NATIONAL QUALITY FORUM

Moderator: Cancer -
May 5, 2016
1:00 p.m. ET

OPERATOR: This is conference #: 86686333.

Amber Sterling: Hi, good afternoon and welcome to Workgroup Call 2. I am Amber Sterling. I'm the project manager for the Cancer Project.

In the room with me I have my colleagues, Melissa Mariñelarena, our senior director; Shaconna Gorham, our senior project manager; and Kaitlynn Robinson-Ector, our project analyst. This is a preliminary analysis, not quite.

Before we jump into the meat of today's discussion, I would like to take a roll call. So do we have Karen Fields on the line?

Karen Fields: Yes.

Amber Sterling: Benjamin Movsas?

Benjamin Movsas: Hello.

Amber Sterling: Hi.

Beverly Reigle?

Beverly Reigle: Yes, I'm here.

Amber Sterling: Great.

Crawford Clay?

Crawford Clay: Here.

Amber Sterling: Brent Braveman?

Brent Braveman: Here.

Amber Sterling: Jette Hogenmiller?

Jette Hogenmiller: Here.

Amber Sterling: And also do we have our developer colleagues on the line from the American College of Surgeons?

(Erica McNamara): Yes, there's a group of us here. And I believe we also have a clinician on the phone as well.

Amber Sterling: OK, great. Was that (Erica)?

(Erica McNamara): Yes. Sorry, this is (Erica).

Amber Sterling: OK, great. Thank you so much.

(Erica McNamara): Thanks.

Amber Sterling: If we have any additional committee members on the line, you're welcome to introduce yourselves. I'm not sure if we do.

Matt Facktor: Hi, this is Matt Facktor.

Shelley Fuld Nasso: Hi, this is Shelley Fuld Nasso. I'm a committee member.

Amber Sterling: Hi, Shelley.

Shelley Fuld Nasso: Hello.

Matt Facktor: Hi, it's Matt Facktor, a thoracic surgeon from Geisinger, also a committee member.

Amber Sterling: Hi, Matt.

All right. It sounds like that might be it, so we'll go ahead and get started.

If you can just mute your lines or if you're on the webinar if you can actually mute your laptop, that way we won't have as much feedback.

All right. So the purpose of today's call is to allow the Standing Committee members to have initial discussions about the measures you will all be evaluating at the in-person meeting on May 18th and 19th. Many of you are new to the NQF process, and this is an opportunity to ask questions about the criteria, the expectations of you as a committee member, and anything else that may come up.

Remember, we do have the developers on the line and they are here to answer questions that you may have based on your review of the staff preliminary analysis worksheets. Additionally, if you simply need clarification on something in the measure, they can also provide that so you're welcome to -- sorry, you're welcome to ask some questions.

The way this call works is that NQF staff will introduce the measure. We will be screen-sharing each of the measure worksheets, so you should be able to see each measure as we go through them if you are on the webinar.

After we introduce the measure, we will turn it over to the lead discussants to direct the conversation.

Just as a note, at the in-person meeting, this will be the role of the co-chairs, to kind of lead this discussion.

The lead discussants will summarize the information presented by the developer and summarize the comments submitted by the other Standing Committee members. These comments are in the sort of peachy pink section of your worksheet in the area that's labeled Committee Pre-evaluation Comment.

So we'll get started with our first measure, which is 0219. The title of this measure is the Post Breast Conservation Surgery Irradiation Measure, the

measure steward is the American College of Surgeons, and our lead discussants for this measure all right Karen Fields and Benjamin Movsas.

So, Karen and Ben, I'll go ahead and turn it over to you so you can give us a brief summary of the measure.

Karen Fields: Ben, do you want me to start or do you want to go first?

Benjamin Movsas: So, Karen, why don't you start?

Karen Fields: OK. So this is a measure that is a percentage of female patients aged 18 through 69 who have their first diagnosis of breast cancer, AJCC stage 1 through 3 receiving breast conservation surgery who also received radiation therapy within one year of diagnosis. It's a facility level analysis, and it's a process measure. And its maintenance -- this is I believe the third -- well, the second maintenance for this measure.

The authors presented no new or the stewards, the measure stewards presented no new data at the time of the submission. However, they did note the previous data, which is Category 1 High Level Evidence that the role of radiation to the tumor bed is -- decreases the risk of recurrence. Of note, it's been long demonstrated that there's a 75 percent reduction in the risk of local recurrence for radiation, so I think this is an important quality measure.

