to 2456 and this is a new measure. And our primary discussant is Russ. Russ, do you want to lead us in?

Russ Leftwich: Sorry, Terry O'Malley was actually going to take the role of the primary discussant (thing here).

Terrance O'Malley: OK.

Gerri Lamb: Oh OK. Great. Terry?

Terrance O'Malley: Sorry, Russ, (inaudible). Yes, all set. Thank you. So, this is measure 2456. It's lovely titled, The Number of Unintended Medication Discrepancies per Patient. The steward in a Brigham and Women's Hospital, and it's a measure of the actual number of discrepancies between a "gold standard," medication list and the admitting medication list and a similar process repeated at the discharge, end of the hospitalization, to identify discrepancies in the discharge medication list.

It is a process – it's an intermediate outcome measurement, and it measures the discrepancies between those two medication lists. And the target population is any adult hospitalized patient, and the measure is the number of medication errors identified in a sense that's a complication and it provides the foundation to achieve improved outcomes which is decreased medication errors. And a study quoted range of 2.78, 4.57 discrepancies per patient with a mean of 3.4 discrepancies per patient. And they also provided evidence which performance can be improved with the proposed process.

So, let me run through the criteria starting with the evidence base. And again, there are a lot of great comments provided and a typical, the comments ranged across the spectrum from high to low on virtually all of these. And if we take an average, they're all about in the middle but the evidence that supports the measure of focus is probably best summarized as moderates. There are a large number of studies of overall fair quality as pointed out by some of the commenters.

There is evidence to support a performance gap which is probably at least moderate, if not strong, and that data presented seem to demonstrate considerable variation on the performance of medication reconciliation process itself. And it seems likely that the performance data from this measure could lead to targeted interventions to reduce medication discrepancies.

High priority or high impact, again, sort of moderate; the nationwide data shows significant variation and have significant – the problem of medication discrepancies are. And it does seem that this is an overall problem affecting

all hospitalized individuals and all sized hospitals. In the frequency of error and adverse event, document the ability to reduce these errors and the fact you can reduce them, I think makes it an appropriate target. So, the range of discrepancies that the authors present, 3.44 as a mean, led one of the commenters to ask, you know, how many discrepancies do you need in a patient medication list to be significant and have a significant impact on their outcome? And that will be one of the questions you're going to toss to the developer.

The next section was scientific acceptability and the specifications (that) the overall rank here was sort of low to moderate, and the (feeling) in several of the commenters was that we needed more specification around the pharmacist's training and the detail, sort of what the details of that training are and (kind of the) standardized and how much certainty is there that the outcome will lead to a similar type of outcome gets to the reliability and validity issues later on.

Another commenter suggested that we really needed some more detail about how to calculate this measure just starting at the beginning of establishing the medication list. And then going through all of the steps of reviewing that list against the admission orders and the ongoing queries during the hospitalization and finally again, with the discharge orders that tells us more specificity about that process would be helpful to allow them to get a better sense of what's being done. And the reliability testing comment was – hopefully, comments are sort of low to moderate in sort of representing the ambivalence, kind of one hand, people thought the examples were very small; but on the other – making it difficult then to generalize.

On the other hand, the data did appear to demonstrate sufficient reliability to be able to determine a difference in performance even with a small sample. (It talks about) data on inter-rater reliability affecting cases with a couple of pharmacists that, again, appeared to be kind of small although it was high inter-rater reliability and likely reproducible. So, the question is whether the other sites that were cited particularly those in the marquee study give some more data regarding the reproducibility of this process.

And next step was validity and it's sort of the summary with moderate. Again, the test sample seems to be too small. But on the other hand, it seems to provide an adequate reflection of the performance on the measures. But what several – what's commented on was (a need) to be a clear connection between the discrepancies found in this review of the medication reconciliation process and the association with adverse patient outcome.

And then, so through the rest – it seems that the exclusion criteria were reasonable, they're very pragmatic. They didn't feel the need for some, for any risk adjustment or stratification, although several of the reviewers cited the question about age and number of comorbodities as well as the number of medications might be reasonable bases on which to stratify.

