NATIONAL QUALITY FORUM

Moderator: Sheila Crawford
April 17, 2014
1:00 p.m. ET

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

Angela Franklin: Hello and welcome to our first workgroup call of the musculoskeletal consensus development project. This is Angela Franklin speaking and I am the senior director for this project.

I have with me Ann Phillips, our project analyst. And we will today be reviewing three measures. Measure number 662, 0514 and 0052.

Before we go much further though, we'd like to do a quick roll call, and I'll turn it over to Ann Phillips.

Ann Phillips: I'm going to start with the workgroup. First, is everybody in on the webinar too?

Angela Franklin: Yes.

Ann Phillips: Yes? OK. Great.

Thiru Annaswamy?

Thiru Annaswamy: Yes, I'm here.

Ann Phillips: Carlos Bagley?

Craig Butler?

Craig Butler: I'm here.

Ann Phillips: Zoher Ghogawala?

Wendy Marinkovich?

Wendy Marinkovich: I'm on.

Ann Phillips: Catherine Roberts?

*Catherine Roberts: I am here but I apologize. I will need to step out for about 10 minutes and a half an hour. So if I go silent, don't worry I'll be back.

Ann Phillips: OK.

Sean Bryant?

Sean Bryant: Here.

Ann Phillips: Katherine Gray?

*Katherine Gray: Here.

Ann Phillips: Christopher Visco?

Christopher Visco: I'm here.

***Please note that quotations attributed to Catherine Roberts may have been by Katherine Gray, and vice versa, due to the challenges of distinguishing between their voices on the teleconference. These can be corrected upon request if the need arises.**

Ann Phillips: Is there anybody else who's not in the workgroup from the committee who's on the call?

Nobody?

Marcie Harris Hayes: Marcie Harris Hayes is listening in.

Ann Phillips: I'm sorry, who?

Marcie Harris Hayes: Marcie Harris Hayes is listening in.

Ann Phillips: Hey Marcie. And then there should be some developers. Some measure developers on the call. Can you say who you are and which organization?

Wanda Johnson: It would be Wanda Johnson, Casey Thompson from OFMQ.

Jenna Williams-Bader: Hi. This is Jenna Williams-Bader from NCQA.

Ann Phillips: And we have the other CMS developer on?

OK. And Juanita, are you doing both 0514 and 0662?

Female: So it's Wanda.

Ann Phillips: Wanda? OK, all right. So we go then to 0662.

Angela Franklin: OK, great. So welcome again to our first workgroup call. And I'd like to quickly before we get started, go over the purpose of today's call which is to have an in-depth discussion about the measures as a small group ahead of our full steering committee meeting in May and really bring to light any specific issues and questions about each measure that committee members may have in advance of our discussion and voting on the measures in May.

Please keep in mind, we are not voting today. We're merely discussing and focusing issues and we're very fortunate to have measure developers on the call today who will have available – to answer any questions that the steering committee may have at this time. And also developers will have the

opportunity if they choose to respond to any specific requests regarding the measures that the committee members may have.

I just like to remind our committee members that we really have to review the measures as they stand before us. There have been some questions about potential tweaks to measures and to the extent that those tweaks are, can be done, evidence-based or it's not a fundamental change in the measure. It's something that of course a developer can entertain but we're really reviewing the measures as they stand before us.

Please also use this opportunity to answer any questions you might have about the measure evaluation process or the criteria and certainly if we can't answer out right, we'll follow up with you after the call.

We're also piloting several new aids for the steering committee during this project including the committee guidebook, and the staff measure, the staff – measure evaluation algorithms and send it to help guide you through discussion and thinking about the criteria. And any feedback you have about these two items will be much appreciated.

So with that, let's see, I would like to turn it over to our workgroup members. And I believe our first measure that is our review is measure number 0662, Median Time to Pain Management for Long Bone Fracture and the measures to work for this one is CMS and our discussant for this one, I believe – I didn't pull it up, is Dr. Visco, let’s see Sean Bryant, Wendy Marinkovich and Christopher Visco.

And again, I think we have Sean Bryan as first if you could start off. And then again, if you could tell your thoughts about the measure going through the criteria then if you could give us any sense of what group comments were overall about the measure and then throw it open for discussion by the committee. And again, we have developers on the call if there are questions directed for the (two) developers.

So, Sean, are you there?

Sean Bryant: I am here.

Angela Franklin: OK, great. Thanks. So if you would mind starting us off with measure 662.

Sean Bryant: OK. So the developer rationale is that this long bone fracture undertreated in ERs, pain management has room for improvement. The patients with these fractures continue to lack of administration of pain medication. When standards are implemented, breaks for pain improve, disparities exist between administration medication from minorities children and several references are given to support all this, the numerator statement time from ER arrival to initial oral, intranasal, parental pain medicine management administration for ER patients with diagnosis of a long bone fracture, I had a question about how parental is defined. It's not clearly specified here.

Is it IM, IV, intra-arterial, intra-osseous and intrathecal? Or is it just IM and IV?

Female: Casey, did you hear that question?

Casey Thompson: Oh yes, I did.

Female: I believe it's intra-osseous, do we have a list of what is considered parental in appendix?

Casey Thompson: We've – yes, it's in appendix at the – here in the glossary. Let me read it to you. One second. It says parenteral, not through the alimentary canal but rather by injection through some other root as subcutaneous, intramuscular, intraorbital, intracapsular, intraspinal, intrasternal, intravenous, et cetera.

Sean Bryant: OK. So any injection. All right. This is a continuous variable so there's no denominator for statement or I mean or exclusions. The measure type, they defined it as efficiency measure. But I think in conversation, Wendy and I had call with NQF staff yesterday. And there seems to be agreement but this is not really an efficiency measure. To me, it seems be more of a process and care measure. But I will open it up for discussion on that.

Angela Franklin: So, this is – this is Angela at NQF. Yes, according to our definitions and for purposes of our review of a measure, evidence and testing. This is a …

Thiru Annaswamy: This Thiru Annaswamy. Can I interject a little bit. The NQF staffs' comments are a little low volume. You can either increase your volume or put the speaker phone closer to you. I would appreciate it.

Angela Franklin: Sure, sorry about that. The issue is that according to NQF's definition of an efficiency measure that is the measure that links both cost and resource. And this isn't – these measures do not fall into that category. So it's definitely is a process measure for the purposes of our review today.

Sean Bryant: OK. All right. And the – this was originally endorsed back in January 2011. So this is a maintenance measure. There were several comments made, many of which I agree with. With regards to evidence to support the measure of focus, the comment was made about just not being an outcome measure.

Also, there was a comment that most of the literature, it's simply that the literature documents that pain is not adequately treated but there's no evidence put forward to support for that (supposition) that measuring and reporting, the time gap between arrival and pain med administration, it actually improves quality of care.

The commenter like my self believes that less time is better but really the evidence is in the form of retrospect to reviews. So this is thought to be in a little bit weaker evidence.

There's also a comment that it was moderately provided in the literature. No systematic review and no grading available. A good number of studies, eight of moderate equality and these comments are recommended passing on this criterion.

And another person commented that the literature evidence is somewhat weak in clearly defining a particular time – what particular timeframe improves care. There are some reasonable – reasonably implied things here. But also they brought up a good point that there are lots of different long bones in the body. And there are different severities of long bone fractures.

So, you know, a small relatively mild avulsion fracture would have, you know, one lower level of pain but, you know, a mid shaft femur fracture or to be a fracture would create an enormous amount of pain. So it's really hard to put all of these different long bone fractures in the same bucket and get meaningful data. So, I'll open that for discussion.

Craig Butler: This is Craig Butler. Can I interject with it? It's kind of fundamental question. And I'm new to this obviously. But they're asking us to evaluate the evidence. And I understand with respect to long bones and even in MRIs that there's evidence that there's a delay in a sort low inadequate treatment of pain for long bone fracture with MRI there's way too many done, say for a – in a (substantial) reason.

Are we asked in terms of the evidence to evaluate whether the evidence supports that or whether or not the measure they're suggesting it equates to measuring – equates to measuring quality to validate measure quality leading those deficits if you will. And I think that's the fundamental question that I had some difficulty asking, and I'll kind of go ahead and at the time, own the first comment in the section on evidence because that's what it speak to, I think.

Angela Franklin: So hopefully this answers your question. We're definitely wanting to see the quality of the evidence presented that this measure is a valid measure of quality. And we want to make sure that the evidence presented supports the focus of the measure which is the improvement of treatment of long bone pain.

Craig Butler: So, you're saying both?

Angela Franklin: Yes.

Craig Butler: OK. So …

Wanda Johnson: This is Wanda from OFMQ. One of the reasons we call this inefficiency measure is because there are no clinical guidelines that support or give a specific timeframe. And so, we did call it an efficiency measure simply

because we don't have those kinds of guidelines that we can rely on for evidence.

Craig Butler: And you mentioned that there were some criteria for categorizing these measures as well whether they're efficiency or process. Is that, somewhere, is that in the guidebook and I missed it or is there some (inaudible)?

