

NATIONAL QUALITY FORUM

Moderator: Reva Winkler
February 27, 2013
11:00 a.m. ET

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

Reva Winkler: Hello, everyone. This is Reva Winkler with the National Quality Forum.
Thank you all for joining us today, as we have a meeting of the eMeasure
Feasibility Technical Expert Panel.

The purpose of this call is for the panel members to review the comments that were submitted on the draft report that was posted for public comment last month. We are looking to get feedback and comment and responses from the panel members to the many public comments that were submitted. We received more than 90 comments from 10 different organizations.

As you could tell, as we were signing in, we do have audience members who are here joining us. There will be an opportunity for public comment later in the call.

For the committee members, you received two documents that are the main materials for this call. The first is an Excel spreadsheet which contains every one of the comments that were submitted. This is a sortable spreadsheet so that you can sort it by who submitted the comment or by topic area or whatever else you may want it to sort it by.

Just so everyone's aware that after discussion with the panel and consideration of all the comments, we will be entering a response to each and every one of those comments in the spreadsheet and posting the final result on our website

with response to each comment. So, that is sort of a final product of the effort we're beginning today.

In addition, there is a memo that outlines the comments and groups them by topic area to facilitate this morning's discussion. Both of those documents are posted on NQF's website.

So, before we turn it over to Dr. Lieberman to begin the discussion, I just want to remind the panel that what we're really looking from you today is a discussion in response to the comments, the comments address a wide variety of subjects, some of them very specific to the recommendations that you have made. We're looking for your thoughts on whether you wish to make any adjustments here, recommendations, any clarifications, do we need to explain or be more explicit at any point in the report in response to the comments.

Some of the comments, I think, are very specific that will be addressed by NQF staff, those specifically addressing formatting or organization report or things like that. But those that are specific to the recommendations made by the TEP, it's very important that we have your input and your thoughts on how you want to respond to these comments.

So, are there any questions from the panel members in terms of what we need to do on this call today? All right. Then, Dr. Lieberman?

Michael Lieberman: All right. Thank you, Reva.

Reva Winkler: ... should I turn it over to you?

Michael Lieberman: Yes. Let me check, still working on that thing we had at the WebEx, I'm very close. Let me do that. And then, how we would do is go through each – we would start with the comments regarding the recommendations, because there's kind of a need of the paper that we're producing, and you know, we only have, you know, less than two hours to get through them.

There are lots of comments. I want to make sure that we got through those first, and then we can circle back and go through some of the more general once. And as Reva mentioned, really what we're looking to do is decide, you

know, whether or not we kind of agree or disagree with the comment and then whether we want to make a change to the recommendation based on the comment.

So, let me – so let's get started by, actually it would be on page three of the comment memo, the recommendation one, which is the assess feasibility throughout eMeasure development, and if you want to – the actual recommendation was that feasibility should be considered early at the eMeasure development process and assess throughout the entire duration of measure development as an agile iterative process, greater collaboration among measure developers, EHR vendors and providers at all stages of development, will promote more rapid evolution in the iterative functionality for current measures as well as to build the capability for more complex EHR enabled eMeasures in the future.

(Paul): Hey, (Mike). Do you want to tell us where you're reading from or where are you? How do we ...

Michael Lieberman: Oh, OK. So, I read back just from the draft report. So, I would be ...

(Paul): Yes. I mean...

Michael Lieberman: ... probably working – I was going to work from two different documents. The draft report has the actual text that people are commenting on, and then the memo document has the actual recommendations.

(Paul): So, can you move first to the page number of one of these documents so we can follow along?

Michael Lieberman: Yes. So, in the draft report, it's page eight, recommendation one.

Reva Winkler: (Paul), this is Reva. There's a link on page one of the memo to the draft report, if you don't have it otherwise available.

Michael Lieberman: And also, I think it was sent out this morning for the call.

(Paul): This morning, OK. OK, that explains it.

Michael Lieberman: And Reva, I assume that you'll then be taking notes and you're getting what's on the conversation. OK.

Reva Winkler: Yes. And as a reminder to everyone, we are recording this call. So, we'll have that for reference as well.

Michael Lieberman: OK. OK, so should we jump in to the first recommendation then, which is – or the comment collaborations require. Now, on the WebEx, Reva, who is driving that, and do they want to pull up the memo on that.

(Kathie): Hi, this is (Kathie). I'm running the WebEx. Right now, I'm having a problem with screen sharing, but I have help on the way.

Michael Lieberman: OK.

(Kathie): In the meantime, I do have links to all of the materials on the WebEx, but one statistics I'll try to show what you're referencing as we go along.

Michael Lieberman: OK. So then, again, we were going to go through the PDF memo that Reva put together just because they kind of, she's collected the comments by recommendation, so we thought that will be a good way to do it.

So again, on the – on the comments PDF, it's page three under recommendation one, assess feasibility throughout eMeasure development, we have collaborations require active and ongoing communication with stakeholders as well as additional time and resources during the measure development process. Obtaining feedback multiple times may increase the amount of time needed at the beginning of the process to send requests for feedback, communicate with stakeholders, analyze feedback, and integrate suggestions into measure specifications. Although we agree that earlier identification of infeasible measures will be more efficient in the long run, these additional steps may impact the timeline for measurement development.

Keri Christensen: This is Keri Christensen from the AMA-PCPI. Do you want us to just jump in if we've got comments on this?

Michael Lieberman: Yes, please do.

Keri Christensen: And so, I will just share from the HHS lead (Kaizen) works that we started about what, three or four weeks ago now around measure development and making that process more lean, we actually determined that this kind of work earlier in the process, we felt, would actually cut out time, because there won't be the same number of rework loops.

And so, although I agree that it does add work up front in the process, I think that I still think we should recommend that that work be done earlier in the process to be consistent with the work coming out of that lean event.

Michael Lieberman: Great.

Debbie Krauss: This is Debbie, I agree. This is Debbie Krauss, I agree with what Keri has just said. I was going to say the same thing.

Rute Martins: And this is Rute Martins with The Joint Commission, ditto.

Male: And this is (inaudible). The measure developers are all agreeing. I would definitely support their motion. I mean, the rework that possibly completely starts resend, I think if there's anything, one of the things we've learned from both this process in eMeasure collaborative is a hardy exchange at the beginning just put them to so many more false starts.

Michael Lieberman: OK, great. OK, so let's move on to the next recommendation. This approach is focused on measures as the root of the issue of the fundamental underlying problem as the lack of semantic interoperability to enhance clinical care, care coordination, clinical decision support and measurement. It is important that the recommendations for eMeasure development also impact data usage within and among EHRs to enhance clinical care.

Any comments on that?

(Paul): Well, I think the measures can be inappropriate or improperly constructed measures can cause a lot of destruction in terms of the efforts on the part of providers. So, I think that's sort of where we're headed. I mean, that was our intent at this work, which is if the measures are put together with a lot of information, a lot of collaboration up front, then it will be much more

meaningful and much more efficient at the back end. So, I think in the end, that reworks that providers should have to both do the work and as well as measure.

Michael Lieberman: Yes, and it seems to me that the eMeasure system is really – does actually help with semantic interoperability. So, I don't quite understand what the comment is trying to get at.

(Paul): Do we have a reference to who made that? Because then we could know the background maybe better understanding background, because sometimes the intent is there and we're just not – we're not able to read it. Do you know where this came from?

Reva Winkler: I will take a little bit of time (Paul), but you can find all of the comments over on the spreadsheet.

Male: I think I already know.

Male: (Inaudible).

Male: That was from a gentleman from the American Board of Internal Medicine, is that one of the – or the American College of Physicians (inaudible).

Male: It's the ACP, ACP.

Male: Yes.

Male: Do we need any other discussion on this one or should we move on? Let's move on. OK.

Male: I think when you read the actual sort of comment, it appeared that they were trying to broaden the scope in terms of semantic interoperability, helping clinical decision support and in many other areas of I suppose, care delivery. That's sort of the gist of what I got from the actually comment itself.

Male: OK.

Male: But I don't know if it fits into the feasibility of just eMeasures.

Michael Lieberman: Right, and I definitely don't think it would change their recommendation. If anything it, you know, it might be something that we would want to play through in a comment in kind of the introductory area where we're – because I think that is one of the, not so much for any measure of feasibility but eMeasure altogether is trying to enhance the use of data for both measurement and decision support.

The next one is, add the inclusion of NLM vocabulary experts in the early stages. Request for new codes from NLM and its supported standards have been made at the tail-end of the process. These requests are the canary in the coal mine for signaling problems with feasibility (if there is not an existing code, the data would never be in a current computer record.)

Male: And that came from (Clem McDonald).

Keri Christensen: This is Keri Christensen, again that's consistent with my recommendations coming as lean event, I think that'd be worthwhile. I don't know where it sits in the paper though, that's my only concern.

Male: Well we could add that we have greater collaboration among measure developers, EHR vendors and providers. We don't make any reference to vocabulary bodies. Should we add that? I think it would be reasonable.

Keri Christensen: Got your point.