There was significant and meaningful difference in performance. And the missing data in minimizing bias that – it wasn't clear exactly what data could be missing and if there were, what impact that would have, but it seemed like the exclusion criteria were very reasonable and basically excluded folks who either died or were discharged before they could have this process initiated.

So, probably the most concerning questions were raised under the feasibility banner. And the concern was the amount of extra resources that would be required to gather the data to do, to establish the gold standards med list with a pharmacist trained to take a good, a better, a gold standard medication history and then review that at several steps during the admission to identify the discrepancies.

A lot of folks thought that that was going to be extra work. The countervailing comment was the extra work appear to be worth it given the fact that this measure seemed to really, for the first time, allow you to look inside the process of medication reconciliation and identify where the two major sources of errors occur. One is either no omission because you didn't know the patient was on the medication or was in omission at the time of discharge when you didn't, you failed to restart the meds that had previously been discontinued. So, again, it's sort of this ambivalence in the review. So, again, the concern, though, this is going to take a lot of extra effort, and they would like to see a better connection between the adverse outcome and the

reduction in discrepancies to be assured that the effort would be worth it in terms of a quality outcome.

As far as usability and use, probably moderate would be appropriate for both but would probably require more refinements. It's not currently in use although the stewards provided a plan to do it. And it's currently used primarily in a research setting, so it's not clear how easily it would to be to implement in a sort of non-academic setting. Certainly, a valuable measure for internal quality improvement, in fact, it's externally (inaudible), it maybe – it may require some refinements to use it as an accountability measure and that some, probably some re-stratification or risk adjustment would be necessary to do that. And it did not appear to be any unintended consequence.

So, I think that is the summary.

Gerri Lamb: Great, thank you. Russ, did you want to add stuff as the secondary discussant?

Russ Leftwich: Yes, I had a couple of comments around the source of the history and the fact that it's not well-defined nor the process for identifying those sources which would cost me to wonder if that should be captured, meaning which family member and how many contribute to that history. That might make it much different as the, who's taking the history.

With respect to the capability, there have been a number of studies that I was aware of previously that show that pharmacy technicians perform much better than nurses, other medication history takers including, I think, physicians, and maybe that would change the feasibility concerns. I don't think that would necessarily mean having a study pharmacist but you would – I guess, I would say, having a trained pharmacist or a pharmacist in (the facility) in this process.

And my other comment is a concern about whether the proposal to take a sampling of 25 random patients per month create some bias or some undersampling depending on the facility because of the references that suggest that the number of discrepancies goes up with age of patient and the number of medications which maybe linked. But if that's true, then 25 randomly

selected patients, depending on the patient mix in a hospital, and I'm not (inaudible) sort of intuited, you might undersample those folder more medications on the list of patient in a facility that had fewer of them. That's all.

Gerri Lamb: Great. Other comments before we invite the measure developer to make some comments?

OK, do we have somebody from Brigham and Women's with us?

Jeffrey Schnipper: Hi, it's Jeff Schnipper. How are you?

Gerri Lamb: Good. Thanks Jeff for being with us. Let me just kind of highlight some of the things that we would appreciate some additional info.

Jeffrey Schnipper: Sure.

Gerri Lamb: One was related to whether there's any idea in terms what the target is in terms of number of discrepancies and the link to outcome. I'm just going to list them, Jeff, if that's OK.

Jeffrey Schnipper: Sure.

Gerri Lamb: Two is any further thoughts on the conversation about the need or the benefit of risk stratification of this measure. You – in the materials we have, there's quite a bit of discussion on reflection on feasibility, anything new on that arena and then – I think that's it right now. Did I hit the major ones, Terry, Russ?

Terrance O'Malley: Yes, those are the (inaudible) …

Gerri Lamb: OK, why don't we start there, then Jeff, and if you could respond to some of those?

Jeffrey Schnipper: Sure, sure. No, (you said) a really nice job of hitting all of the, you know, hitting all of the important points of this measure. You know, we know this is a new measure and that it's, you know, it's got some rough spots but we also

acknowledge as the reviewers do that this is really the first measure that really is attempting to look at the quality of the med rec process.