Angela Franklin: It's in the guide book. And it also in our – is basically in our guidebook, and also in our measure submission form. There's note included which could spell out the differences between health outcome intermediate clinical outcome or process measures, structure measure and efficiency measure.

Craig Butler: OK, I'll have to review that again. Thank you.

Male: Can I offer comment to (John Sentira)? I'm glad Craig's brought that, pointed out, we faced the very same thing with one of the other measures on functional status and rheumatoid arthritis whether we were looking at the evidence about the patient report and outcome or the evidence about the actual, just their process improves clinical outcomes. So, I'm glad it was expressed.

Angela Franklin: Thank you.

Sean Bryant: I'll move on to 1b performance gap. Again if your comments were submitted, one that I really agree with, pointed out that there's been very marginal change or improvement over the last five quarters that this data has been reported which brings into question in my mind, is this measure really driving improvements in quality or do we need to be thinking about a better measure approximately and outcome measure, that may drive improvements in quality better than this measure.

Male: Is that's something that we should regularly consider as part of our evaluation from adoption or maintenance.

Angela Franklin: For maintenance measures, we definitely want to see, any differences in the gap between the endorsement, original endorsement date and the resubmission for re-endorsement submission? And I see some comments about that in the

comment form. Are there comments from other workgroup members about this?

Christopher Visco: So, Chris Visco here. I heard with the last sentence that you said that I didn't understand it, could you repeat it?

Angela Franklin: Sure, I'm sure, I'm sorry. There was a question about the performance gap and how we should be looking at it with all of these measures and across the maintenance measures that is the measures that are previously endorsed. We are expecting to see information regarding the performance gap since the initial endorsement. So any changes in the performance gap, we do expect to see that.

Christopher Visco: All right.

Wendy Marinkovich: So this is Wendy Marinkovich. We – from the information that they provided it looks like there is some very slight improvement and the time decreasing nationally. And if the benchmark, I think that's the median. There was again slight, one minute improvement. But it's been very minor improvement.

(Catherine Roberts): This is (Catherine). I guess I would ask of the NQF staff. I mean, what does that mean to us that as reviewers of this if we are not seeing any changes. I mean, what are our options? I mean is it that can we – where do we go with that? Can we just …

Angela Franklin: Sure.

(Catherine Roberts): … make that comment and it's a stand alone? Do we have to reject the whole thing? I mean, what do we – what's possibly …

Angela Franklin: No, no.

Christopher Visco: Yes, I guess – Chris Visco here, I second that. I'm just trying to understand what are – what my capacity is at this point in terms of recommendations because, yes.

Angela Franklin: Sure.

Christopher Visco: You know, given that, there hasn't been much change.

Angela Franklin: Sure. So the committee would want to discuss, you know, the fact that there's a change – that there isn't much change amongst itself and amongst the measure rather. And we'd want to have a discussion about what that means. And we could certainly pull the developer into that discussion. Why hasn't the gap moved much? And that would be a basis for discussing what the implications are. How the measure might improve in the future? A lot of times, we'll, in this particular category, we'll look at whether a gap or a lack of a gap in the case the measure is (topped out) or whether there's additional rooms for improvement and the reasons why or why not.

So, it's really for committee discussion. And we'd also want to hear from the developer about reasons for the amount of change in that gap portion of the measure.

Wendy Marinkovich: And this is Wendy again. One other point that is brought up within the description is there is some disparities between median times to medication as well as interquartile readings for the race ethnicities that they have listed. So we do see a little bit of gap there.

Wanda Johnson: One of the comments – this is Wanda from OFMQ. That we believed might be affecting the fact that those – the median time is not moving very much is because ED volume keeps increasing. And we – we are at the belief that as long as the volume keeps increasing in EDs then these times may not be – that they may not show much of an increase.

(Catherine Roberts): This is (Catherine) again. Then, does that mean that I mean, what does that mean then? I give you, you know, I mean it could even be that it gets worse if there's another variable that impacting this.

Wanda Johnson: Right, that's true. And one of the – the request of some of the, you know, the hospitals where they have a large volume is that we report some of these by ED volume. So that their – the distinction can be shown.

(Catherine Roberts): So this is (Kath). I still struggle with what this is actually measuring. And it would vary depending on what type of ED you have. So I live in not the ED world but certainly the imaging of long bone fracture world. And there are lots of people who come in and, you know, they fell three days ago and their wrist is still hurting or their ankles still hurting. That in no way will be triaged to be imaged quickly, diagnosed quickly and certainly not given pain medicine.

You know, within an hour or hour and a half of them showing up because there are lots of low level long bone fractures that do not require that. And it's still appropriate good treatment. So I really am struggling with what this is actually measuring. And that's – yes, so that's why I suspect it's not really changing much.

Sean Bryant: Yes. This is Sean. I couldn't agree with that more. I think there are a lot of these, you know, less acute fractures that come in and their pain level maybe well enough that not giving them pain medication in the ER is the right thing to do. Yet, you know, if you don't give it at all, you know, what does that do to this metric if all you're, you know, able to do is capture a timeframe and someone is in the ER for an hour and a half and they never get it? How is that skewing the results?

Casey Thompson: This is Casey from OFMQ. I would like to speak to that. We have the specific list of ICD-9 codes that only patients with long bone fractures fall into the population. And there is an exclusion if the patient refuses pain medication or if the physician, nurse, or other professional documents a reason for not administering pain medication. They are excluded from this measure. They abstract no to pain medication and they are excluded.

So all of those patients, if there is a large number of patients that still have a long bone fracture code that's in our list, that happens to fall in here but they don't need pain medication, we still – they're not included in the data, they're excluded.

Male: Was it …

Sean Bryant: If the nurse or the doctor says there's not a need for the pain medication?

Casey Thompson: Right, or if they refuse. If the patient refuses.

Male: Was there a way to pull that from the, electronically, because that – I would think that would be a bit of a challenge for patients that were deemed inappropriate for it to be able to glean that from the data mining for this.

Thiru Annaswamy: This is Thiru Annaswamy. I think from the way you're describing the collection of this in the algorithm, in the process, it says you are – the pain medication is administered, only those patients are going to be included in this measure. So I guess if there is no pain medication administered, am I correct in assuming that it wasn't even in the denominator?

Casey Thompson: That is correct.

Female: Right. But you still might give them a Percocet two hours later or something.

Thiru Annaswamy: Then that would (inaudible) into this measure, no?

Female: OK. No, it wouldn't be included either because unless the patient is under 18 years of age, oral pain medication is excluded for patients 18 years of age and older.

Thiru Annaswamy: Can you speak a little louder? I …

Female: Sorry, I missed that. Oral pain medication is excluded for?

Female: For – I'm sorry, is this better?

Thiru Annaswamy: Yes.

Male: Yes.

Female: Oral pain medication is excluded for patients aged 18 and over. Our technical expert panel believes that patients that enter through the emergency department with long bone fractures, major code – only the major long bone fracture codes are in our table that they should have parenteral pain

medication unless they refuse it or there's a reason they shouldn't for patients 18 and over. Now, oral pain medication is included in the measure for patients two to 18 years of age.

Male: OK.

Female: So I guess maybe I need help with the codes because I looked at those and I still saw codes that said like fibula, you know, (NOS) or, you know, something like that and you can still code a distal fibular avulsion fracture as (NOS), can't you?

I mean, I didn't see anything on there that excluded low level long bone fractures

Male: Yes. It's (inaudible), I completely agree. All the (NOS) codes are included so you can just about put anything, it in fact (ranges) about everything.

Female: And that's – the next one of the reasons that going to ICD-10 would have made this more precise. But because the ICD-10 was delayed, we are stuck with the ICD-9 codes.

Female: So then your measure includes the low level like avulsion, nondisplaced, tip of the fibula, nondisplaced radial styloid process, I mean, it will still include all those things, you know, a radial head fracture that's, you know, nondisplaced, fairly visible. It's been there for three days.

Female: Yes that's possible.

Male: So I just want to have a question related to the transition from ICD-9 to ICD-10. I know it's been pushed back one year but my understanding is that these measures if endorsed are active for multiple years before they're re-looked at.

So if this measure work to get endorsed, how would the transition from ICD-9 to ICD-10 be handled, should that occur next October?

Female: I'm not sure whether you want the measure developer to answer that or not, but we have already figured out the appropriate ICD 10 codes that would apply to the measure and have that table ready. But I think that was a

question for NQF, what happens when you do transition from ICD 9 to 10. As a measure developer, we would just provide those ICD 10 codes to NQF in the annual update.

Male: OK.

Angela Franklin: And that is the answer, my answer was going to be the annual update provides an opportunity for developers to update their measures with any additional information. Now, we included ICD 10 codes.

Craig Butler: This is Craig Butler. I have another question on the developer. How does the measure account for in the – some of the clinical variables that tend to like, whether or not their patient is conscious, has other injuries that might be relative contraindications that, you know, sedating to the fact that they might want to give an informed consent for something before they give a medicine that often (inaudible) things kind of on the ground level that there might be considerations that might – are quite to delay even to get the orthopedic surgeon in to see the long bone fracture or he might not want him sedated or medicated before he has a chance to do a complete neurologic exam or something like that.