Male: So, I'm wondering if we, in our recommendation about the collaboration at the front-end, if we both enumerate the number, the kinds of stakeholders we're interested in and the process for this iterative feedback, we might be addressing a lot of this comments at the same time and also that's probably one of our biggest recommendations.

Shannon Sims: Hi this is Shannon. I just, I think I agree with everything that's been said. I don't know if it necessarily needs to be an NLM expert per se but just a vocabulary person in general (inaudible) cautious that the existing of – existence of a code and centralized vocabulary doesn't necessarily translate to automatic feasibility within an EHR.

So, I don't, I think clearly there's not a code that's bad but the fact that there is a code doesn't that it's good either, so.

Zahid Butt: So, Shannon this is Zahid. I agree with you but I think that (Clem's) comment was specifically directed at the need for new elements or new code sets. That would ...

Debbie Krauss: I think, this is Debbie Krauss. I think keeping it general including the stakeholder terminology expert at the beginning where we explained that collaboration early and often is good. And if it doesn't have a code, it just signals that we may need to request it. That may be an issue but it should not stop us from pursuing the measure development.

So, I would not think we would want to get into the details of the end of the comment where they talk about, if there's no existing code. Yes, they're not capturing it currently, but it doesn't preclude us from capturing it in the near future.

(Paul): So, I think there's two, this is Paul, there's two aspects to the suggestion. One is explicitly including vocabulary experts in the process. But, two is explicitly including NLMs, they're the curator of the vocabulary. So, it works both ways in the sense that if you want, one, they know what's in it and two, they would have an influence over the work to be done. So, it's sort of both things are helpful.

Michael Lieberman: Yes, and I think, I mean what I'm certain with a little bit is whether the comment is trying to get at, I mean I think involving people familiar with terminology really in the process is a very a good idea. But it's, you know, doesn't really need to be a, doesn't need to be the NLM or for that matter, you know, the SNOMED folks, if it's a SNOMED code or is it more that, you know, as you're developing measures it's not whether or not you really need a code and in some ways whether that code should exist is part of the question as well.

But I think, so I think having the comment that you want terminology specialist involved early in the process is good, but then again I actually don't know how you, how you necessarily operationalize that (Paul). You have kind of mentioned the process for that as well. Do you have any, you know, any ideas about how that process would work?

Rute Martins: Mike, this is Rute from The Joint Commission. I think, I think what you just said makes a lot of sense. There are actually two aspects to terminology expertise and the measure development process. The first one is the user or an expert user of the vocabulary, making sure that we're using existing codes that are the right codes and appropriate code and for other representation of a particular concept.

Then, there's the editor of, I would say the vocabulary editor piece. In which, if there is no appropriate code, then when a new code needs to be requested, we need to make sure that it fits in with the vocabulary. And that's where I guess I see the preponderant role for the NLM. At least for SNOMED they would be the brokers for that.

Michael Lieberman: OK. Are there any other comments on this one or should we, or we're ready to move on then. The next one is, recommend that developers map to the QDM and codes early on to allow for time to request new codes, if necessary. I mean, I would have assumed that that is already kind of the process. Is that not the case?

Male: I think it is part of the process and probably doesn't need to be specified by us.

Zahid Butt: Well, I think maybe what the (inaudible) and the measure developers can let us know is a measure definition can proceed far along just on concepts alone and then at the end they say, we don't have a standard code for that and then, end up having to request it at the very end. These two comments seem to say that. So, if that's true, then they're suggestion is that to involve people with vocabulary expertise early on and one of the ways to do that is to check to make sure the concept you're asking for in the QDM.

Rute Martins: This is Rute. I'm just not sure how this relates exactly to feasibility. It seems more of a process comment in terms of the efficiency or the order in which the steps are performed.

Keri Christensen: This is Keri. I agree. I think it's true and it's a good point that let's maybe led to lean process work cover instead of this work.

Male: What's the output of the lean process that would effect, that would get back to the measure developers.

Keri Christensen: One of the deliverables is a standardized testing plan for testing at all points in measure development. So, that guidance would be given in that and it is covered. I'm just not sure where it's at. I agree that this just not relates to section specific things we think measure developers should start doing.

Debbie Krauss: And, it's also covered in a number of other processes that closely align or overlap with the testing, so it is covered maybe data elements (eSpec) creation or one of those two processes.

Male: Yes. It's most of these actually refers to kind of eMeasure data element feasibility. If we're saying that that needs to occur early on in the process, then we are saying that they need to be mapped to QDM and standard codes set early in the process to know, you know, whether they can be.

Michael Lieberman: All right any – I think we're ready to move on to recommendation number two.

Howard Bregman: I would like to make some comments for recommendation number one.

Michael Lieberman: Sure.

Howard Bregman: I think they're somewhat related. This is Howard from Epic.

Recommendation number one is really a process recommendation and at the end of the process it says, it mentions one goal of feasibility measurement, which is the rapid evolution and EHR functionality. But really, our document does not have a section that lifts the goals of feasibility measurement, which I think is a deficiency.

And I think really what this recommendation should say is, we recommend this process in order to achieve the goals of feasibility assessment rather than what it says now, which is that we recommend a certain process and it only highlights one goal.

Of course, another goal could be the measure set aside because it's determined by provider feedback or other feedback that it's not feasible at this time and it should not be endorsed. So, right now it just calls out rapid evolution of EHR functionality. What I would like to see it say is, it achieves the goals of measure feasibility assessment, and then in a separate section which should be at the beginning, we also add what all of the goals of the assessment should be.

Reva Winkler: Howard, this is Reva. Aside from the two you mentioned, are there other goals that you have in mind?

Howard Bregman: Well, I think in general, we do mention a goal in our overarching principle. We say the fundamental goal of performance measurement, oh that's performance measurement, to improve the quality of care. Our goals are to provide the users of the product with the ability to collect, generate data for their own purposes, for their own quality improvement measure efforts.

Another goal of the assessment process is basically to, basically endorse the measure and say this measure is feasible at this time so it's ready, it's ready to go, and I agree that promoting new functionality in EHR vendor is another goal. I think others can come up with more, but I think there's several that we could, at least even if we have a limited number of goals, we should call them out.

Reva Winkler: OK, we can do that.

Michael Lieberman: No, actually Howard I like that a lot. I don't, in another goal that I would have for feasibility assessment is I think it's somewhere in the introductory area as well, but it's to expose issues with feasibility early on in the process to and throughout the process.

As we kind of discussed earlier, people may choose to use a measure that is less feasible than in another measure if they feel there's enough value out of it. But, we want, we want people to have the information necessary to appreciate the problems with the measure or with the issues with the feasibility throughout the process.

Howard Bregman: Right. So, I would like to see recommendation one say more of about here's the process we recommend and the end result is that our goals are achieved without calling, without linking a specific goal to this process because it really should benefit all goals.

(Paul): So, this is Paul. I agree with the multiple goals. One of the things we are, up on the front, we are trying to solve, I don't want it to get lost is we're not looking to stay stuck in the past or with the current systems.

So, I think the reason that the phrase for rapid evolution and EHR functionality is because we want the EHR vendors to try to go towards the future needs. That's one of the very loud voice is about the current state of affairs is that people are not happy with the measures or the ability to implement the measures. So, we're ...

Howard Bregman: Well, I think that's fine.

(Paul): ... trying that toward the new world.

Howard Bregman: I think that's fine. I think that's an overarching.

(Paul): (Inaudible).

Howard Bregman: I agree with that. I think that should be an overarching, sorry, go ahead.

(Paul): I don't want to water down that message.

Howard Bregman: Well, I think that message should be in an overarching principle. That should be one of the principles.

Reva Winkler: OK. Howard, I got note of that and I'll be happy to revise things along those lines.

Michael Lieberman: OK. So, for recommendation number two, framework for eMeasure feasibility assessment says the TEP agree that a framework for assessing eMeasure feasibility would provide a common language and provide decision makers with valuable information by the technical feasibility of eMeasure. Building on the work from the (inaudible) proposed framework addressed assessment of technical feasibility of data correction only and does not address the value of the data on eMeasure.

Assessment of the feasibility of data owned as well measures being conceptualized and specified can identify significant feasibility issues before the measure is field tested. The feasibility assessment should address both the data owned as it's subsequently the measure logic aggregation and reporting.

And then the recommend, the comment was, to promote an open dialogue between all stakeholders and allow for greater flexibility, this report should outline the standards and types of questions for feasibility assessment, rather than outlining a prescriptive approach to obtaining answers to those questions. Such guidance would better support the identification and discussion of the obstacles to feasibility than a potentially game-able scoring system.

Comments about the comment? So I think that, I mean I guess I would. I think that the – that we want – that what we're trying to do is give people more of a framework to do, and I don't think we're entirely prescriptive about how this should be done.

And the game-able scoring system is kind of another issue that probably warrant some discussion and that some of later comments talk about you know, how this is used in promoting a measurement so on. But overall for this comment, I feel like the – that we do – that we do need to give some framework and some more structured format to do the assessment and but still have it be in the spirit of doing exactly what they say, promoting an open dialogue.