And – but you have to realize that every one of these hospitals that we've studied using this methodology is in complete compliance with all the process measures that are out there. You know, (they're still committing) three hours per patient in their (orders). So, clearly, the process measures are inadequate to know whether or not a hospital is actually doing a good job and that there's at least a twofold variation from the best to the worst performing hospitals.

As far as target number of, or, you know, how many discrepancies you'd like to see it reduced by, you know, to give you a sense of the scope of this, you know, if there are three errors per patient of any type, it's about a third of an error per a patient that has significant potential for harm, so about 1/10 of all errors are potentially harmful. So – but, you know, all of these sort of dwarf every other category of med error that's out there. If you think of the amount of money that's been spent on bar coding where the rate of administration errors is probably like one in a thousandth, you're talking about one in three patients having a potentially harmful error. I think you can get that down by, you know, another, you know, 30 percent relative reduction, that's going to be a substantial number of patients who you're saving from harm. So, I would do it in terms of relative terms. I would say, you know, somewhere between a 25 percent relative reduction would be significant, you know, meaningfully from a clinical standpoint.

As far as links to outcomes, you know, I would say two things, you know, one is that, there's a pretty robust literature out there in general on the links from medication errors to potentially harmful errors to actual adverse drug events. And they track very closely in terms of high versus low performers and they track very closely in terms of response to treatment. And then regarding med rec errors, specifically, there are at least a couple of studies out there that show that if you improve the med rec process, you do result in fewer readmissions to the hospital into emergency departments. And so, you know, you can at least find a link there to hard outcomes of the intervention, you know, addressed at the med rec process.

In terms of risk stratification, we agree that if you wanted to actually compare hospitals and their absolute members, maybe some risk stratification would be in order. We didn't mandate it because we wanted to get the info from NQF, but we agree that, you know, age and number of meds would probably be a good place to start, it's pretty robust data that link both of them to the number of discrepancies. If it's like I wanted to use it for internal QI purposes, obviously, it's not necessary. You would just sort of track improvement over time.

And then finally, feasibility. You know, I would use the analogy of, you know, the NSQIP measures that are out there. Every hospital now has some research nurses who's job it is to, you know, read charts and to figure out whether hospitals are, you know, how they're doing with their NSQIP measures. So, this would be sort of a similar idea whether it'd be a pharmacy technician or a pharmacist that, you know, you would have on site to do this to finally know whether or not hospitals are doing a good job with med rec.

So, you know, we've done it now in eight different hospitals. Some of them are community hospitals, some of them are small as a 100 beds, as large as 900 beds, and anywhere from, you know, Sioux Falls, South Dakota to Atlanta, Georgia, so – and community hospitals, VA hospitals, and academic medical centers. And, you know, to them, they view it as a QI project more than a research project.

So, I think it is feasible. I'm not saying that it's easy to set up but, you know, we did create a five-year plan in sort of how do we get to a point where, you know, hospitals are doing this around the country.

I'm happy to address some of the other issues that came up as well or I can stop there.

Gerri Lamb: OK, thank you. Other, you know, in – I know that we have got to also open this up to public comment. So, let's see if there's any additional comments on the measure or for Jeff, and then I'll go back to Angela and Lauralei and see if we need to open the lines at some point here.

Ellen Schultz: This is Ellen Schultz. Yes, I just wondered if you could maybe speak a little bit more, Jeff, about sort of what the process for hospitals to start using this measure? I mean do they need a lot of support from your team to figure out how to implement it? Or we're they able to just kind of take it and put it in place mostly on their own just thinking about, you know, what's that learning curve look like for this one?

Jeffrey Schnipper: Yes, that's a good question. We had a teleconference with the study pharmacist in each of the sites where, first of all, in advance, they went through a slide deck and they went through a sample case. And then on the phone call, we answered any questions that they had, we reviewed the information. And then we went over the sample case, where they had to, basically, you know, go through this entire process documents, all the errors. And it was a nice opportunity to make sure that they were doing it correctly.