There's just a lot of things, kind of ground level thing you'll say that might explain any – a delay and I mentioned pain medications that I know, I just want to need to understand how was it captured, eliminated or isolated and identified.

Casey Thompson: Correct. Any reason, any documented reason by physician or nurse for the delay in pain medication, we instruct the abstracter to abstract no. So, those patients will all be excluded. Even if later on, they administered pain medication and there's the delay, they still exclude it from the measure.

Craig Butler: So, if they – an altered mental state was the reason and altered mental state, mental state was documented as a diagnosis along with a lung bone fracture. But they didn't just fairly document that reason that they delayed giving pain medicine. Would that be included or excluded, how would that be treated?

Casey Thompson: We do have a few examples like patient unconscious and those sorts of things. Anything like that that's documented, we would accept as a reason for not administering pain medication, and would allow the abstract turn the abstract no.

Sean Bryant: Thank you. So, we shall now move on to the next subcriteria 1c high priority previously high impact. There was a little bit of discrepancy here on the comments. One person noted that while 78 percent of ED visits, pain is part of that. Long bone fracture is extremely small percentage of overall ED visits. And don't really see evidence of this as the high resource use of being a leading cause of M&M or resulting in severe patient suicidal consequences.

Others fell like, because 2 million people per year were affected by long-bone fractures that are good being a high priority, so, open that up for discussion.

Thiru Annaswamy: And this is Thiru Annaswamy. I just felt that while the time to receiving pain medication is an important outcome, the efficiency of pain management is a better, more important outcome. So, this being a process measure is only part of the high priority of the overall pain management sort of outcome measure.

Christopher Visco: Yes Chris Visco here. I would second that in that the – it does seem like this is a small component of something that is high priority. So in and of itself, the priority is difficult to recognize because the time to pain medication administration, I mean their – and as we addressed earlier in terms of the evidence, the evidence really is in the realm of quality of pain administration not necessarily time but this sort of falls into equator context of that.

Sean Bryant: So at this point, can I turn it over to Wendy to continue the discussion on this one?

Wendy Marinkovich: Yes. All right so we've now moved into the scientific acceptability of the measure properties. And that being the reliability and validity. So the, for the reliability, there was mentioned that they did not test reliability so they did not perform reliability testing as the individual critical data elements were tested for validity. And they have a notation and I guess that is going to go back to

NQF to make a statement whether that is something that should have been done because there is no reliability testing performed.

Angela Franklin: So this is Angela, yes, the developer is correct in stating that if measures are tested at the data element level for validity, reliability testing does not need to be presented.

Male: OK.

Angela Franklin: We would just ask that the steering committee then review the validity testing for sufficiency.

Wendy Marinkovich: All right so then we'll move to the validity and the comments there that the validity testing was used by (chart) abstraction and the elements were subject to that validity testing, that pain medication was the lowest agreement rate across all of them at about 85 percent. And that they did check no exclusions but 0.32 percent exclusions were noted in table 2. And then no risk adjustment regarding the comments – come or whatever, that's neurologically compromised that might be existing at the same time, I think the developers already mentioned that they would not be with – it would be excluded if there was notation of that status.

Other comments were that does the measure inadvertently measure time to diagnosis? So is there not necessarily the measurement of time to pain medication. But that the time that it would take to diagnose the long bone fracture treatment could be delayed pending the diagnosis. And then there was the exclusion regarding polytrauma patient again I think that was partially addressed in the original or the discussion we had.

Previously, there is a comment about contraindication to pain medication and maybe the developers could address that. Any other discussions? Go ahead.

Thiru Annaswamy: This is Thiru Annaswamy. The developer answered questions about how the abstracters would exclude patients in home that pain medication maybe contraindicated. But I would – it would probably be better if the exclusions are clearly specified in the way the algorithm is described in the measure form

in page 6, where the algorithm narrative is described. Perhaps those details could be included there as well.

Male: I would just like to add my little two cents the fact that the comment, I didn't make it, but as the measure inadvertently measure time to diagnosis, I think that might be a significant component of the measurements because I don't know that all long bone fractures are clinically apparent especially when they're near into the bone. And I don't know, and maybe the developer can address how they would extract those where to diagnose. Do they have a time in there to look when an x-ray was done perhaps, that might be a time to confirm the diagnosis of both the (inaudible) and the time that the patient arise at the emergency room if they're not allowed distressing they'll not be triaged to the head of the line. That's a little weight which would impact the measurement at that time to administration of that pain medication.

Wanda Johnson: So, this is Wanda from OFMQ. That would make it a measure, but we could – because what you're asking really is to collect the time that we made the diagnosis. And that would introduce another level into the measure. But, yes it could be done and yes it's possible that this measure is inadvertently also measuring time to diagnosis. But that at this point in this measure it is just – it does not pick up the time to diagnosis.

Male: So, if a patient had the diagnosis made later on, so it wasn't initially made and say they were admitted for something else. And then they have to make the diagnosis. Would that patient then, even though they were an in-patient at that point that they had come to the ER with that indefinite data set.

Wanda Johnson: It would. Well, I'm sorry. So, this is outpatient measure. So, if they got admitted they would not be included in this measure. So the diagnosis actually would have to be made in the emergency department. So these patients are sent home after their ED visit or else, they're transferred elsewhere. So if the diagnosis wasn't made until after they were admitted, they would not be included in this measure.

Male: Most long bone fracture are – I mean true long bone fractures are (inaudible).

Wanda Johnson: So this would only include those patients that visit one ED and are diagnosed. And then maybe sent to another facility for inpatients. But this measure would look at those cases that discharged from the ED. And what time and the time they gave their pain medication.

Male: I have to say I like the comments on the contraindication. So if pain medication is contraindicated, certainly that should be an exclusion. And if none, medication treatment interventions are used that could be excluded as well. And those are too, probably very reasonable things.

Female: There are some guidelines from NQF about the, with the validity testing and the summary of that that they provided, what is the agreement rate that we would be looking for?

Angela Franklin: So, the agreement rates that we're looking for is – I think we noted in one of the comments that there was a low agreement rate regarding pain medication, 84.8 percent. The remaining agreement rates were relatively high. And what we might want to ask from the developer is some discussion about why the pain medication element had the lowest agreement rate?

Female: That's kind of like the gateway data element, it's yes or no. And so, it's possible that like you've mentioned whether a reason is explicitly documented or whether it's picked up as a reason for not admitting – administering pain medication. I think the fact that the agreement rate is low, it's possibly because it's a yes or no question. And it's the gateway question.

Male: Can you explain what you mean by gateway question?

Female: Well, if your answer is yes, you're in the measure. So, that patient did receive pain medication. If you answer no there is documentation that there was a reason for not giving it or else they just didn't get it. And so, I would think that the agreement rate between what the validation contractor picked up and maybe what the hospital picked up is, you know, the hospital maybe looking for reasons for not giving or it's not explicitly documented. And that the CDAC who is the validation contractor may not be picking up those reasons that the hospital may have picked up.

Male: Thank you

Wendy Marinkovich: Any other discussion? Thank you. Do I turn it over for criterion three? Are we doing this in (thirds)?

Angela Franklin: Yes, that would be fine, thank you.

Wendy Marinkovich: Right, I believe – Chris Visco?

Angela Franklin: Chris? Chris Visco, do we still have you?

Christopher Visco: Oh I'm sorry, geez, I was talking, I was on mute. I can hear myself but you guys can't hear me.

No, I'm here. So, in criterion three, (inaudible) the specifications including the measurement logic. Shall I go through the comments? OK, so just to expedite, going now to comments. When the comments, the cost administrative burden has not been assessed. And the evolution of electronic health record in emergency documents which should be feasible.

To find data elements are originally generated and cost administrative burden for updated facts not assessed. Stated facilities with EHRs have less burden. So the questions is, what percentage of the facilities have EHRs? It would be interesting to see and I can imagine that there are many without at this point, but – and then you know, the time it takes to share abstracted data. Otherwise, there are no concerns. So, you know, mostly around percentage of facilities with EHRs.

Thiru Annaswamy: This is Thiru Annaswamy. Am I remembering this correct that this is a voluntary measure that not all EDs who have EHR will be participating in it? Are – is all EDs participating with EHR automatically in this measure?

Female: This is a voluntary …

Christopher Visco: Once I get the peak at 27 measures and this is one of the options.

Angela Franklin: Yes, this is Angela. And I would say that this is a voluntary measure, but I'd also ask the developer to comment about the useful measure and or the …

Wanda Johnson: As far as I am aware this measure is not retooled for EHR's. So, it is collected by chart abstraction. And then is not specified for EHR.

Male: And do we know if that's in the plan for this measure? Or …

Wanda Johnson: Well, and if they allow us to give our opinions it's very difficult to collect, you know, as we've talked about on this call the reason for not administering pain medications. They're pretty much has to be a field, that says there is reason. So, I think that CMS has passed over this one. So I don't know of any plans to retool it, but I do not speak for CMS.