Anybody else have a comment? Anybody agree or disagree with the comment that the – that the comment were made?

Rute Martins: This is Rute. Would it be allowable for measure develop to use a tool other than the score card to answer the same question and I guess, then we need to better formulate the questions that we're trying to answer in the score card so that we would allow for different approaches to answering those questions.

I'm not sure it's clear from the report right now that the score card is the suggestion rather than a requirement.

Male: Well, are we – are we, this is a little bit like standard. If you have a lot of optionality then you don't have a standard. I think NQF does need a standard, it's just like its criteria, its criteria set. And it's fine for people to develop things and have different attributes that they want to measure for their own purposes, but in the end there has to be something that they even know how they're going to be met, how they're going to be assessed.

Zahid Butt: This is Zahid. I think Rute made a good point that if you even look at the subsequent comments; it appears that people feel that this will be mandated tool which is not the intent, as Mike just mentioned. I think, this is providing more of a framework, but not just totally sort of without any structure. So, it's more of a structured framework within which people can use this to develop tools as this see fit to actually implement it.

So, perhaps some clarification of that, if that is the intent, and if everybody agrees on the TEP that this is a framework that's just sort of structure within which tools could be developed to actually operationalize this framework. If that is the intent and that is what we're trying to convey, then perhaps, you know, making it clear along those lines might be helpful.

Male: Could NQF maybe speak to the needs here? What would be of use?

Reva Winkler: This is Reva. I think, you know, NQF has the criteria to evaluate all measures for feasibility. eMeasures have their specific characteristics and issues around feasibility. I think we're trying to identify some greater guidance about what kind of information will provide stakeholders, implementers, also to the end users with some way of assessing a measure's characteristics around feasibility.

It's certainly is a subject that has been brought up and we get feedback all the time about. You know, greater assessment, greater testing for feasibility. We need to know more about feasibility and how this will work to better understand how realistic this measure will be for implementation. And certainly, that kind of information when a measure is being evaluated for endorsement, will speak to the criteria for feasibility.

So, feasibility has never been a required must have or any kind of absolute threshold associated with it. But the guidance to help both the stakeholders to understand feasibility of a measure as well as to how to have expectations for implementation down the road. So, you know, we are looking for a little bit greater guidance, specific to eMeasures that will help all the stakeholders understand as eMeasures come through.

Helen Burstin: Yes. And just to add to that – this is Helen. Thanks, Reva. The other pieces increasingly, as the eMeasures are moving in to being used from major accountability programs, there are concerns about how feasible these measures will be, particularly the, you know, issues around the ability of – reliability and ability of the measure result.

So, I think, being able to provide this information on, you know, ability to at least capture the data is a really important input for multi-stakeholder group what they consider measures.

Male: It sounds like the more precise we can be, the more helpful it is. It goes again to standard let me give a whole lot optionality just because it's less helpful. So, we may have to better explain or potentially revise this category, availability accuracy standard. But are we saying we don't agree with having these listed?

Michael Lieberman: No. It sounds like what we want to say is that an assessment using this framework is required, that there's no absolute, you know, scoring system. There's no absolute path fail associated with it. So, it's not that you have to meet a certain threshold in order for the measure to move along, but you do need to provide that information so that it's easily accessible, so people, you know, during the evaluation of the measure and the endorsement process, and

then even pass that, you know, whether or not you want to use the measure particular program.

Male: And these are suggested attributes that some groups have found useful.

Michael Lieberman: Well, I mean, if that – I think, if you are just pointing out, I think, we could still make the assessment using the scorecard required for all data elements and then have the report, you know, as we've discussed, on this data on it's required part of the measure and the measure submission process.

Male: I think that's where we're headed, right? I think that would be useful.

Michael Lieberman: Yes.

Debbie Krauss: This is Debbie Krauss from CMS. I just caution though, requiring a scorecard to be produced for all the data elements, because you don't want to have any rework. We have many data elements that we've already used that have passed feasibility and we're implementing, and so we don't want to do rework at them just to fill out a scorecard as new measures come along.

So, we want to, but we want to stand up a framework and I think state that you'll have to give explanation of how you assess the feasibility and here's to recommend a framework to assess it as you take the measure forward for use and endorsement. So, I'm a little concerned that we require a scorecard.

Helen Burstin: Debbie, this is Helen. Just a brief response, and that's a great question. And we already have built in to our criteria. If there's evidence out there that a data element is already determined to be reliable and valid, and that can just be put forward, and maybe we can think to be more explicit here that, you know, as long as there's some evidence that somebody can put forward, you know, of the, you know, 11 data elements in this measure, you know, prior analysis, you know, and point to where that analysis might be, would suggest those are, you know, can be reliably and feasibly can be collected, then you know, we could ask them just to focus on the ones for which we don't have that prior information.

Male: And in fact, the information for the scorecard on existing data elements could exist in the QDM, one of its original purposes.

Helen Burstin: Yes.

(Paul): And what should have – then we'd actually encourage behavior that if you're looking forward to, wow, these things that are readily available, reliable and authoritative, why don't I just use that. And hola, you'd be reusing things that exist and then offer some comfortability. That was sort of part of the thoughts behind the QDM.

Howard Bregman: (Inaudible) also, part of the feasibility assessment is not just evaluating a data element in isolation. It's evaluating it in the context of the measure. So, I agree that you may have data elements that are already vetted and judged to be feasible, but in a certain context, they may be problematic. And this is a communication mechanism to convey to the developer that there may be a problem.

I mean, I have to say, I just, yesterday on the CMS website, raised an issue about the feasibility of a very established data element that just does not appear to work in a certain context.

Helen Burstin: Could you possibly share that. It would be helpful. I'm not sure I can wrap my head around the context issue, or just a general example.

Howard Bregman: The general example was – are you asking me?

Helen Burstin: Yes, please yes.

Howard Bregman: The general example was the principal diagnosis which is a, you know, diagnosis as a – for a standard data element. Principal diagnosis, you understand where that comes from, but in the CMS – I forgot what they're called, the notes that were sent out recently.

Helen Burstin: The release notes or the project document.

Howard Bregman: Yes, the guidance, the guidance document.

Helen Burstin: Project guidance.

Howard Bregman: It basically said that the principal diagnosis from an inpatient encounter is basically the diagnosis that appears on the billing statement that's coded as – well it's flagged as the principal diagnosis. And of course, diagnosis has an attribute of a start time and a stop time, start time and an end time.

But a billing statement just includes the diagnosis and a flag. It does not have any start or stop time. So, if we specify that the principal diagnosis is the diagnosis from the billing statement, we can't then say that it has an attribute of a start and end time and then design a measure saying if a diagnosis started on a certain date, then that defines the denominator.

Helen Burstin: Yes, that's helpful.

Michael Lieberman: OK. So, we have – so I think that we decided that the scorecard would be useful, especially and maybe, you know, one of the options for describing feasibility of a data element and other options might be, you know, evidence of prior use and validity.

OK. So, now we're going to move on to 2.1, which is the data element feasibility assessment, and I don't think I'm going to read the entire section, because there's little too much here. So, let's jump into the comments, and then, if we need to go back and figure out what part exactly they're commenting on, we can try to do that.

So, the first comment is, it is unclear how the score card will be used to drive meaningful use of EHRs. A potential negative consequence of this tool may be the development of eMeasures that are possible, rather than eMeasures that could be possible.

And you know, I think that we address that by saying that we were going to do the feasibility assessment, those with regard to what is out there now, as well as what has been specified in meaningful use criteria and so on. Does anybody else have a comment on that?

Howard Bregman: This is Howard from Epic. During our discussion in the in-person meeting, I think we did say, and everyone agreed with the idea that a negative – a low feasibility rating does not mean that a measure cannot go forward. It is just one factor among many that need to be considered and it is not the dominant factor, nor do we have a threshold below which you cannot go forward with the measure. And I think that that addresses this point, or we should – if we don't say it in a document, we should.

Kevin Larsen: This is Kevin Larsen. My personal goal for this is that we make intentional choices about those attributes and measures that we want to have as stretch goals rather than accidentally stumbling on things that are really difficult to do and may they don't – they're not really moving us forward or driving a lot of value. We just invest a lot of energy in doing something that's hard. And the other thing is don't we still have the column of three to five years?

Michael Lieberman: Yes.

Kevin Larsen: OK. So, I mean, that was sort of our intent, it could be made possible.

Michael Lieberman: Right, right. OK. The next comment, the use of a prescriptive and standardized process does not allow measure developers and their partners to tailor the feasibility assessment approach to their needs, nor does it encourage innovation and advancement within the field. Listing, separate from the scorecard, the important components of data element feasibility that must be considered and modifying the scorecard description and instructions to indicate that the resource is a tool and its use optional is recommended. A rating system such as the one proposed is subjective and could be gamed through the purposeful selection of more technically advanced or aligned EHR systems in the assessment process.

I think we've kind of already discussed this issue in our initial discussion. Are there any other comments about this? All right. I'll move on to the next comment. It is unclear which entity would complete the scorecard (EHR vendor, measure developer), how that entity would complete the scorecard, and how the process for completing the scorecard would be operationalized. Since the scorecard proposed within the report would only apply to one data

element, we are concerned that the process for completing the assessment would be highly labor intensive.