And then the next step we did after they had done about five or 10 cases, we, you know, talked to them again by phone, went over the cases, answered any questions that they had. And then over time, we've developed an FAQ document where, you know, I can see it almost like case law. You know, as questions come up, you know, how do I, you know, rule, in this particular case, so what about that particular case, you know, we, you know, we're able to provide them with that information as well.

But one thing that we did do that was higher intensity is that we did go to each of the sites and observe the process and action. And I think the question would be for feasibility, you know, we do actually do that or not, certainly helps for our research study, you know, say that we really the valid process that's reproducible and reliable from site to site whether you would that in reality, you know, I have questions about. But those are the kinds of things that we've done.

And the simulation case is actually pretty easy to do by phone where, you know, you play a patient and they interview you and they ask what additional medication sources they would want to have and you provide them that information in real time, and you get to evaluate them on sort of the process of taking that history and then eventually on the gold standard history they come

up with and you compare it with ours. So, I think most of it really could be done remotely, fairly efficiently, not that expensively, and like I said, we've done it now in a number of sites.

Gerri Lamb: Thank you. Any other comments or questions?

Ellen Schultz: I had this one quick follow-up question, like do you have a sense from the site of how long it takes the pharmacist to do a particular case?

Jeffrey Schnipper: You know, it's an average of 20 minutes per case. But there's a wide range, you know, some patients, it takes five minutes and some, it takes an hour. It depends on how long a medication was there on, how will they understand their med, how available the sources are, you know, how much thinking you have to do.

You're going to have a family come, you know, bring in all their pill bottles from home and dump them on the bed and do what I call brown bag rounds, that obviously takes a long time. But that's a small number of patients. But yes, on average, it's about 20 minutes.

Gerri Lamb: Great. Angela, Lauralei, should we open it up at this point what you saw here?

Lauralei Dorian: Definitely. Operator, can you open up the lines, please?

Operator: At this time, if you would like to ask a question, please press star one on your telephone keypad. We'll pause for just a moment to compile the Q&A roster.

Again, for questions, please star one on your telephone keypad.

We have no questions at this time.

Lauralei Dorian: Thank you. We do have some comments via the chat. Would you like to read some of them?

Zehra Shahab: OK, so we have some comments from Don Casey. He wanted to emphasize that there are instances when patients arrive too quickly ED to an ICU or

inpatient unit without clear and well-defined plans of care which can create confusion on the receiving end in the interest of reducing ED wait times.

There's also another question to keep in my mind. How does this measure ensure that the medications are the correct ones as specified by the various clinicians especially non-hospital clinicians who care for the patient? Also, how does this measure perform within the context of current and future meaningful use requirements? And these aren't public comments. This is a comment from a committee member, our co-chair Don Casey.

Gerri Lamb: OK. Jeff, can – do you – would you be able to respond to Don's questions about within that measure? Has there been any attention to kind of the issue related to accurate medication use and whether it's the right medication and anything like that?

Jeffrey Schnipper: Right. I mean we train our study pharmacists to gather as many medication sources as needed. You know, the general principle is that, ideally, you've got a list, say, from a pharmacy or from your outpatient medical record or recent discharge summary plus the patient has their own list which is, you know, by memory or list they have or their pill bottles or something like that and, you know, if the sources agree with each other and the patient seems to know what they're talking about, you know, the process can be done fairly quickly, if they can explain away, any discrepancy as it may not take that long.

But if there are issues and you don't have two independent sources or you can't reconcile the sources or the patient doesn't seem to know what they're on, that's – and you have to go digging. At the end of the day, we can't guarantee that the list is perfect. You know, we use the term best possible medication history to acknowledge the fact that we try to gather the best history we possibly can, but that includes outpatient medication sources, and what they're filling at pharmacies.

So, at the end of the day, you know, we hope that that's as reliable as we can possible get. Is it perfect? No, I don't think anything is perfect. But I think this is a standardized and, you know, and reproducible as much as possible, and we think it's pretty high quality.

That was the second question that he asked, is that correct?

Gerri Lamb: Yes, actually I was going to ask about that, if you remember it, go for it.

Jeffrey Schnipper: Second question was – no, I can't remember now, sorry.

Gerri Lamb: Lauralei or Wunmi, can you give us the second question that Don had about this measure?