The Committee was asked if since there's no new evidence presented. Does it require a re-discussion or a voting? I don't -- I don't think it requires any re-discussion. This is many large randomized trials with years and years -- decades of follow-up to demonstrate the importance of post surgical radiation in these patients. And the guidance from the algorithm is that this is moderate based on the evidence since no new evidence was presented. Although it's been previously noted to be high level evidence, so I would rate the evidence as moderate.

Do you want to go through all of the comments all the way through or do we want to stop at each point and have a discussion with the committee?

Amber Sterling: Yes, let's go ahead and stop there, Karen. We can -- excuse me -- ask Ben if you have anything you want to add to what Karen said and if not, then we can go ahead and open it up to the Committee if anyone has any additional questions or discussion about this first section.

Benjamin Movsas: No, that was a great summary.

Amber Sterling: Great. Any other committee members have anything they want to add about the summary of the measure?

All right. If not, which it seems like not, we will go onto our first real criteria, which is the importance to measure and report. And this section consists of two sub-criteria, which is evidence and opportunity for improvement/gap and disparities.

Karen Fields: So, moving along, there was no new data presented since the last maintenance process. And the reason the developers stated that no new performance data was provided was because the adjuvant therapy data is likely to be incomplete for the most recent year until all the programs have had time to collect this information.

Previously, when this was reviewed, there -- the committee did note a moderate gap in performance. Although in 2012, the mean performance rate was 90.7 percent, which is probably approaching maxing out on that gap. Clinically there's no reason for it to not be 100 percent since that's part of the treatment plan. Disparity gaps were seen in multiple groups and what were presented in this application. And it also looked at Medicaid and uninsured patient's low income, low educational levels, and the type of facilities, seeing gaps there as well. So I think it's a very important measure.

As I said, if the prospective plan is a lumpectomy with radiation, then all patients should -- 100 percent should be the target. So I'd say a high measure.

My only other question for the developers may be -- would be, are we going to start to include some of the newer kinds of radiation schedules? There's some -- one of the things is (brachytherapy) as one of the measures, but inter-

op radiation or other kinds of measures like that for the future to continue to make this measure relevant for the future.

Benjamin Movsas: And this is Ben. I would just add to that. You know, the use of hypofractionated or abbreviated breast radiation as opposed to maybe more conventional, you know, prior -- you know, the longer course that might be something -- that might be of interest perhaps moving forward.

The other question I had is if this is looking to see what percentage of patients actually receive breast radiation within that timeframe, but is there any effort to see whether or not the radiation oncologist actually recommended it? I mean, if a patient declines the recommendation, does that count the same kind of towards non-compliance, quote unquote, "as if the radiation oncologist never made the recommendation."

Amber Sterling: So I think …

Benjamin Movsas: Because obviously we can't always control for what patients decide to do. We can -- we can strongly recommend something, but if a patient decides they're not going to do it, you know, there's always going to be a percent of patients that don't follow the recommended plan.

Amber Sterling: Sure, absolutely. I mean, if the developer wants to respond to either Karen Fields or Ben Movsas' questions, you're welcome to do so.

(Erica McNamara): OK. Thank you and thank you for reviewing this measure. I did want to note that the data that we provided even though both for 2012 is actually new data. It's the most recent data that we had available at the time of this application.

So as far as new radiation therapy, the way that we collect information through the CoC or Commission on Cancer-accredited centers is through cancer registries. And so they collect radiation therapy. They don't collect the specific agents of radiation therapy. So this measure includes any radiation therapy that is administered.

And then the recommended versus administered therapy, this measure does not -- patients are considered non-compliant with this measure if treatment is recommended but not received. And we do get a lot of feedback on that, but you can see that compliance is still extremely high, and we do know that there are a proportion of patients who -- a small proportion of patients who plan to have radiation therapy and as much as it is recommended may not. And we can ask our clinicians to speak more to that at the in-person meeting.

Karen Fields: Yes. Do you have a target of topped out in the measure or the goal is try to get to a 100 percent?

(Erica McNamara): We would like that. Yes, so we hope, right, to get to a 100 percent, but realizing that that's not …

Karen Fields: OK.