(Paul): This is (Paul). Just a comment about this comment. The two choices they gave are, I think, are a little bit off, at least, what I think we were headed towards. The entity that would “complete the scorecard,” I think should be more from the perspective of either of the user rather than the EHR vendor or a measure developer, because I think that’s where the workflow and what does this data element actually comes into play.

Keri Christensen: This is Keri from the AMA-PCPI. Built on that, I might disagree slightly, and that I feel like the measure developer is responsible for getting the scorecard completed, but yes, it obviously needs to be in collaboration with vendors and actual providers with EHRs and you can't state that in the back end. Can we just clarify that?

(Paul): And maybe the overall theme here with I think our whole body of work is this collaboration is important because there's the experience and onset each perspective had. We'll have a far better product if they start to consider all with this three, (inaudible) plus the user are evolved almost from start to finish, and finish in the sense in the scorecard.

So, maybe that’s a theme we could strike, we can make more explicit. And it suggest yet another way of showing how that would work.

Reva Winkler: OK, (Paul). This is Reva. We can certainly make that more clear in the report.

Howard Bregman: This is Howard from Epic. I do think that there's a couple of small changes that can be made to address this comment. We do want to convey the idea that a number of different types of entities can complete the data, the feasibility assessment and yet in our data availability paragraph, we say the EHR vendors can determine how often data is captured.

But really it's the whoever is doing the assessment is who's doing it, not necessarily an EHR vendor, that’s one of the possible entities. Also in the first sentence, we say, the extent to which data is readily available in the

structured format across EHR systems, but really any feasibility assessment is from any one feasibility assessment is really only likely to address one EHR system.

So, while the overall combination of feasibility assessments would likely spend more than one, really any given assessment is only going to look at one.

Debbie Krauss: This is Debbie. Actually, when we did feasibility on data elements, we did work with multiple vendors on the same data elements. So, we did look at multiple systems for the same data element. So, I imagine that they're different. Developers may do it different ways, but we felt, in order to get a large picture, that was important to do.

Rute Martins: And this is Rute ...

Howard Bregman: Well, that's – I think that's (inaudible).

Rute Martins: ... I think if your point is that you're going to do an individual assessment and then, or the measure developer will do individual assessments with multiple systems, multiple providers perhaps, and then aggregate the findings into one scorecard for that data element.

And that goes back to my point that we discussed on the draft report prior to it's release, which is that the scorecard should be able to work for a single assessment as well so that there is some standardization in the individual assessments. And then, there must be some sort of way in which we can aggregate that, but that the assessment, I agree with you, Howard, that the individual assessments need to get to the aggregate findings.

Michael Lieberman: So, it seems like in general, the consensus is that the measure developer ultimately is going to be the one that is responsible for the feasibility assessment of the measure, and in collaboration with EHR developers and clinicians, you know, the users. Is it worth – so Reva, is that what you said you were going to specifically call out earlier in the draft?

Reva Winkler: Sure, we can do that.

Kevin Larsen: This is Kevin. One addition to that is that the National Library of Medicine is interested in figuring out how they can also help with the feasibility of the data element level and potentially even store some of that information in the value set authority center. So, they have been engaged and willing partner in this, so that if it's done in one context or one measure, it could potentially be that information could be leveraged for other measures and other context.

Michael Lieberman: I think that's great. I think, actually the context issue is a really important one there. So, we think it would be useful to have kind of repository and the scores for the data element, but then I think we would also need the ability to comment or somehow assess the application of that assessment to that particular measure.

And I think, as Howard was mentioning, I'm not sure if it's the exact, same sort of issue, but where a measure may be feasible, data element may be very feasible in one measure the way that it's used and may be less on another. And I think we talked about, I don't remember exactly which it was, if it was vital signs or temperatures during the labor process that they were collective, you know, many, many, many times where it may be a feasible data element as being collected once.

When you collect it a lot of times, it's going to be less feasible. So, you'd want to have some way of kind of capturing, perhaps the measure-specific issues with that data element as well.

All right. Ready to move on to the next one then. Ideally, assessment should be performed by clinicians using EHR systems today rather than vendors. It is important to differentiate from fields vendors make available and actual data capture. I think that we all, from the previous discussion, I think, we all agree with that.

Zahid Butt: I think, Mike – this is Zahid. I think we agreed that it should be vendors and their clinicians, right.

Michael Lieberman: Yes. In collaboration with the clinicians and not vendors.

Zahid Butt: So, ideally, it should be the two of them together.

Michael Lieberman: OK. And then, the next is, It would not be appropriate to expect all measure developers to obtain a representative sample of EHR implementation, due both to the great variability in EHR systems and the burden placed on the vendor community; in addition, feasibility assessment efforts under federal contracts often are subject to abbreviated measure development schedules and constrained by Paperwork Reduction Act regulations on data collection.

And I think, I mean, this gets at a larger question of, you know, what is an adequate example to do feasibility assessment.

Rute Martins: This is Rute.

Keri Christensen: This is Keri. Go ahead Rute.

Rute Martins: Go ahead.

Keri Christensen: I was just going to say we're not prescriptive in reliability testing as to exactly what needs to be done more is better. And I don't think we'd want to be more prescriptive for feasibility than we'd want to be for reliability.

Rute Martins: And I completely agree with that statement, Keri. This is Rute. I would add to that, that we already had some discussion about the key factor being transparency. So, when presenting feasibility or reliability results that there is transparency in the number of sites, the number of EHR vendors, but because that will frame the constraints and limitation of the testing result.

Howard Bregman: This is Howard from Epic. If I can draw our attention on page seven, there's actually a specific recommendation where we apparently prescribe the number of systems that should be – that should be reviewed or assessed and this common erases objections based on the (O&B) requirements that that would entail. So, that's – I think it's directly related to this point. So, we should ...

Michael Lieberman: In which – Howard, which ...

Howard Bregman: This is on page seven. If you just click, if you just skip to page seven.

Michael Lieberman: The comments to document?

Howard Bregman: Yes, the document that you're in.

Michael Lieberman: Yes.

Howard Bregman: And you look right where it says down about the middle of the page, it says remove the recommendation that steering committees consider assessments of fewer than three ...

Female: I absolute agree.

Howard Bregman: ... EHR systems and 10 installation is insufficient.

Female: I don't remember that we agreed on a number of sufficient installation at all.

Female: Yes, especially 10 because 10 is more than nine and nine is (inaudible). So, we've got a mixture that we're not recommending something that the government contracts can do.

Female: This was a conversation that you all had on the conference call we had in early January where you were talking about what you thought would be sufficient. You're hesitant to make a threshold, but felt that there needed to be some recommendation around where you thought meaningful information would be obtained.

Keri Christensen: And that's fair. This is Keri. I mean, this is what we do in operation. We try to get three EHR systems and more installations is better. Typically, that's somewhere between five and nine. But yes, I'm comfortable with even making a recommendation that's specific, because then, that sounds like that's all you have to do and that you have to do that many and I'm uncomfortable with both sides of that.

Male: So, if we have come to both.

Female: Hi, this is (inaudible), and I completely disagree.

(Paul): So wait, are you saying you want to go back on number three?

Howard Bregman: We'll not specifically say three.

(Paul): I mean, we did – as Reva has pointed out, we did come up with the number three and now we're no longer in agreement with that.

Female: I'm sorry. I'm confused now, I'm looking at 10, not three.

Male: Well, three EHR, 10 installations. Yes.

Female: Oh, I'm sorry. I see that now.

Male: The sense that it apprise to.

Howard Bregman: I understand, and Keri, you can maybe clarify, it's the 10 that's the issue, 10 installations and not the three EHRs.

Male: I don't know if it's number 10 because it may be months old, but I do number three.

Keri Christensen: Yes.

Male: I'll just read the sentence that this applies to.

Howard Bregman: (Inaudible) and we felt that it should be at least a certain number of installations that should also be important ...

Keri Christensen: Yes.

Howard Bregman: ... or that would be important. So, Keri ...

Keri Christensen: This could in fact ...

Howard Bregman: ... this is an issue in terms of the (O&B) that it's five to nine installations.

Female: It can't be more than nine if anything.

Howard Bregman: But could we say that it is recommended that a minimum of nine and, you know, that would get out of that ...

- Female: No. I'm super uncomfortable with that, and I think it gets back to our definition of feasibility.
- Male: Right.
- Female: If we have to find nine places that can do this, that's incredibly expensive, as well as, you know, for some measures, it just may not be possible. And I'm thinking the cutting edge measures, so like what we're doing with functional status assessment. If I have to find nine – there's no nine places in the country that are doing that.
- Male: Sure.
- Female: We can never ever measure around it. I mean, I think it's good to see a variety of EHR systems, it's desirable. The more installations you can assess, the better. But to set a hard number is going to limit the tapes and measures that we can do that I think we want as a country. And I wouldn't want to have that (inaudible) come at this report.
- Male: That went back up to three?
- Male: Well, I think we could say – could say – like you said, a variety of the EHR vendors might be, you know, that's obviously more than one ...
- Male: Perhaps, I believe, it's multiple implementations.
- Male: Yes. And idea – and I would say like, and ideally, that's multiple implementations, but I don't know that we even need, you know, realizing that there are times we may only have one implementation up in EHR. But again, back to the ...
- Male: It would be possible with multiple implementations.
- Rute Martins: And I guess – this is Rute. I guess, it goes back to the point of whether we're trying to convey what's desirable and ideal versus hard minimum requirements. And my perception of our discussion in general in January was that these were more of a guidance element than a requirement for feasibility.