Wunmi Isijola: About the meaningful use question.

Lauralei Dorian: That was …

Jeffrey Schnipper: Oh, the meaningful use, yes. You know, meaningful use specifies, you know, doing the process, it specifies documenting it, and, you know, that there's a – there are some steps around post discharge med rec that requires, you know, never in meaningful use or any other measure does it actually say how accurate a job did you do. And, you know – and that's why, you know, this is really the only measure that sort of gets at that. It certainly is not in conflict with the meaningful use requirement. But I would say it's complementary to link, you know, process of documentation with actual quality of your process.

Terrance O'Malley: Hi, this is Terry O'Malley. I'd go one step farther and just say that this is a measure of the integrity of the process itself of establishing the medication list, and that's the point we've never gone to previously. So, we have this whole med rec process predicated on a (flood) medication list. And this is the measure that takes that and pushes it back to sort of the best list available which is really a remarkable step.

Gerri Lamb: Thanks, Terry. You know, one of the topics that I'd like to see us discuss – we don't have time now but with one that both Terry and Russ raised – and I think this maybe Angela, Lauralei, Wunmi a question for you for the March meeting is they raised the question of the difference between accountability measures and the usefulness for quality improvement just to make sure that we have some clarity on that and when we come back to review this measure.

Angela Franklin: OK.

Lauralei Dorian: Yes.

Gerri Lamb: OK.

Lauralei Dorian: Yes, we'll definitely take note of that. And …

Gerri Lamb: OK.

Lauralei Dorian: OK.

Gerri Lamb: All right, well then, we have gotten through it. We actually are close on time. I would just like to compliment everybody, what a – just a wonderful discussion, thorough reviews and really excellent input from the measure developers, so thank you all for that. This has been a really productive session.

I mean let's bring it back over to Angela and Lauralei and Wunmi now.

Lauralei Dorian: Yes, thank you again, everyone. We do really appreciate all of your feedback. And just to echo Gerri's comments, we really appreciate just your input and some of the responses (inaudible) that we do, in fact, look forward to our next steps.

And, Zehra, could you just kind of talk (to) some of our next steps and then the anticipation for our two-day full committee meeting.

Zehra Shahab: OK, so just to recap, as Lauralei had mentioned earlier in the call, the two-day full committee meeting will be on March 18th and March 19th, and it is set right now for 2:00 to 4:30 p.m. You will be receiving an e-mail within the following, in the coming week with an e-mail including the agenda and then the next steps as well, and you'll also receive a calendar invite with the time and these two dates as well.

Lauralei Dorian: And just an additional note, we will be posting the workgroup summaries. So, a lot of the discussion that we've had today, we will be inputting them into the measure set (stats) located on SharePoint. So, please feel free to look that over prior to our two-day committee meeting.

And with that said …

Angela Franklin: I think that's it for our call. Gerri, this is Angela, I just want to thank you for your excellent leadership on the call, as well as thanks to all of the committee members for your hard work on this complex measures.

Gerri Lamb: Great. And if Don is listening, Don, we missed you and we look forward to getting back together in March. Those times that you gave us, those are Eastern, yes?

Lauralei Dorian: Correct.

Gerri Lamb: OK, because we're going to be going through daylight savings between now and when we do that again. So, the times are just going to weird across the United States, so we should keep that in mind.

Lauralei Dorian: Definitely.

Gerri Lamb: Thanks …

Shari Erickson: Hi, this is Shari. I couldn't hear the times very well, I'm sorry. So, is it 2:30 to 4:30 on both days, so it's just the afternoons, or is it all – I'm sorry, I just didn't hear it very well.

Zehra Shahab: Shari, it's 2 o'clock to 4:30, both days.

Shari Erickson: OK, thank you.

Zehra Shahab: And this is …

Lauralei Dorian: We will be sending additional information in the coming weeks, so you can prepare accordingly.

Gerri Lamb: Great. Thank you and thanks to everyone. Great call.

Female: Thank you.

Lauralei Dorian: Thank you.

Female: Thank you, too.

Female: Thanks, bye bye.

Male: Thanks, bye bye.

END