Angela Franklin: So this is Angela. I just had a couple of clarifying questions about the data elements that are used to collect this measure. And I just want to be clear. Are they routinely generated during the healthcare delivery? And I think I heard that they are or are not available in any electronic form or from any electronic source. And that's for the developer?

Wanda Johnson: Yes so, you could if you just kind of take a step back and look at this measure, if they were provided pain medication, so you would know that from the in electronic health record, you know, the eMAR the medication administration record and then you could just look at the time to pain medication. So it could be retooled if you just would not pick up any rationale for not administering pain medication.

Wendy Marinkovich: Great. So the spirit of understanding of the actual burden on a facility to collect this type of data, looking at the benefit of collecting data for this measure in terms of burden versus benefit?

Wanda Johnson: And that's not something that we typically have ever looked at, excuse me as far as burden is concerned.

Sean Bryant: And so any additional comments at this point?

Wendy Marinkovich: That would – I didn't hear any that would move us to our usability and use question?

Sean Bryant: OK, so shall I continue or is that, so I'm going to …

Wendy Marinkovich: Yes that's fine.

Sean Bryant: OK. So this is the usability and use to the extent in which special audiences are using (inaudible) where we could use the performance results for accountability and performs improvement to achieve quality and efficient health care. And the comment on the accountability improvement, unintended consequences are given a continued performance gap. At first I wonder they're reporting measures, have to choose and in the form – in the HOQR how many actually choose this one, it's one of (their 27) required to report. And so, there's some comment about the optimism that fill the optimistic view of the demonstrative progress on improvement. But not a lot more here, so I will just open it up to a comment then.

Wanda Johnson: This is Wanda from OFMQ, I do want to mention that this is not a voluntary measure in the HOQR program. In the future it could be voluntary for EHR if they choose to retool it. But currently, if you participate in the HOQR reporting program, you will submit this data.

Male: OK, that's a good important point and how I guess as the NQF staff, how do we know what is required, not required in a column like the HOQR?

Female: We really.

Male: Or is that important?

Angela Franklin: You mean, yes, we really look at whether it's being used publicly for public reporting that's as far as we looked, we don't look any farther from that.

Male: All right so if we say part of the HOQR, it wasn't required, and nobody chose it?

How would that, how would we know that, I guess?

Angela Franklin: At this time, we don't really, we would rely on – we don't look at that from the NQF perspective at this time how the measure is actually being implemented. And whether anyone chose it when they – it was presented in the public

program. We rely on questioning some information presented by the developer.

Male: Is that (inaudible) determining usability where you can use it?

Angela Franklin: Yes, approximately for determining use, whether the measure is in you.

Thiru Annaswamy: This is Thiru Annaswamy. I have a question on the benchmark. It says the benchmark is based on the measure performance in the top 10 percent of the hospital's reporting. I was wondering if that's the better way to do a benchmark or if there's an actual number that's recommended based on analysis as the median time benchmark because the top 10 percent rate may vary based on practice other than quality. It maybe based on the type of long bone fractures they see in their hospitals.

And other factors that may not – may influence the measurement, may not necessarily mean the same thing for the rest of the 90 percent. So any thoughts or comments on this from the developer?

Wanda Johnson: I'm not sure, we do see achievable benchmarks of care methodology for quite a while for CMS and well I might agree, it's the way the benchmarks are reported. So, whatever the top 10 percent of the hospitals, whatever they are performing, however they are performing, they are used as the benchmark. It's just the recommended methodology that CMS has used.

Male: Thank you.

Sean Bryant: OK, so any additional comments at this point? Should we go down to criterion 5?

Angela Franklin: Sorry this is Angela, for criterion 5, we would – we won't get to that today that's the comparison to related or competing measures so we would just be – there is no competing or related measure for this one. So we just conclude if there are any concluding comments about this measure. We'd entertain those but then we move on to our next measure.

Male: And I guess I would just speak for myself. I may – had a loss for how we are doing as a committee.

Female: Well, you guys are doing great.

Angela Franklin: You're doing great. These are all.

Female: Really, yes.

Angela Franklin: These are all the questions – this is the purpose of the worker calls is to really start to have the discussions and understand the questions that we want to ask but of the committee and of the developer as we go through this process. So this is perfectly fine.

I will just make a time check that we have two more measures to get through though I would also encourage folks if they have questions about this measure that you like to wrap up with, that would be great. And we can also continue the discussion offline and engage the developer that way as well.

Are you all ready to go to 0052 which lead discussants are Catherine Roberts, Thiru – oops sorry, Zoher Ghogawala and did Carlos Bagley get on the call?

Female: Yes.

Angela Franklin: So that's 0052.

And I just wanted to check and be sure we had the developer NCQA also on the call, Jenna Williams-Bader?

Jenna Williams-Bader: Yes. Hi. I'm here.

Angela Franklin: Great, great all right.

So I think we're staring with – so are we starting with you again, Catherine?

Do we still have Catherine on?

Katherine Gray: You have Katherine Gray but I think it might be Catherine Roberts.

Angela Franklin: Roberts.

Katherine Gray: Yes.

Angela Franklin: Uh-oh.

Katherine Gray: She may have dropped off …

Angela Franklin: All right, who is the next discussant?

Since we lost Catherine Roberts, Zoher would you like to lead off on this one?

Zoher Ghogawala: All right, this is Zoher. I'm sorry I just – I didn't – I was pushing my secretary to change my schedule and I did not prepare for this meeting today. I really apologize for that.

So I don't think I can lead the discussion but I'm more than happy to participate.

Angela Franklin: Well how do people feel about we can also look at the next measure which …

Female: We can into – why don't we go to, unless Carlos Bagley is on the call and he wants to lead, Carlos?

Female: Hi. I'm his nurse actually. He like just stepped out. He will be back in about two minutes so I can see if he likes to …

Angela Franklin: Why don't we jump in to 0514 and then we can go back to 0052, how does that sound?

Zoher Ghogawala: That sounds great. I apologize

Angela Franklin: That's OK, no problem. We're flexible.

Male: Thank you.

Angela Franklin: OK. So 0514 is – let's see, do you want to go ahead and start with this Craig?

Craig Butler: Sure I can start just correct me as I go. First, I probably need to apologize with some on my comments. I kind of dumped a lot of content and it (inaudible) kind of guide, posting notes to myself. I didn't realize it. Other people think would be – it may not, you know, they don't make a lot of sense to other people a lot of times, but, so my apologies but let's go forward. So this – I'm sorry.

Angela Franklin: No worries.
]

Craig Butler: OK. So our measure is 0514 roughly measures the percentage of patients with the diagnosis of low back pain that had a conservative treatment documented via the claims data prior to the MRI.

The first part of the first criteria is measuring importance of the measure and report I think the – one, I think there's some continued question about what type of measures they are. Are they efficiency versus process and what I think most people has acknowledged that this is an important area due to the high stand – high cost, high resource use and often utilization.

What else, here? There, I know there are some comments about the update by the ACR and just one of it – one of my colleagues who made those comments might help me understand, are they suggesting that they be added to the measure or can we do that during this process at this time?

Angela Franklin: And this is Angela, just a quick note. At this time, any additional specifications or changes could not be added to the measures, unless the developers are able to do so, but with our – within our time frame, probably not be possible at this time.

Craig Butler: All right, all right.

Angela Franklin: We just have to review the measures?

Craig Butler: OK, all right. There's also comment that quote on, I'm somewhat concerned about the detail that the evidence of various quote, guidelines enclosed with different criterion, meaning there's noise and evidence and they're either used to be addressed to explain that it does not make a difference or what

difference it means in using the data. I wonder if we can have this developer respond to that comment.

Angela Franklin: I guess that's Wanda.

Female: I didn't the measure developer from CMS was on the call.

Angela Franklin: Yes. Wanda was listed for both – I thought we had Wanda on the call.

Female: Yes. So we have Wanda for the health care.

Angela Franklin: Well, so we are – we'll take a note of that. And propose that to the developer if we could continue then our discussion?

Craig Butler: OK. They – in some of this I think evolved the question I had about the – they didn't identify any harm but I think most of (inaudible) utilization can cause harm. I'm just wondering how a measure would account for that or does it – is that not something we should be (boggling) down with to this (day).

Angela Franklin: How the measure – let's see …

Craig Butler: You said, one of the similar details in the – I'm sorry.

Angela Franklin: Yes, I was just looking at the comment. No harms identified for overutilizing MRIs.

Craig Butler: Right, is that – I thought they were supposed to identify the harm in the …

Angela Franklin: Yes. And that is a valid question for the developer, you know, whether they have identified any harms and for the discussion for the committee whether this is something that gives them real pause. And so I'm just considering this measure if it's not there.

Craig Butler: And I'll just tell you, my bias as an orthopedic surgeon is I realized that fine (inaudible) is way, way out of hand. You know, MRIs do show abnormalities that are currently correlated to produce great harm. And so, not to mention any harm from overservicing or overutilization to me (inaudible) mission.

Angela Franklin: Great and that's exactly what we're – yes, and that's exactly the discussion that we want to have amongst the committees.

Craig Butler: OK. So does anybody else have any comments on that to the other committee members?