Male: Right.

Rute Martins: And going back to Keri's point earlier, since we definitely have these request (inaudible), I don't see (inaudible) through feasibility.

Male: Yes.

Michael Lieberman: I think one question that we could, you know, (inaudible), there is, it sounds like (Paul) you're still in favor of having a specific recommendation around a number in saying three EHR systems and others want to keep that more – don't want to put a specific number. Are there any other comments around (inaudible).

(Paul): Let me see if Keri agrees about three, but no defined number of installations, and I certainly understand that point.

Keri Christensen: I would not want to recommend that we say that no measure can move into, you know, the national forefront if only two EHR vendors are currently able to do it. We all know that EHR vendors responds to measures who are endorsed, and it's putting their requirement on the measure developer to get EHR vendors to pick up a measure before being able to ask for endorsement when most vendors will not put the measure into their system and tell it it's endorsed and that's just a circular logic that I think is going to be very difficult for people (inaudible).

Howard Bregman: This is Howard. Can I – I would like to read the sentence in the draft that refers to this, because it does not say that it has to be endorsed.

Keri Christensen: Right.

Howard Bregman: All it says is as a guide to steering committees, the TEP suggested that assessment of fewer -- assessment of fewer than three EHR systems and 10 installations would not provide sufficient information to evaluate the feasibility of an eMeasure.

Keri Christensen: Right. Doesn't that mean that can't be endorsed, that's the logical conclusion if the ...

Howard Bregman: But it doesn't say that – it didn't say that it has to receive a certain score in feasibility. All it says is that you asked basically.

Michael Lieberman: Right. But I would – I would rephrase it.

Keri Christensen: Then, you could make that clear if you can ask and say that ...

Female: Right, right.

Keri Christensen: ... the answer is no. It's not the interval.

Female: Keri, I was basically recommending the steering committee – it's recommending to the steering committee, if you don't have – if you don't meet this threshold, then you can't say you really evaluated feasibility which is a problem.

Female: Yes.

Female: And I would also – I'm sorry, I'm just want to – it's a little bit out there. If you think about some specialty measures where the EHR market maybe smaller and perhaps there are two vendors that cover 90 percent of the market, isn't that physician's market coverage of those EHR vendors. This has the feasibility of a particular measure. And I think this question will come up over and over again, depending on the measure and the context that is in the area that it's focusing on.

Michael Lieberman: Yes. And so, then I would, as opposed to saying, you know, that's not sufficient or insufficient in our draft, I would just make it a more, you know, that a recommendation that you'd test with more than one EHR vendor and in multiple installations. But that doesn't amend, as Rute pointed out earlier that we make it very transparent as to how the testing was done.

So, if you do have, you know, very general measure that does apply to a wide, you know, wide variety of EHR vendors and you've only tested in two vendors, then that will be an issue and can be, you know, we will be obvious that's occurred and the question is why did you only do with two vendors as

opposed to especially measure where there are only two vendors and you've used them.

Female: This is how – I think that's fine. I just want to follow up on Keri's (inaudible). I think it's slightly different, and Keri, correct me if I'm wrong. I think part of what you're saying is you have a real challenge of getting vendors to even play and doing the testing, because they're not yet put into their systems.

So, I guess, I mean, this is a fundamental issue, I guess, for Kevin and Debbie and others on the phone about the ability for those in the midst of measure development to be able to partner with vendors to just really kind of check out if that (inaudible) could work.

Female: Yes.

Male: Yes, it did.

Female: (Inaudible) comments on both the vendor side and the HHS side.

Kevin Larsen: This is Kevin:. I'll defer to Keri and others, the more robust testing that we do, and at least in the current framework that traumatically, we increase the cost and the time to market. We typically have had challenges in not getting enough input from enough diversion of group because as we try to get these questions answered, they're challenging in complex and often require kind of high-level people. And those high-level people typically don't have the time or haven't prioritized a time like we've done in the business model to figure out how we get that input from those people.

Keri Christensen: Yes. And this is Keri. We're working as part of the outcome of the lean event to try to put together what the model would be to have a standing test pod to be able to engage folks on a regular basis for measures that are, you know, being funded to your government programs at least, which would really help that, but we don't quite have that figured out yet, it's in progress.

Kevin Larsen: And you typically need to know a fair bit of the details of how the measure works. So, if we would go to a site, and you know, medium level person at a

vendor and say, could you do principal diagnosis, they'd say, sure we do that all the time, we do it in claims. Once you really dive into the context and the intent of the measure, that's when you start to get to some of the questions and how it's raised. And we absolutely want to get there, but it needs a level of sophistication that is expensive and time consuming

We are exploring possibilities of using large data sets, have already aggregated this data to do a more quantitative numeric approach, but that requires somebody that has mapped data semantically across a large number of different vendors and different installations. And so, there are only a few such kinds of groups in the country.

Female: OK. Helpful.

Male: Can we move the number to two? I mean, I'm sympathetic, although objections that were – would two make sense, or you said the word multiple?

Male: Yes, I'd like the use of the word multiple better.

Rute Martins: I do too because again we're not saying that who's going to staff either.

Male: Right.

Kevin Larsen: Yes. And also, I don't know that we have to even mention the number three (inaudible).

Female: Yes, I agree.

Michael Lieberman: Yes. OK. Let's move on because we – I think, we're going to typically get into all the rest of these. The scorecard assumes that all data elements must be collected through some sort of data entry - including check boxes. Very often the EHR system is aware of clinician activities and can provide indications of these activities automatically.

I mean, I think that the comment is correct, except that I don't think we're really talking about doing at your check box that we are looking at trying to grab data that's already there in EHR.

Female: Right. And I would assume that's a comment regards for instance, time stamps related to certain activities. And then, I would certainly assume that those time stamps would be part of the evaluation in the scorecard and not just simply something that is explicitly introduced by a Humana negotiator.

Michael Lieberman: Right.

Female: We could maybe make that more clear where appropriate referring to data elements and amount of data about the data elements that just might make it more clear, that it's not just something entered into an EHR.

Michael Lieberman: OK. All right. The next one is, provide greater clarification on the definitions of current and future feasibility ratings. For instance, is current feasibility limited to the capacities of EHR systems at present, or does it include changes that could be made to the EHR within a relatively short time by purchasing existing EHR modules or through new certified EHR technology requirements?

Furthermore, we recommend that NQF provide alternative definitions of the scores for assessing the future feasibility of the measure. For instance, the definition of a score of one on data availability could read: The ability to collect this data element is not expected to be required for certified EHRs in the next 3–5 years, nor is it likely to be widely collected within that time frame.

I agree with the spirit of this comment, and I don't know, you know, we kind of gave it our best shot.

Kevin Larsen: This is Kevin:. I think there's – we have some commitment to an evolving standard here. We know that what we have isn't what we want and we know that what we want, we can't achieve instantly. So, there are some commitment to how do we – how do we commit to making this more robust as we discover pathways to do that.

Michael Lieberman: OK. And I thought we kind of do that in the – don't we make reference to certify EHR technology in the scoring system?

OK. Let's move on to the next comment. Clarify how the market share for the assessed EHR systems should be determined to inform documentation of the feasibility assessment process. Although we agree that this information is valuable, we have found it can be difficult to ascertain, particularly as EHR vendors are unable to provide such statistics.

I don't – (Dee), does anybody know where this comment originates from in the draft document?

(Dee): Yes. At the top of the scorecard, there is a question about providing your assessment on – I'm trying to see where it is exactly – about the market share. That was the discussion that the group had fairly significantly. Right above the scorecard, it says, the feasibility assessment should provide transparent detail on how the assessment was made, including the number of EHR systems assessed and the percentage of EHR market coverage for that clinical domain.

Michael Lieberman: Yes.

(Dee): I agree with all that.

Male: Why don't we add if available.

(Dee): Yes.

Michael Lieberman: Yes. Right. You know, your best testament or something like that.

Rute Martins: I agree.

Male: Shall we say if available, not just cover, you know, if they have no idea that I really have a way to reliably get that number. You want them to guess if they – no, they still might – they still might estimate it, but at least it wouldn't be required if you say if available.

Female: OK.

Michael Lieberman: Yes. Or even just put in the approximate percentage of EHR market coverage for that clinical domain. OK.

Rute Martins: Shall we go with both approximate and if available.

Female: OK.

Michael Lieberman: Fair enough. Yes. OK. Next, under data availability, The assessment should not limit the data availability to only structured data, because free-text data can be useful as long as the EHR vendors have capabilities to transform free-text into query able data.