(Erica McNamara): … yes, so …

Karen Fields: That just helps us answer the question about is it still remain important to measure if that is new data. And I didn't understand that from the application, so I apologize.

(Erica McNamara): No, thank you. The other thing that we were concerned about is that we do see some disparity still occurring, so we think that needs to be also tackled.

Benjamin Movsas: I mean, to me, that's probably the most compelling reason to continue studying this is that there's a gap based upon disparities, which, of course, is always a concern. And so by continuing to collect it hopefully that will help to address that issue.

Karen Fields: Don't worry.

(Erica McNamara): Thank you. I appreciate it. We do agree with you.

Amber Sterling: OK, great. Thanks so much.

Excuse me. So now we are going to move on to scientific acceptability. The first part is reliability, and here we're looking at two aspects. First, we're

looking at the specifications, specifically how well are they designed. And then we'll be looking at the reliability testing and whether the testing used is appropriate for the measure.

Lead discussants, you can go ahead and provide your feedback on the specifications and reliability testing.

Karen Fields: Ben, do you want me to continue? I'm happy to.

Benjamin Movsas: You're doing such a great job. Keep going.

Karen Fields: OK. All right. So, the specifications, I think, are described. Everyone has them in front of them, and I think that those are -- will delineate the numerator and denominator. And so I don't have anything to add. I think that they meet the criteria.

The data source would be the tumor registry or in paper or electronic medical record, so obviously it's not at all electronic measure so that, obviously, it's something to talk about later when we get to feasibility. And -- but I don't see any issues with specifications or coding.

I reviewed the algorithm. I had no questions about that, and I think that this could be consistently implemented. So I thought the specifications were fine. My only question remains will the developers be adding some of the new treatment strategies that we both discussed.

And then reliability testing, again the original data used and how the reliability testing was conducted was presented. And the new -- the developer updated the performing rate -- performance rates from 2012 for all of the CoC-accredited cancer programs and noted that in the tenth percentile there still remained 81.1 percent or less alignment with a range from zero to 98 suggesting that this measure still remains important to measure.

The only comment and I need help, and I put this in my little evaluation as well. I need help from the committee to understand how to interpret that this reliability testing and all of the American College of Surgeon testing was -- doesn't meet the current standards for NQF reliability requirements and that

they are measuring only two of the variables. And so the guidance from the reliability algorithm becomes insufficient to assess, so I'm assuming that we'll just vote on whether or not we think the right reliability data is adequate to pass this measure. Otherwise, I'm at a lost about how to use that algorithm.

Melissa Mariñelarena: Sure, Karen, this is Melissa. So when we have reliability data, patient level testing for reliability, we can actually rely on validity. And even then we said it was insufficient just based on the information that was provided to us at the time and that is because we do ask for validity testing on all the data elements or, at the minimum, numerator, denominator exclusion, a little bit of more information that was provided to us.

So when we have data element validity testing, percent agreement, kappa scores, then we apply those -- that rating to reliability, so that would have applied to the reliability rating. But if we go down to validity, we got some information. It just wasn't enough. So we -- we're not saying that they're not valid, they're not reliable. We're just saying that what was in front of us was not enough information.

So if we go down to the testing results, this is we got that -- we got some percentage agreement for two data elements that were included in the numerator. It looks like timing of radiation therapy and therapy recommended but not received. But again we ask, based on our criteria, that all -- we asked for information on all the critical data elements. And so we didn't have that information.

And then we only received percent agreement for these two data elements, and it looks like -- it said 91 -- let's see. It's -- the last word says, "They provided percentage agreement statistics. While these are fairly high for radiation therapy, a percentage agreement of 59 percent for missing radiation therapy." That was a little concerning because that seems low.

And then there were no additional results were provided besides that. So, you know, the developers, they like to respond or they have additional information, the Committee also can provide additional information. Your experience with these measures since they are used in registries, you know,

how easy are they to obtain the data elements of the information so you can see to that.

Karen Fields: And so to go on to the incomplete radiation data, that's a chronic problem in the registry, and they talked about their strategy to identify those patients, flag them and ask for follow-up. However, they didn't tell us their success on answering those questions and were they able to still tell if the low compliance with data, 59 percent, would affect the validity. So my comments were going to be I wanted to ask the developers those questions that you outlined, and I thought that the committee needed to discuss this and decide if we would revote on the validity of the testing.