Thiru Annaswamy: And this Thiru Annaswamy. The basic antecedent conservative therapy definition I have some problem with or doubts about, so the way I understand it is, if the MRI was done within two months after physical therapy claim or chiropractic claim or any kind of E&M claim, then that would be excluded from this measure, from the denominator. But, let's assume someone had an E&M code for low back pain, and then, the MRI was more than two months after that, whatever delay, reasons for delay in getting the MRI.

Then, that person would show up on this measure as a positive MRI without antecedent conservative therapy which I think would be inappropriate. So there's a fundamental flaw in the way the measure is being defined, I felt …

Male: Are we just completing the two measures on the MRI? I felt that the ones from NCQA had that time limitation. This one just had the percentage of patients with the diagnosis of low back pain that had conservative treatment and they didn't – I don't recall a time except for a 90-day exclusion for, in the denominator related to surgery. Is my recollection accurate?

Thiru Annaswamy: No actually, and the description of the measure is pretty clear. It says the antecedent conservative therapy is defined as claim for PT 60 days before the MRI …

Male: OK. OK.

Thiru Annaswamy: … claims for chiropractic, 60 days before. And claims for E&M greater than 28 days but less than 60 days before MRI.

Male: Right, OK. I recall that now.

Thiru Annaswamy: OK. And then I'm also fundamentally having a problem understanding when would this happen? The MRI is ordered. You would think it's ordered

after some kind of evaluation. What would be a scenario in which an MRI would be performed without somebody talking to the patient and doing some kind of E&M and billing for an E&M. Fundamentally, I don't understand when this scenario would happen.

Male: So we have maybe the developers on to that?

So, we don't have developer on this call?

Angela Franklin: It looks we don't have a developer on, I'm e-mailing …

Male: Yes, you know, Chris Visco, OK that's a huge, huge flaw for this. We definitely want to get back to a developer on that.

Angela Franklin: Yes. Yes.

(Catherine Roberts): This is (Catherine) too, I mean if this is Medicare. There's a huge flaw on this because anybody over 70 can get an MRI period, that's agency criteria. So, it's a very small group of people that would not either fall into one category of having something done to them or they're over 70.

Male: I think that (sums up) again when, and the fact that they did not risk-adjust this measure. And meaning, you know, I think at one point I alluded, can you really have uncomplicated back pain and somewhat just chronically over the multiple medical problem and over 70, yes, another …

Angela Franklin: Well, which is (inaudible) but then, according to the guidelines, over 70 is clear. You get an MRI without questions.

Male: OK. OK.

Angela Franklin: Regardless of anything else.

Male: Yes, and another is defining an evaluation and management code as the conservative therapy is also inaccurate. If someone saw the patient and if the intent of this measure is to, for uncomplicated low back pain in the Medicare population, have some kind of antecedent conservative therapy before you get

an advanced imaging study like the MRI which is what I'm assuming the intent of the measure is to capture that population.

Then, just doing an E&M is not sufficient because it doesn't mean you got any kind of conservative therapy, it just means you were evaluated. So it still doesn't capture the intent of the measure, I think.

Craig Butler: I think that's a point will-taken. You can manage a clinical scenario with the PCPs as well. If it doesn't get better, call me back and we'll get an MRI.

Male: Yes.

Craig Butler: You wouldn't catch that, that would pass the test if you will.

Male: That's right, yes.

Craig Butler: OK. Any other questions on the evidence to support the measures, then you'll hear the comments down there, but I think we have enough questions for the developer to move on to a performance gap. Some of the comments in there point to the fact that there's a continued gap in quality. One of the questions I might backup is we're doing maintenance evaluation on the measure, is the assumption that the measure already made it to this consideration with a previous must go scalable standing committee and path and therefore was adopted, or are we kind of a second round, but a new level in-depth of review of the measures that were made.

Angela Franklin: That's an excellent question. Actually, what happens is this is, we should review this measure with fresh eyes and I'd like to advice the committee that are criterion, the way we apply has become a bit more stringent since the last time this measure was endorsed. So, we really ask this measure – the committee at this time to look at the measure with fresh eyes but also look at areas in the measure regarding gap whether there's an improvement there or as well as use and what happened in use the measure.

Craig Butler: OK. So, we're going to look at the performance gap. One of the comments was that the way that many change them, 32.1 in 2007 to 31.4 in 2011, obviously suggesting a continued opportunity for improvement. Question is,

how do we, you know, relatively small changes between pre-measure and post-measurement is that, how significant in the way that in terms of the use of the evidence for a continued use for the measure, is that a factor or not a factor?

Angela Franklin: I'm sorry, you're asking about the …

Craig Butler: That was probably articulately (ordered), let me try again. The difference that we see different measures, you know …

Angela Franklin: Yes.

Craig Butler: … what the benchmark was perhaps if they have been measured or a previous measurement and did the measures implemented, it comes up for maintenance and we see little or no difference or perhaps even regression in the measure.

Angela Franklin: Yes. Yes.

Craig Butler: How do we weigh that in judging the measure?

Angela Franklin: So, both again, a question for the developer to discuss why there wasn't a change between these years 2007 to 2011 or in this case a regression. And we also wanted to ask the committee – well, the committee won't be able to really discuss this in -depth without that input from the developer who we have reached out to see if they were able to join. So, at this point, just looking at face value it's really up to the committee to discuss whether the measure is doing what is intended. There is of course a continued gap in care quality and issues around continuing to have the measure and endorsed given these issues. And of course we'll reach out to the developer for answers.

Craig Butler: All right. So, I guess in the presence of questions for the developer, we just need to identify those and keep moving to …

Angela Franklin: Yes.

Craig Butler: … criteria, OK.

Angela Franklin: And I'm sorry about that. And there's also – no, I see there's a note there's no information on disparities, that will also be a question both for the committee if you're aware of any information regarding disparities, regarding the focus of this measure. And then, a question for the developer as to whether they have discovered anything.

Male: Here's actually information on disparities in the article cited by, cited in these measures. This information about people in the minorities, there's more delay in getting advanced imaging, primary care physicians, there's 78 percent greater likelihood, getting early MRI, whereas if the initial provider was a chiropractor, so 50 percent reduced likelihood, getting an early MRI. So, there is some disparities in the information provided.

Angela Franklin: OK.

Craig Butler: OK. So, they are not – necessarily have to extract that from their (recommends) to include it for us to stay too kind of past that (inaudible).

All right. Then the 1c is about high priority, I think most people believe it is a high-issue or (too high) volume and therefore a clinically-developed issue. So we can get that where – is that part of criteria one. And move to priority criteria number two, is one of my other lead discussants want to take that?

Thiru Annaswamy: Yes, sure this Thiru, I can take – go to criteria number two. This is for reliability and validity. There was some public comments submitted which included that in the denominator, definition is potential to severely underestimate the information. So, the imaging studies coded with the diagnosis of low back pain, et cetera which are uncomplicated back pain, but once they get the MRI, if it's a Medicare population particularly as previous comments indicated, there's a high likelihood of positive abnormality. In that case, the diagnosis would be recoded with a positive finding on the imaging claim. So then, you would not hold them uncomplicated in low back pain. That also prompted me to kind of look at the value set that was enclosed along with this measure. And the values that included not just low back pain, nonspecific low back pain codes, but other degenerative codes as well. So I'm not sure that those patients would be necessarily excluded, so that that was a

question I had for the measure developer, if it is just including uncomplicated low back pain codes or any low back pain codes without the exclusion of neurological disease, et cetera. So that was a question I guess we can table for to ask the developer at a later date.

Ann Phillips: This is Ann, and what I'll do for these measures if we don't have a developer on the call is when they get the transcript for this call, I'll pull your questions out and I'll e-mail them to the workgroup, and that we can make sure that we bring them up in the in-person meeting.

Thiru Annaswamy: OK. Moving on. The other comments in the pre-workgroup comment section was, could there be additional exclusions to be considered which included unexplained weight loss, unexplained fever, prolonged use of corticosteroids, et cetera. So that's something that could be used for future versions of this measure I guess, but will not be applicable for this current discussion. The MRI of 52 modifier for a limited lumbar exam, that's another question that the developer needs to answer.

Another comment that I have many issues and questions about this section, a precision of the criteria does not seem to make sure that those in the study can be consistently selected for comparability over time.

I have trouble categorizing the measure as well defined because of long potential list of valid denominated exclusions that were not listed. And once again, another question about exclusions. So, history of surgery should be an absolute exclusion criteria as opposed to just within 90 days, because post-operative back surgery patients cannot be categorized as uncomplicated back pain. That's certainly a point well-taken.

The new denominator exclusions for this endorsement maintenance was lumbar surgery codes and neurological impairment code. So one of the concerns that I had about it was if you excluded lumbar, spine surgery within 90 days? But …

Male: Correct.

Thiru Annaswamy: … if you – does that mean that automatically that you can get a new MRI for back pain within 90 days after spine surgery without an evaluation and management code? That's what that seems to imply which also doesn't make sense to me because why would you just go ahead and get an MRI on the phone if you had recent spine surgery without checking the patient out? So to me, it doesn't seem very appropriate to exclude those patients.