Well, I mean, I think once they've transformed it from free text into acquirable data, if not structured.

Male: Agree.

Michael Lieberman: OK. So, I think I would kind of reject that comment or just say that once it's – exactly once it's been transformed, it's now structured. OK. The work would be easier and more certain if all developers were asked to pick their data elements from a fixed universe, and if the CMS/ONC definitions for what had to be carried in the EHR were concordant with the part of the universe that was used in the eMeasures.

You know, I think that as Kevin pointed out, I mean, if we have universe of known measures and known data elements, that would be very useful, but still I doubt that we can constrain it to that and still be able to grow the ability to create noble measures.

Keri Christensen: This is Keri. Our measure (inaudible) struggled with it for 10 years, thinking about should we restrict ourselves to what's available or should we create the best measure to move clinical practice to where we want it to be and they consistently stick with the latter one. So, agree.

Michael Lieberman: Yes. And that's ...

Kevin Larsen: This is Kevin. I think that the goal is that the bulk of the measure, the behind the scenes part of the measure is established and known, and the small part, the five percent of the measure that one percent of the measure that is

stretching that care, that's the part that we're OK with requiring new data element.

So, if we build a measure on feasibility assessment, the whole rest of the measure should be easy to do. The feasibility assessment part we know is new, but we'll ask for that, but the rest of it shouldn't have a high burden because that's all the rest should be feasible.

Michael Lieberman: I think that's a great point, and I don't know how exactly to put that, but you know, there should be a premium put on using available data elements and that known data elements, known feasible data elements, and maybe that's part of the assessment where, you know, is this one that's already been evaluated and used versus if it's new. And if it's new it's going to great greater scrutiny.

So, these people, you know, first of all, if they choose us and it's already used, they don't have to do the feasibility assessment for that particular data element, which should be a big help. And then also, you know, they would – you would get the experience with that, you wouldn't have to – you would see that it has (inaudible).

Kevin Larsen: Yes. I kind of think of it the same way I build order sets which had a highly variable portion and a fixed portion just like an antibody. And so, ideally, there is a whole lot of fixed portion and a little tiny bit of highly variable portion.

Female: And I think that speaks to the concept of undo burden. If at all possible, we're using what already exist and make sense, so that no one's creating additional data elements for no good reason. It's all about the good reason for the new data element. What I would also say, though, is that it's kind of reductive to think about the ONC-CMS definitions in meaningful use and saying that that's the base for all eMeasures to be.

Zahid Butt: So, this is Zahid. Just I think this sort of those raise that issue that we discussed quite at length in terms of the cause versus benefit, also in element. And I'm just still thinking that, that was a very good concept when we discussed it. Is there some way to incorporate that framework within this, that

OK so this is infeasible or difficult to get element, but the value of it is such high in terms of it's benefit as it's worth the cost.

So, perhaps, I don't know whether there's still some ability to incorporate that concept. I know, Mike had over each had brought that issue up, and we have quite an interesting discussion around that.

Michael Lieberman: Yes. And Zahid, I think that that would show up as a, you know, a data element that would have a low feasibility score, but it would be one that you would look at and say yes, it's critical to this measure and we feel that this measure is important, so we still want to use this measure.

Zahid Butt: Right. But exactly what you just said, did reserve incorporate that concept somewhere in the report, perhaps clarifying it ...

Reva Winkler: Maybe not.

Zahid Butt: ... in those terms that's just expressed.

Reva Winkler: Zahid, this is Reva. I don't know that it's that explicit, but we can certainly revise it to be more explicit.

Zahid Butt: Because that might sort of address this comment in some fashion.

Michael Lieberman: OK. Let's see, where are we. Strongly suggest asking eMeasure developers to note in what part or what field in the EHR the data would be found, not just whether it would be present, e.g., problem list, discharge diagnosis, or lab results, radiology, reports, orders.

So, that's an interesting one. And I guess, do we have that – is that part of the – is that part of the QDM? Sometimes where it is going to be found.

Christopher Millet: Hi, this is Chris Millet from NQF. Can I chime in on this one? I think this is a really ...

Michael Lieberman: Yes.

Christopher Millet: Particularly in the way they frame this, and I'll address the QDM message after. I think this is interesting question because they're asking what field in the EHR, and you know, I think mostly here we're referring to EHR system and as we all know and we've seen in the previous comments, different EHR systems don't have the same self field necessarily. They vary from system to system and from installations and installation.

And I think an alternative way to look at this where we can ask that question is by looking at a different electronic data source when we say the EHR. And as an example, I'm thinking of electronic data source that does have a standardized set of fields that all EHR system should be able to produce like the standardized documents included and consolidated (CCA).

Those are just examples of structured documents, but the standardized set of fields that, you know, technically all certified EHR should be able to produce, and if they all have those same set of fields, then we can start referring to those in measures and then testing those fields for, or testing those fields for feasibility and even other criteria like reliability and validity.

But I know that's quite different in the way we've been thinking about a lot of these comments, but I think that's an alternative that we should really start to consider.

Ginny Meadows: This is Ginny Meadows, and I apologize for being so late to the call, but just a couple of comments to that. I'm trying to understand the difference between what you're describing as a field versus what are you looking at as required data element sort of process and why it would be important to know exactly where that field was located if we're using the QRDA to actually export that data to a quality measurement module.

Michael Lieberman: All right. I think – I think that makes sense. I think field and the element are used quite interchangeably.

Ginny Meadows: Yes. To me, it's kind of a different connotation, but maybe it's just me.

Saul Kravitz: This is Saul Kravitz. I just want to comment on that, and like Chris said, you know, the fact that we have a format that says we know where you – we know

how to represent an assessment in a file, really doesn't shed much light on where precisely in McKesson system or Epic system or (inaudible) system where and how that particular element is represented or whether it's even representable, which just says after that class of data, we've defined the format. So, I think they're completely different things we're talking about.

Ginny Meadows: Yes. I think so too. And I guess my question would be kind of to what you just have told was why is it important to know exactly where that's located.

Kevin Larsen: So, this is Kevin:. And a couple of examples that have come up, and I don't actually have a particular opinion about this, but times I've heard this reference where certain ECQMs are, we want to identify this by the problems on the problem list. We want to identify this based on the discharge medications or we want to see this in discharge medications.

So, those imply in some ways a data element, but in most systems I'm aware, implies a very specific part of an EHR. In fact, sometimes that the part of an EHR we've called out that has to exist to be a certified technology.

Ginny Meadows: And that actually makes sense, Kevin:, but I just wonder if that would be more in the attributes of the data element itself versus what we are calling a field.

Howard Bregman: This is Howard from Epic. Let me tell you how I frame this comment. If a measure developer wanted to send me a measure for assessment, I would want to know everything they're thinking. I would not want to hide it. I would not want them to hide it.

So, if they happen to be thinking that this data is going to come from the problem list, telling us that would help us a lot to respond to the assessment request. And furthermore, if I were to fill up the form that appears on page 10 and 11 of our draft, and all I did was say 312 and I send it back to the measure developer, that would not be very helpful. They would want to know what I'm thinking, why do I think it gets a three or a one.

So, I think that we should be encouraging, if we don't already, both the requestor of the assessment and the provider of the assessment to provide all

kinds and whatever relevant information they have to help the other party understand where they're coming from. And this is just an extension of that.

Female: That's completely agreeable. You just say how much information is helpful. So, if that's really more of what we're talking about, then I absolutely do agree.

Howard Bregman: Doesn't the QDM that we have a lot of that information in there. I mean, it's like medication prescribed, medications ...

Kevin Larsen: Well, it does in a sense, but for example, the example diagnosis that I gave. You'd find the diagnosis in different places. If you just say a diagnosis during an admission and you're thinking coding statement, billing statement, and I'm thinking, well it's probably going to be on the problem list, but maybe the problem list isn't going to have it or there's other issues to the problem list, that kind of communication.

Even though that's a standard element, it's in the QDM, it has a definition and attributes. It has variability practically.

Rute Martins: Yes. And I would add to that – this is Rute, that I really think this is an unresolved issue, the issue of meaning versus source and how a source can assign additional meaning to certain data element. My understanding from QDM is we, to some extent, may be able to specify the sources that are more of the HQMF pattern do suggest specific lead or context for this information as far as I can tell from the way that HQMF release those are intended to be suggestive only and not necessarily mandatory.

And I think we're still testing the water in what should be mandatory or what shouldn't that can be effectively communicated. So, I think that's really important to have that sense of context when we're evaluating the data element's context input.

Howard Bregman: So, I would suggest that on page 11 of our draft, after we have data element feasibility score, we have another item that says comments on how the score was obtained or how the score was generated, was the score is based on.

Female: Howard, actually, if you look at the appendix where (Aldo) and Ginny gave us examples of the sort of applied scorecard to some existing data elements, we actually added that column. So, I agree that the one on page 10 needs to include that as well.

Howard Bregman: For me, I think we can ask for, you know, comments, you know, sources, source in the EMR or something of that nature for when there's a data element, that if a data is that measure developer has an idea of where it's coming from, they can list that. But is it, do we want – I mean, would that just be freeform or will we want to link it to some other standard like CDA or not.

Rute Martins: No. I guess I struggle – this is Rute. I struggle with the – in the CDA. Let's say that we're looking for gestational age, and then the CDA document, you can put gestational age, but really what you're looking for in the measure is the gestational age, courses to delivery. If the mom was admitted and that confines to (inaudible) two months before delivery (inaudible) that admission is of no usefulness to the measure. So, that's the kind of new answers. And they can be pretty data element specific or measure specific, so I would go with freeform for now.

Michael Lieberman: OK. So, the next one, for a score of 3, the guidance states “Data element is routinely collected as part of care process and exists in the majority of EHRs.” For measures not achieving a score of 3, it will not be determinable if the low score is due to structural issues with the EHR or workflow issues such that clinicians do not document in the structured data fields even though the capability exists. Since the latter concern (clinicians not documenting) is addressed by the workflow question, I suggest that the data availability ratings be revised to clearly represent only the issue of the EHR's structural capability to capture the required data elements.

Male: I actually think that's what we met, right? That data availability, it's a structural capability and that workflow says. Well, there's a little bit of complexion of this thing (inaudible).

Michael Lieberman: Yes.

Rute Martins: I guess, my struggle is, and I truly believe that we should separate technical capability from actual user behavior and workflow issues. Under the data availability, though, I thought that we had talked about the idea of how often is the field populated, and the answer to that question, the answer is not often. It may be a workflow issue or it may just be that the clinician or particular clinician are boycotting the EHR system, I don't know. But I would be happy to see a clear separation of the technical versus user input.

Kevin Larsen: This is Kevin. Although I agree with the separation, I think they're both hugely important. So, as I think about what field is feasible to an end user, what field is usable to them as a combination of can my tool do it, and is it something that has been made a natural part of my work.

So, both the technical and the use, I think, are very important. And in some ways, I think, you should get the highest, you know, weight.

Male: If we do ...

Zahid Butt: I think that in the description of the data availability, we implied that it's really a technical capability. It's further down, but it gets a little bit more confusing in the 3-2-1. And so, I think, since we have workflow as a separate element, if we including workflow here as well, then inadvertently, workflow will be counted twice.

Male: Maybe I'll say that I think part of the history is that we were a little bit hedging on whether workflow was even going to be included. Yes, it's included, then I think we can ...

Zahid Butt: That's what I was saying, that since workflow is specifically called out as a third component that perhaps we should revise the 3-2-1 scoring to sort of align with what we say up top, which is the data readily available in structure format, i.e. it resides in fixed field within EHRs. So, that kind of ...

Kevin Larsen: Zahid, I would make the new ones different between workflow and what I'm of as use. So, I think workflow is a really important pathway to get to use, but my mind uses the outcome of a good tool, a good workflow and a good incentive. And what I really care about was, was this – is this populated, is it

used in this context which means it's easy for me to leverage for measurement because all those things have been lined up and happened.

Because you can have a great workflow, but if your tool doesn't match the workflow, it's still – it's still you don't get any data.

Rute Martins: (Inaudible) workflow in this context. Isn't this workflow within the tool?

Kevin Larsen: I think presumably, but we don't have a good way to represent that, right? So, we're looking at the – so, if I want to use problem list to identify patients in the hospital that have a certain condition to put them in the denominator of a measure.

I got a tool that has problem list, I have a workflow that describes what a problem list looks like, and as I talk to my colleagues, I think the best I've ever seen is a 50 percent population of that particular item. Most places that even focus really hard have trouble getting uptake of hospital-based problem list.

So, the question is what do we do with that 50 percent number. We have a data element, we have a workflow that is validated and used routinely across sites, but we still only have a 50 percent uptake. And how do we represent those three things?

Michael Lieberman: I think that they like to use a little bit of double counting of a workflow, but I actually – I think that's OK, and that's such a crucial part where we're really looking for around feasibility is question of is the information going to be available without putting burden on clinicians to collect itself. We're kind of having that twice, I'm fine with that.

All right, ready to move on. The next one is, the current availability of structured data is largely dependent on the quality and prevalence of vendor-developed templates. Well, I think that's true. A matter of its only vendor-developed templates, it could be developed by individual users as well. But that is – I don't think that really requires any changes. Everybody else OK?

Next, we're going to move on to data accuracy. Data accuracy conflates the concept of feasibility with the criterion of validity. As currently defined, it is unclear whether this item refers to the accuracy of the specifications (already captured in NQF scientific acceptability evaluation sub criteria 2b1) or the accuracy of the data element relative to what care actually was provided to the patient.

Furthermore, we believe the descriptions of ratings shown under data accuracy make broad assumptions on which sources should be considered most accurate. Although we agree that data accuracy should be considered early in the measure development process, defining it as part of feasibility likely will lead to confusion.

I think the silence is telling here. I think it's a complicated issue, and that there is a large overlap and do we, you know, how do we – how do we deal with that. I mean, one thing is that we can stay up front and if we haven't already done so, that we realize that there's an issue, though we think that looking at accuracy or kind of practice yearly on the process if useful, and so we want to leave it in the assessment.

Rute Martins: I agree. I don't – this is Rute. I don't think that we're trying to evaluate accuracy in the context of feasibility, but rather use whatever information we have about accuracy of a particular data element to decide whether there's a better way or to choose between two data elements if we have that ability, but it's not where we are assessing the accuracy of the measure or in the sense of the scientific acceptability criteria.

Male: I think it might have been like – I think that specific to the comment, I think we're not asking for accuracy of those specification, right? I think we're trying to get at whether the way the data is captured or within the EHR whether it will be accurate the way it's specified as to who captures this and where it gets captured.

Kevin Larsen: Yes. I think, for example, we would consider structured data coming from a lab system to be accurate, and where I say probably used to type in number attributes, we are unclear about the accuracy of undifferentiated vital sign

where it could be entered by host of people. That's certainly what we met. We probably haven't well delineated differences that it is not related to the specification.

Male: Right. So, at least that part of it we can address.

Male: I think the way the part is defined in our document, I think it's very clear, it gives examples. I think the definitions are rather clear, so I think it does currently explain our intents.

Michael Lieberman: Right. OK. The next one is, accuracy is a function of the measure intent. In determining feasibility, it will be important to determine if the source required by the measure developer is available in the EHR. I think that's what we're getting at. I mean, in terms of data availability as well.

Zahid Butt: No, it's availability – they're comment is not availability. We met actually the validity and reliabilities of that, but what is typically (inaudible).

Michael Lieberman: OK. Any other comments on that one? Then, authoritative source is judgmental. Given the context of a data element with all of its associated attributes (metadata), the question should ask about A, Does the information available in your system reliably meet the definition of the data element purpose and definition, B, Does the information available in your system have the potential for meeting reliability based on the definition and purpose of the data element, C, even with workflow changes, the information in my system cannot be expected to reliably provide data as required by the data element.

I don't know what's that. I think that's – but I think that's for getting our data accuracy. It's not what they're – I think what this comment is getting at is kind of mismatch data somewhere, so you might be able to find something in your system that looks like the information being asked for, but it might not really be the same thing, you know, based on the definition improves to the data and that gets – and that, I mean, I think when we talk about it being correct, I think that's what we were getting at.

Male: I do agree that we already covered this in our existing categories.

Michael Lieberman: OK. Accuracy will be most influenced by currency, something that might have been true in the past is not necessarily true at present. I think that's true, but that would – it seems to me that that would be addressed by the definition of the data on that, you know, at what point in time are you looking for that information.

Female: Agree. And that's just one aspect of accuracy. It's not just the time relevance.

Michael Lieberman: All right. OK. Next, under data standards, Mapping eMeasures to the nationally-accepted vocabulary standards is good. But if one has to create a new vocabulary code (as mentioned above) to capture something that is not ordinarily captured (as mentioned above) that should be a signal that this is a burdensome measure or to look for other EHR data that would be a sufficient surrogate.

I think we've kind of addressed this. I think that's true, but there are times when we are going to want a new data element. Other times, you might – you might look for already existing. Any other comments on that?

Rute Martins: Yes. This is Rute. I actually disagree because it's actually specific address – it's specifically addressing the vocabulary piece.

Michael Lieberman: Ah-huh.

Rute Martins: That assumes that vocabularies are perfect, and they're not.

Michael Lieberman: Right, right.

Kevin Larsen: So, we've run into this a number of times in the meaningful use two measures where we're looking for very specific concepts in validated serving instruments, and those concepts don't already exist and it's no more their link. So, we can't just use a different concept and think it has the same meaning as a validated serving instrument.

Michael Lieberman: OK. So, it's actually – so, we're disagreeing with this comment in that, and I think that they're already – what they're getting at is oftentimes, we want to use what's already available, and if you can, that's good. But what we're

saying is that, I mean that there are many occurrences where we do need a new vocabulary element, and it's not – and it's not necessarily a new burden.