Moving on, unless there are any other comments about this?

The reliability testing using algorithm 2 as the reliability was moderate? Someone said they couldn't open the measure testing form. Hopefully, that person was able to open it later. Another person, rate of the reliability is low using the algorithm.

So there were some discrepancy in the rating of the reliability among the community members. The validity, someone mentioned concerns about changing inclusion of more places to gather data from any inpatient/outpatient carrier claims. We have interest to know the differences – what differences these adjustments would make to current study period?

Another person mentioned that the validity is insufficient because elderly patients with multiple chronic conditions will undoubtedly impact clinical decision on MRI ordering or utilization. So chronic conditions and other factors should also be included in the risk adjustment. So there's a validity of all this measure. And then probably similar to comments endorsed earlier is a chronically ill elderly patient in the Medicare population ever truly because they have an uncomplicated back pain.

So repeat the occurrences of these comments and thoughts. So the other comments in this section are similar to what we have previously talked about. Do I go on to 3, or is there another workgroup member available to take it on?

Angela Franklin: Well you can go on to 3, but are there discussions, are other comments from other workgroup members?

Male: OK, we assume that some of these questions raised are going to be addressed by the developer when they get to the …

Angela Franklin: Yes, we're compiling them and we'll be after this call sending them to the developer for response, so.

Male: Do they get a copy of these worksheets ahead of the call?

Angela Franklin: I believe they do, I'll just (inaudible) we really …

Female: They get the question.

Angela Franklin: They get the question. OK, the question and usually the developers who are on the call and we're going to bring the developers for this measure into loop.

Female: Yes.

Female: Via email and then includes you all on the responses.

Male: OK.

Female: … and because (inaudible) dialog in that way, prior to the meeting and get some of these questions answered. So, on to criterion 3, is anybody going to take criterion 3?

Male: I can continue if that's OK.

Female: Yes.

Female: OK. That's fine.

Female: Sure.

Male: Sure. A short comment here on feasibility. Feasibility is high because it's claims-based. And the other comment said is feasible for data collection analysis as it's part of the workflow data collection in outpatient settings. But it may not be feasible if few Medicare patients would not qualify an MRI based on low-back pain, not quite sure what that means but it looks like there was a question about a part of feasibility.

(Catherine Roberts): All right, this is (Catherine). That's my comment and I'll just clarify. There are two criteria that's according to the guidelines and one is trauma and there's very few elderly patients who would come in that have not fallen or something that's causing their low back pain. And or that they're not over 70 and both of those criteria would qualify you for exclusion – I mean, that in effect, you should have an MR. And so that means that the population of who really could be included in this turns out to be very small if people are really following these guidelines.

Now, whether or not they're documenting it correctly or anything else, is another question but it certainly, it's – at the very root of this that the criteria, you know, yes you can collect all the data but what does that mean?

Male: I'm also not sure what you mean by someone over 70 should get an MR. I mean I think they should get an MR based on the evaluation done but I'm not sure if just being 70 would automatically qualify them to get an MR.

Female: It does according to the guideline rules that they have listed in all their – in the different guideline section.

Male: I see.

Female: So, so that's the thing. That was part of my question to hopefully the developers could tell us that but that would, you know, that if – I mean if you would think about that they are the real physician or somebody, you know, ordering this, that if they are reading these guidelines, that alone would be sufficient, you know, to say yes to get the MR. So they shouldn't be – they should not be penalized for following the guideline.

Zoher Ghogawala: This is Zoher Ghogawala. I'm also a little troubled by this guideline that suggest that advanced imaging is indicated for patients over the age of 70 just because they're over the age of 70.

Male: Yes. I think that's in the presence of a complaint of a back pain.

Female: So it maybe true, it maybe true. And they could even have trauma, but either way, they get there.

Male: They have trauma that it completely makes sense. But a 70-year-old with low back pain to me could be evaluated clinically and would not necessarily need advanced imaging. Am I wrong about this or – I mean, I'm not aware of any real guideline that says otherwise.

Female: If you – I can look it up but it's one of their guidelines. I don't know if I remember the page but I can find it for the group but yes, it's in there.

Male: OK, yes. I would also like to know – which guideline you're referring to because I didn't read it the same way so definitely that clarification would help me.

Male: Yes. I mean from the – (inaudible) neurological surgeon's perspective as well as from that, I mean that – this would be news to me.

Female: OK. I'm flipping through the pages here. I'm trying to find it.

Male: So, should I move on while you do that?

Female: It's on page nine and it's under guideline number five. In second variant.

Male: Page nine of the evidence form?

Female: Of the whatever, of the middle whatever the measurement form is, yes. Yes. That's the evidence form. There is one or more of the following that means age over the 70 is one.

Male: I'm not aware of any evidence that would support that.

Female: It's just what I'm saying. It's – whether they don't discuss it so I'd like to sort through that and it doesn't appear to be in any of the summaries of the references. We're going back and reading the whole reference. But it's in there and it certainly draws a huge monkey wrench into selecting data – from anybody in Medicare that would not …

Male: Sure, no I understand that. It's just that I think this should be (inaudible) a little bit because I mean, I'm curious – I'm curious to hear what other people on the committee think but this is absolute news to me.

Male: Is this guideline 5 you're talking about?

Female: Yes.

Male: OK. To me there's a fundamental question buried in all of this which is the – judging the evidence used to support the measure, you know, when you get into the references and you kind of, you know, because obviously some of them can be known as the expert opinion that's included in one of the references you need to support this measure and then you were kind of on another level kind of a (guideline) on the (inaudible) of that is adequate for the measure to be adopted or maintained.

And so I've struggled with that content as well that, you know, where the 70 year-old automatically you know is given an MRI but I can understand how it comes about understanding, you know, in part, yes there's adjustment of what scientific high level evidence but there's also room for (content), there's expert opinion and all that in formulating a clinical (STPG) on a given disease or disease entity.

Male: OK. May I move on to usability and use? Are there any other comments on the feasibility? OK. Moving on, public reporting and quality improvement with benchmarking are well-described. There is a concern that the weighted means haven't changed much. This serves as benchmarking information but also reflects a limitation of the ability of this measure alone to change practice patterns.

So, there, I guess there needs to be more attraction on what does measure means and how important it is in order to continue to have usability. Since the data collection is easy but the impact is difficult to measure or potentially making large numbers, the measure's usability is troubling as it's currently constructed, was not comment.

If the impact is only on outline facilities then it would be difficult to use the changed performance for accountability or improvement performance in any expensive way so that limits how much this measure can do to change practice. Those are all very valid comments.

Any other thoughts or comments about this?

(Catherine Roberts): This is (Catherine) again, my one point or thought about this that I would have asked the developer is why did they not tell us anything about these outlier facilities or why did they believe that they could make a difference in the more, you know, I mean there's something and we don't even understand who they are. I mean, are these, you know, facilities that are, you know, the disparity facilities, you know, where people are not getting good care, you know, what about them? Why would you even want to try to make a difference because if you're not moving the, you know, the mean or the median, you know, you're only reducing the variance.

I don't think that's going to save any money or do, you know, even make any, you know, that it's really hard to see any impact unless you can specifically break those units out and look at those which would be a different use of the data but if they're advocating that they can't move the middle but they could, you know, do something to impact the outlined facilities or outliers, you know, we just need to understand what would that – who are those people and how, what would be the theory of what could be impacted?

Male: I think if you look at some of their references, kind of reading between the lines, I think those are the physician-owned scanners which tend to be the high utilizers. You know, at least that's what I kind of gather reading between the lines but again, I think kind of clarifying who are the high utility users that seem to be using it inappropriately would be very helpful.

Angela Franklin: This is great discussion. Is a developer on the line because we have an email that they were going to dial in? Are they on? No – OK.

(Kelly Anderson): Hi! Yes. This is (Kelly Anderson) from the (inaudible) group. I believe we may have a few other members of the developer team joining but …

Female: OK. Thank you.

Female: All right.

Angela Franklin: Great. (Kelly) I don't know if you have heard any of these questions so far?

(Kelly Anderson): I just recently got on the call a moment or so ago, but if there's any questions I can help out with, if you could repeat them that would be great.

Angela Franklin: OK, great.

(Colleen McKernan): Yes. This is also (Colleen McKernan). I'm also from (inaudible) group with (Kelly).

Angela Franklin: Great thanks.

(Charlie Broodman): Hey, (Charlie Broodman) also from the (inaudible) group side.

Angela Franklin: Great thanks. So the committee actually had several questions and I'll leave it open to the steering committee members. So we want to start at the beginning where we had some questions and just quickly summarize, we had some very critical opening questions for the measure, or do we want to …

Male: Sure, we can start there. That's fine. All right, several of us have a question on what the intent, the fundamental intent of this measure was rather than us asking the questions if the developer can concisely state what the main intent of the measure was that will help.

Male: Well, (OP8) is basically there's, analysis shows that there are two major issues. There's a huge volume sort of impact and there are clear guidelines on what should be used or when it should be used, the use of MRI that it should not be the first indication and that there should be other prior treatments. And therefore, we – there is a huge, with a high utilization, and it was the goal is to decrease their inappropriate use of MRIs for low back pain.