Debbie Krauss: And this is Debbie. I agree with you, but I think we did adjust part of this earlier on when we mentioned about involving terminology experts early on in the discussion on the feasibility of the data element, if there's a code already exist, that could be better served or if we need a new code. So, we did talk about that in recommendation one.

Michael Lieberman: OK. And this comment I think is one of my favorites. The QDM are like the islands in the film Avatar that float in the sky, disconnected from anything. The eMeasures are not grounded in data that would be found in an early medical record. So specifying “where” the data element “is” in the QDM does not help except for those QDM elements that are also specified in the ONC MU2 rules.

I don't know quite what to do with this comment, though.

Male: Can we give out a prize to the most creative comment?

Female: I love that. And I suspect all of you can guess who said it as well. It's (inaudible).

Male: So, I think the response to that comment is that we have – we have a wide variance in what exist in the current world. And so, because of that wide variance, we need to come up with some common core and we're committing ourselves to empirically improving that and wrap in improvement processes. But there is no way to pick things that occur everywhere because we didn't create systems that are clones of each other.

Juliet Rubini: And this is Juliet with NQF just to also state that this – well, creative, I agree. Comment may also have been addressed in our previous discussion regarding EHR context.
So, it maybe solved with that discussion.

Michael Lieberman: Yes. Although I wonder if – I mean if everybody's thought (inaudible) about this because if the idea that he doesn't feel that QDM is a reasonable,

you know, this is the method that they're using to specify data elements. I mean, is that what those foundation has commented or that we need to define more QCM elements in ONC MU2 rules or what?

Juliet Rubini: I think concerns go even go higher than that, and he really is questioning a lot of the measurement paradigm in the greater respect of sort of even when we're starting to consider concepts for measurement or we're really considering the right concepts.

Male: And I think when he says disconnected from anything, he means disconnected from reality. He's objecting to the QDM fundamentally, I believe.

Male: Yes. I think he things it's a conceptual model without grounding an empirical evidence.

Michael Lieberman: OK. So, next comment, clarify whether the term data standards refers to the measure specifications themselves or the regular use of the specified terminologies within EHRs. I think it's regular use of specified terminologies within EHRs.

Kevin Larsen: So, this is Kevin again. This is my personal opinion. I think those two things can't be disconnected. If I again use my analogy of the antibody, there's some fixed portion which is the regular use of standardized terminology in EHR, and then there's the highly-variable portion by which the ECQM will cost used some small amount, ideally, of new data on it.

Female: Well, I guess, this is my question, what we mean by data standard is actually the vocabulary. So, it's not whether its normal link and so forth, correct?

Michael Lieberman: Yes. That will be my understanding.

Male: Yes.

Michael Lieberman: OK. The next one, (inaudible), Define the term "nationally accepted," – there is great variation among providers in the types of coding systems adopted. For instance, although sites may be using NDCs for billing, some use 11-digit codes while others use 10 digits.

So, is there, you know, either a standard set of coding systems that we can refer to? Are these specified in meaningful use?

Female: I think we've all – as eMeasure's evolved now, I would encourage other measure developers to jump in, but I think what we've been following is I believe the high tech recommendations or the – I think it was the vocabulary task force and the quality work groups that provided trunk recommendations on the vocabularies and entice them to actual QDM category. So, I think that too is a nationally accepted standard right now.

Michael Lieberman: Great. Yes. We should spell that out specifically then.

Female: And also, there's disagree though.

Ginny Meadows: No, I would agree on that. This is Ginny. And we definitely, I know that we specifically look to what's been unspecified in the certification standards. So, in this example, we would always expect to use RxNorm, even though we may for prescribing purposes use, and it's because we mentioned RxNorm because we know that that's the specified standard.

Michael Lieberman: Yes. OK. Next, under workflow ...

Female: Mike.

Michael Lieberman: Yes, go ahead.

Reva Winkler: Yes, this is Reva. Somebody else want to say something first?

Male: I do want to just say something before we leave data standards.

Reva Winkler: Go ahead.

Male: I think that we, in our three point definition, we use the term always, and I think we should back off from that. We should never say always as a criterion because there'll be – it's too easy to just say it's not three points because it's not always. So, we should say something like almost always or a large majority of the time.

Michael Lieberman: OK.

Female: And I guess, I have one additional question, I'm sorry. I don't want to derail this. When you say regular use of the specified terminologies within EHRs, my understanding is that there is a lot, there are lot of concepts that are specified in local terminology or terminologies other than the nationally accepted, I should say, terminologies. But there are mappings to those terminologies that are current and actively maintained.

So, I guess my question is what we mean by regular use is it at the interface level or is it at the back end or both. I think we should clarify that.

Male: Where do we use the term regular use, I just see we say is the data element coded.

Female: So, I guess my question is where would data element be coded if the data element coded in the standard vocabulary is that within the EHR interface or is that at the back end.

Female: You mean the user interface?

Female: Or is it outside of EHR.

Male: It should be – well, I think the understanding should be that it's anywhere. There's no technical limitation to if it's coded it's linked in some way, whatever way that may be.

Female: And that's what I think we should clarify, because I think some may perceive the question to be at the interface level.

Zahid Butt: So, this is Zahid. I think that's an interesting point because perhaps coded is too strong a word, which implies that it has to be natively coded at the point of capture, perhaps maybe would be that the data element is available within EHR using nationally accepted terminology standard. Would that make it a little bit more encompassing whether it's an interface or a map or a primary capture?

Female: Or just spelling that out, Zahid, I think it would be helpful just to give example vocabulary available at the point of care or local vocabulary use with the capability of mapping to standard nationally accepted vocabularies. I think that would ...

Zahid Butt: I think what I'd like to do is to stay away from the implication that the local vocabulary could be used directly in eMeasure specification. So, I think, we want obviously at the specification level, we want national terminology used. And within the EHR, there will be different ways in which the national terminology will be captured. Sometimes it will be captured as a, you know, coded element at the point of capture, and sometimes it will be mapped to local terminologies.

So, I think if we say that is not coded necessarily, perhaps that implies that it has to be captured that way, and I may be just reading too much into it, but I think that's kind of where this question sort of is getting at. So, if we can somehow say that as long as it's available within the EHR for use and it complies with nationally accepted terminology standard, I think that should cover it.

Michael Lieberman: OK. I think now, Reva, we have five minutes left.

Reva Winkler: Right. That's what I was just going to say. I just want to get some sense from the panel because we do want to take the opportunity to have some public comment. But you've had a wonderful discussion on the topics you were able to cover, and I would not want to short change that for the other comments. Do we need another opportunity for this group to get together to review the remaining comments.

Female: Yes, please.

Male: Yes.

Reva Winkler: Yes.

Male: Yes.

Reva Winkler: That's what seem to be the reasonable thing. So, we'll try and work with schedules and see if we can get something together to map very soon so that we can continue doing it because I do think there is incredible value in having the opportunity to have this robust discussion around all these comments. So, that's – if everybody's in agreement, we'll plan on doing that.

Female: Reva, just a reminder that most of us will be out of the office next week.

Reva Winkler: Right.

Female: So, maybe with the following weeks.

Reva Winkler: Yes. OK. Thank you, (Carrie) for a good reminder because you're right, we are. So, we'll do that, but before we close out, we did have a fair number of audience members join us. And Mike, we do want to have the opportunity for public comment.

Michael Lieberman: OK. Are those lines all open?

Reva Winkler: I believe they are.

Michael Lieberman: OK.

Reva Winkler: (Nathalie), operator, are all those lines open?

Operator: Yes, ma'am, they are.

Reva Winkler: Great.

Michael Lieberman: All right. So, I'll take these – are there any comments from the public?

Female: No.

Michael Lieberman: OK.

Reva Winkler: Well, for the folks that did join us on the call today, when we reschedule another conference call, we'll be sure and post it and let everybody know when that next call will be so that you are all aware.

Female: Thank you.

Michael Lieberman: OK. So, it looks like we made it through. We'll pick up next time with workflow. For people in our – Reva, can you just a minute and kind of keep track of what we've commented and what we haven't. Would it be – is it possible to push this PDF document back into an Excel spreadsheet so that people can usually comment on the ones that they haven't had a chance to yet, if that would be helpful.

Reva Winkler: Well, I think that would be. One of the things we'll be doing is taking this conversation and beginning to populate the response column in the Excel spreadsheet. And I think we're going to have time to be able to do that and get it back to you all.

Michael Lieberman: OK, that will be good. Now, we can just see what's still open on that, so people don't have – if they're not able to make the next call, they can still make comments.

Reva Winkler: Right. I think that would be helpful.

Michael Lieberman: All right. Anything else you want to add, Reva?

Reva Winkler: I don't think so. Thank you to everybody on the panel and for all of you're very thoughtful comments. I think that this feedback, and now your consideration and thoughtful discussion is going to make this report even better and stronger. And so, that's why I don't want to cut off and make an incomplete discussion. So, thank you for willing to come back and take another go at the rest of them.

Michael Lieberman: All right. Thank you all for joining today, and you'll be hearing from us.

Female: Great. Thanks, Mike.

Michael Lieberman: Bye-bye.

Female: Thanks. Bye-bye.

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

END