Male: OK. So with that intent, if your measure states that antecedent conservative therapy may include claims for physical therapy in 60 days before the MRI chiropractic before 60 days before the MRI and any claim for evaluation

management code greater than 28 days with less than 60 days before the MRI, how would the just an evaluation alone suffices conservative therapy and what would be scenario in which a patient would be getting an MRI without even an evaluation if you can explain the possible scenario with what would happen?

Male: Well, I think those are the evaluations, E&M are proxies all the potential having access to that after the diagnosis is that they have been on the claims as we don't get all the information is that there is the possibility of having some treatments before they go into indication directly to the MRI. That's why we look at those E&M codes and that has potential proxies for antecedent before the MRI is performed. That's why we use that coding. Is that clarified?

(Colleen McKernan): And I can – this is (Colleen). I can give some additional detail on the – the change actually occurred back about four years ago that we added this kind of broad E&M exclusion. And that was something that we talked through significantly with out tech list or panel. And the rationale for including the broader definition of evaluation management is that we were trying to absorb all claims for E&M visits that may involve multiple complaints.

So we're making the assumption that there may not all the – the complaints may not be captured appropriately according to the diagnostic coding on the claim. And so, using the broader definition of evaluation and management, again because using the – we're using a claims data source; we were trying to be more inclusive of a visit that could potentially be evaluation and management that maybe missing a specific code for evaluation management.

Male: So the scenario in which a patient would not even get an evaluation management before they get an MRI would be that the person ordering the MRI ordered it on the phone without checking the patient out?

Male: On the first visit. They can – the patient comes with the – comes to get a visit and there's an MRI indicated. And therefore with MRI indication, we did not find evidence of a prior treatment and that would be the indication that there

was no prior antecedent treatment and they want to do (inaudible) MRI (Colleen) can you …

Male: But (inaudible) from the visit directly to get an MRI without any treatment.

Male: That was a first indication, yes.

Male: OK. And what if there is a delay in the MRI to more than two months after the initial visit, how that would be in captured in this incorrectly?

Male: Well (Colleen) – that's exactly how – I don't – when we look at that claims data, all the visit, how we capture that specific …

(Colleen McKernan): So, yes. So the – we allow for definitions of conservative therapy to this I think you've already been and so physical therapy 60 days before the scan chiropractic evaluation to the patient before the scan. Or any of those blanket E&M claims that are more than 28 days, but fewer than 60 days. Before the scan. And the reason for that again, is we're considering all these things as appropriate care so if the patient does the scan more than 60 days after the evaluation and management visit, we're considering that to be not – either not associated with or not matching the evidence used to find that evaluation and management area.

Female: Solution.

Male: Yes. OK. I mean, that could be – that could actually reduce the validity of this measure because you might be inappropriately excluding or including patients who got an appropriate E&M but they just had a delay in the MRI. The other question was about lumbar surgery within 90 days of the MRI, once again, if the patient had recent spine surgery and they get an MRI without appropriate antecedent conservative therapy that alone we are questioning whether just 90 days recent spine surgery alone would be appropriately excluded because they would need some of kind of evaluation as well.

The other question was about the diagnostic codes and the value set. We – are you trying to capture uncomplicated low back pain but not degenerative disorders that do not have the exclusions listed? Are you capturing any back

pain code including the degenerative disorders listed in the value set, if they don't meet the exclusions?

Male: All right, that's another layer to the question and it backs up to the exclusions for surgeries and I would comment on that. And I thought that surgery should be an absolute exclusion just because once you've had surgery, I don't think you can ever be qualified with uncomplicated back pain and I just wanted to hear perhaps some comment on the developers on that.

Male: (Inaudible) any surgery, I mean, any surgery would be considered as a (derived) exclusion any patient with a prior therapy?

Male: Prior surgery.

Male: Yes. I'm sorry. Yes. Prior, yes.

Female: Prior surgery of the low back.

Male: Correct, of the low back.

Female: Yes.

Male: But isn't the exclusion only if it is within 90 days of the MRI?

Female: That's stated currently.

Male: It is currently but I'm saying even if it was 120 days, they should be excluded. I don't know how you draw the line at 90 days for someone supposed to have a lumbar fusion or discectomy because there's a whole differential diagnosis, a little different for someone who comes in with, you know, uncomplicated back pain without a, it's just surgery and I'm speaking as an orthopedic surgeon.

Male: Right. But isn't the question, if they have had prior back surgery, you should still perform an evaluation management before ordering the MRI. Isn't that what this measure is trying to look at?

Male: That maybe true. I guess it depends on the nature of the complaints. And I guess there had to be a lot of variables in that that makes me think that they're including anyone with prior back surgery would sort of degrade the accuracy of what's written on the measure.

(Catherine Roberts): And this is (Catherine). Let me – and I think the – I mean I'd have to go back and look at the criteria. But it seems like the E&M visit though, you have to wait at least 28 days from the date of that to the MR to allow for conservative therapy. So that wouldn't make any sense like if it was 120 days you have to wait another 28 to actually get the MR, that does make sense …

Male: Right.

Angela Franklin: … that back surgery. So yes, there is something wrong about how this all fits together.

Male: Any comments from the developer about the diagnostic codes?

Male: And I can, you know we – the 90 days, and I understand, you know, that the orthopedic surgeon has concerns about the 90-window. This has been discussed with the extra (inaudible) included clinicians and it was set on this time frame for any patient at the 90 days. That's how it was even at that prior NQF meetings when it was originally endorsed. And that has not changed since its initial endorsement. And (Colleen) you can clarify if that's …

(Colleen McKernan): That's correct.

Male: I assume that that has not changed in our latest specifications at all.

(Colleen McKernan): Correct, it is still 90 days. It's been 90 days I believe the last time as mentioned (inaudible) the last time we received endorsement.

Male: And is that with the technical advisory panel that included that?

Male: It was discussed with the technical advisory panel. Yes, and this was in 2000 – be at 2009 probably around that time. I think it was endorsed in 2010. I'm not incorrect or it maybe it was in 2009 endorsement. It was a few years ago.

Male: I would just have a question about, you know, considering the different procedures that are dominant in lumbar spine just likely seen no matter what you have done. Even a fusion at 90 days, it's not even expected to be healed or fused and you're saying that if they a problem they should get an MRI or they will be downgraded if they do. And that's a little worrisome. And I would, you know, love to, I don't know how you really address that. But that would be something that I would question at the – tell the foundation (inaudible).

Male: There was an exclusion for a lumbar spine surgery as a measure exclusion.

(Colleen McKernan): Yes.

Male: And the reason stated is that candidates for surgery to address low back pain on radiculopathy. But there is a – there's a possibility that when you exclude – the reason for the exclusion is potential candidates for surgery then you maybe inappropriately excluding patients where MRI is used to screen for lumbar spine surgery without an appropriate evaluation management. So there seems to be a little bit of circular logic there or maybe it's my problem in the way this – the reason for the exclusion is written. But that's one of the questions I had as well.

Male: It's OK and you're looking at – this isn't hard. Unfortunately I don't have that document you are looking at. Which part was this? Can you repeat that?

Male: I believe it's under exclusions in the reliability testing. But I'm also sensitive for time. And I was going to ask NQF if we should table some of these questions because we still have another measure to talk about or what do you guys think?

Male: Would it be possible if you're going to table to get all the questions and send them to us so we can address one by one?

Angela Franklin: Yes, NQF staff will compile the questions and send them out so that the developer can respond and we'll forward those responses from the developer to the committee or the – prior to the in-person meeting.

Male: OK, we appreciate that.

(Colleen McKernan): Great, thank you.

Angela Franklin: Next week, Yes.

(Catherine Roberts): Wait and this is (Catherine). Could we ask that when you send it to the – do send the questions so we can review those. And just in case there's any edits or any clarifications we need to make this is kind of a long, windy conversation …

Angela Franklin: Yes. We will certainly do that, absolutely.

(Catherine Roberts): They can start working on them but that might help us clarify some things too.

Angela Franklin: We will get that out as soon as possible to the workgroup.

(Catherine Roberts): Thank you.

Angela Franklin: As soon as we have the transcript from the call which will be at the beginning of next week.

(Catherine Roberts): OK, thanks.

Angela Franklin: So I also want to do a quick time check. I think I heard that we're going to move on to the next measure. I also want to do a quick time check with Jenna at NCQA to see if you're – if we go over a bit – if you are able to stay on call as well as the committee members.

Jenna Williams-Bader: Hi, this is Jenna. I am on the call and I might be able – I can stay a few minutes over. Not too long. I have to – I absolutely have a drop-dead need to get off at 3:30.

Angela Franklin: OK, great.

Female: OK.

Angela Franklin: So we'll try that. That's what we'll try to end before then. And also for the steering committee members that are on?

Zoher Ghogawala: This is Zoher Ghogawala, unfortunately I have a hard stop here in about ten minutes.

Angela Franklin: OK.

Female: OK.

Thiru Annaswamy: Thiru Annaswamy, I have to leave at 2:00 also.

Angela Franklin: OK.

Female: OK.

Catherine Roberts: As do I Cath Roberts.

Craig Butler: And as do I Craig Butler. Same here.

Katherine Gray: And as do I Katherine Gray.

Angela Franklin: OK.

Angela Franklin: All right, we'll move on to the next measure then that's measure 52 use of imaging studies for low back pain and NCQA is our measure developer who's on the call. And who's going to be the lead discussant for this?

Female: We do have Catherine.

Female: No?

Female: Roberts?

Female: Catherine Roberts left.

Female: Is Catherine Roberts back?

Catherine Roberts: I am back.

Female: Very good.

Catherine Roberts: I reached out some of these but I can go ahead and get us started.

Female: Yes, please.

Female: That would be great. Thank you.

Catherine Roberts: So this one is the percent of patients with low back pain without imaging studies. We can or without imaging within 28 days of diagnosis. It's a – it's interestingly done. And it's nicely done that it's an invertebrate. So it's, you know, one minus the number of people with imaging over, you know, all the people. So you get a percentage that how much is appropriate. So it'd be like a 75 percent appropriate. So if we go – let me go straight to the worksheets.

Importance to measure in reports. We have some comments. So 1a evidence to support the measure focus. We had a little bit – a little bit of a spread on how things were ranked. We had things – had it ranked moderate. Also had it ranked low and there was a question whether an analysis of the (hidden) data constitutes the systematic review.

1b was the performance gap shows that there is a room for improvement and there's disparity in this. And it's for priority – this is felt to be a high resource use and a high volume of patients. So what people thought about importance to measure in report, comments.

OK, so that was not too terribly controversial. OK, our criteria two, scientific (expectability) of measured properties.

We had an external comment from Dr. (Rueben) and he points out that there is potential to severely underestimate the information that's being tried – that they're trying to capture. So he points out that (inaudible) studies that would be coded with some of the diagnosis they list like lower back strain wouldn't really be captured as un-indicated. But they may well be un-indicated. And then he also points out that on MRI exams of the lumbar spine there's often especially in older patients some kind of finding. And say it's, you know a small disc herniation or disc protrusion.

And that he states that most practices will lift to that positive finding on the imaging plane. So instead of leaving it with low back pain which won't get paid for they actually would go back and change it and say, you know, arthritis or disc herniation which would be reimbursed. And thus that would make interpretation of the data challenging. So it's a concern that would not quite take what the measure of what they were looking to measure. Additional comments from the committee included the need to update the American College of Radiology reference to current version which is by (inaudible) set all in 2011. They also have two additional references that were not used in this one so that could considered.

The American College of Radiology has an (appropriateness) criteria that is based exactly on this imaging for low back pain. But basically additional exclusions that don't seem to be included on this measure. So it was suggested that if the measure developer could potentially comment on why those were not included or not excluded – included as an exclusion, that would be useful. Those are unexplained weight loss, (insidious) onset, unexplained fever, history of urinary or other infection, three immunosuppression and diabetes mellitus. Four, prolonged use of corticosteroids, osteoporosis and five, prior lumbar surgery.

Other comments, there's a concern that the denominator and the ability to accurately determine whether a patient has a negative diagnosis for low back pain in the six months prior. Reliability testing again varied by how algorithm 2 is applied either ranking it high or low insufficient. Validity testing again kind of spread the gamut to depending on how algorithm 3 was assessed high, moderate and sufficient. So it's a little bit of a challenge.

On the next page there's another comment, the measure excludes patients with a diagnosis of low back pain in the value set. But includes diagnosis other than nonspecific low back pain. I'm not sure if the other diagnosis including degenerative conditions and sprains/strain has to be included in the set because the (inaudible) for this measure is specific to nonspecific low back pain and not for sciatica, radiculopathy, or degenerative disorders of the spine causing low back pain and related disorders.

They comment that excluding neurologic impairment was really not sufficient in their opinion. Let's see, comments about them acknowledging data collection and calculation method, variability of making it a little challenging to have a reliable apple to apples comparison between health plans. And, let's see, comments on the statistical differences between 25th and 75th percentile. And then another person thing that immunosuppressed patients and patients with other risk factors for malignant low back pain do not seem to be excluded.

So several comments in that Section number 2, scientific acceptability of measure of property. So would anyone like to expand on their thoughts about some of these items that I've kind of summarized? It's a quiet group. Everybody is looking at the time. OK. Well, so it sounds like there are a couple of things that would be nice for the measure developers to respond to, to help us understand some of these challenges. Let's move on to criterion 3, feasibility. There's a couple of comments, no issues, data elements available, electronic claims. There is a cause and a burden associated with reporting measures. But given the prevalence, it appears to be also not really a lot of concerns voiced here, thoughts on feasibility?

Hearing none. I'll move on to criterion 4, usability and use. Let's see, a few comments reported usage only by the measure developer. Also those reports are widely distributed and significant to report at the health plan level. Question, could reporting at the provider level decrease in appropriate use with more substantial improvement? And there's another question, minimal improvement over the previous time period. So there – so are there adjustments that need to be made to increase the impact of this measure?

So usability and use thoughts. OK.

Male: Do we have a response from the measure developer on the question that you spoke?

Female: What? I mean …

Jenna Williams-Bader: Hi this is Jenna from NCQA. I've – we've actually covered a few different questions. And I didn't want to jump in because they never short on time. Which specific question would you like answered right now?

Catherine Roberts: Do you want to go back to the exclusions?

Angela Franklin: Let's start with the exclusion.

Jenna Williams-Bader: Super. OK. So, with regards to the American College of Radiology appropriateness criteria, they are really in line with what you're saying. Although they do have five additional exclusions that they list in the 2011 citations. And so we were wondering, would those, or why were those not included or, you know, are we perhaps overlooking them and they included things like unexplained weight loss, unexplained fever, immunosuppression, prolonged corticosteroid use, prior lumbar surgery.

Female: So, for some of the exclusions we actually looked at it when we did the original field test of the measure. And they had – in claims, you know, they had such minimal impact on the rate that we do not end up including them. So that's actually on page 13 of the test report.

Female: Hold on. Thank you. OK. All right. And then, now I'm flipping back and forth. The external comments regarding this, the potential to underestimate information …

Female: Right.

Female: … we're not able to capture additional unindicated studies. And people changing, which I was not aware was being done. People changing what the indications was for the imaging claim.

Female: Right. That particular question I'm going to have to go back to our coding staff and ask about that. We might have less of an issue for this measure because the age range is only up to age 50. So we're not including older adults. You might seem more likely to have indications that or other – abnormalities identified using the imaging study.

Female: OK.

Male: I would just question that, you know, that idea that the diagnosis on there as a reason for request in MRI is changed once we get the reading. I don't know, we're in the process with anybody's relevance cycle they would do that and including your radiology group.

Female: I'm in the radiology group. I have no knowledge of that being done.

Male: Yes. And – that sort of thing. I know it's not done on a clinical basis because I'm an orthopedic surgeon. I go back and I change the reason I request an MRI that's out the door and somebody else did it after I ordered the test. So, I just wonder about the validity of this or the occurrence rate of this actually happening as being something we need to bog down on.

Female: That's a good comment. I suppose we can actually look on our own or I can actually circle back around with the person. I've, I know who made that comment. And see if there's any kind of reference for us to refer to, to get a sense of how big a problem that is.

Male: Now, if they can …

Angela Franklin: So, other questions?

Female: You know, the only questions that come to my mind is that, for us internally we've got quite a range for how we score reliability and validity testing. So, you know, probably we might want a little more discussion on that at some point in the future. And it is 12:00 noon.

Male: Yes.

Female: Yes.

Male: You know, those algorithms read in plain English obviously. But, how we each individually interpret some of the words and act on and decide. I think with what varies. And, perhaps the, an opportunity to kind of walkthrough one of the measures through the algorithm and understand, you know, how people are interpreting those so we get on the same page and do the same

interpretation on the words that are in those boxes and other than diagram might be helpful. But, that's why I think a lot of the variability might be coming in.

Angela Franklin: Yes. And this is Angela. We understand if there's going to be a variability between how members interpret that. And that's what, you know, is part of the process. We also have some additional guidance that is being drafted this week that we will also provide to steering committee members probably the next week or so regarding how to better guide you through the algorithm. But we do understand and appreciate the variance in that part of the process.

Female: OK.

Angela Franklin: Were there additional burning questions for the developer about this measure?

Female: So it does sound like we're going to end on time then. We'll go ahead and work on some of the additional questions, it sounds like for the measure we talked about, 0514.

Angela Franklin: That's correct.

Female: And …

Angela Franklin: This measure as well, if there are additional questions …

Female: Yes.

Angela Franklin: … we will be putting together a summary of this discussion. And, so a little bit out of our process and make sure that we solicited questions that we can then gather by e-mail and send to the developers as well as to the committee for any editing. And make sure we get those responses to you before the next or before the in-person meeting.

Female: For the in-person meeting. Thank you all.

Female: Thank you.

Female: Thank you.

Male: Thank you.

Female: Bye-bye.

Female: Bye.

Angela Franklin: Thank you all.

END