We use this obviously and report this data regularly. But I think that there are some gaps certainly and was noticeable in the radiation data. So, comments from the developer.

And, Ben, if you want to weigh in.

(Erica McNamara): So, yes. Oh, Ben, did you want to go first and then we can …

Benjamin Movsas: No, no. Why don't you go first? Thank you.

(Erica McNamara): OK. So we did provide the information from previous testing results. Previously when the reliability has been tested and the previous measure submission has been rated as high or moderate, and so we originally just submitted the previous results and added to it based on some additional information that we had.

We did select outliers purposefully because we wanted to make the best use of improving the data within the data capture. So that's why maybe for some of the information we're going to see as high of -- you don't -- you don't see it to be exactly 100 percent compliance. You know, we're not seeing direct in our interrater reliability based on just the cases that we selected.

And as I said, we did that purposefully and we've seen overtime capture this information grow. That's part of the reason that we report back to our facilities and give them time to update the adjuvant therapy information. We

point out those cases so that they can go and look and confirm whether or not treatment was received or not.

We also have provided the original sensitivity and specificity -- sensitivity testing that we did when the measures were originally submitted to the NQF, which looks more at can do the measures and the performance rates show changes across the board, so more of the rate testing rather than just the height on testing. And we could provide that -- those original sensitivity results to the committee as well.

Benjamin Movsas: Yes. So my thought is that this kind of -- as noted, this is an evolving process. It's gradually improving in terms of, you know, the reliability, et cetera, although there are still some gaps. But we probably -- we probably have better data now than we had, you know, two to three or four to five years ago. And so I think at the end of the day, my feeling is that, you know, we still want to pursue this and continue that process improvement even if we understand that there are some gaps.

(Erica McNamara): And …

Karen Fields: Go ahead.

(Erica McNamara): I would say we would wholeheartedly agree. One of the reasons that we have released one of our newer quality tools, which is called the Rapid Quality Reporting System, is so that we actually provide alerts to our program saying that, you know, that this patient should receive radiation therapy and to give them an additional avenue to track that care both to receive it and get it into the cancer registry, but also to provide it to cancer committees as well.

And with that, we have seen and increased in the capture of these variables.

Karen Fields: So my main question was to the NQF staff, I mean, we -- this is the first time we've evaluated these two measures -- validity and reliability -- using the new algorithms. And it sort of puts us at a loss about how we're supposed to interpret that. I'm assuming going forward the next time this measure comes up for maintenance, everybody is going to be used to what your algorithm

needs to have to run. So that's my main question to the NQF staff. I agree that it's important to measure.

I don't -- if this is being used as -- for payment models right now, I think that would be more important for us to be concerned that there was -- that were these reliable and valid measures and was the methodology sufficient, too, for payment models, but I don't think it's reporting but not payments that I'm aware of right now.

Comments from the NQF staff or from the developers?

Female: It's not used for payments. It's used for quality improvement.

(Erica McNamara): Public reporting.

Female: We do hope …

Karen Fields: That means -- yes, so I think that over time then everybody is going to need to shift to different ways to present data for measure updates then, correct?

Melissa Mariñelarena: Correct. This is Melissa at NQF.

Karen Fields: OK.

Melissa Mariñelarena: Yes. Our criteria has changed quite a bit even some from -- some of the measures that we have from 2012. Some -- you know, this is from 2007, 2008, so there have been a lot of changes. And for the maintenance process, we don't require new testing. We do ask that the testing that is available that it needs our current criteria, but then we also -- you know, it's up to the committee if you -- if this is acceptable to you then that's fine, you know, you vote on that.

Again, you know these measures and how it's captured in the registry. Sometimes the argument is that with registry measures, they are -- it's different because it's easier to capture the data elements. Again, we look to you and your expertise. And you can also make the recommendation to the developers that, you know, next time they came around that you would like to

see some new testing. You know, those are all discussions that we could have at the meeting.

(Erica), does have some additional data that -- and I know you sent some stuff the other day with our -- I think it was you.

Female: Yes.

Melissa Mariñelarena: And we haven't had a chance to look at, but we can also include that and provide it to the committee as well prior to the meeting.

Female: Yes.

Karen Fields: And that's fine. I'll -- I'm done with the conversation. I just -- going forward, all -- most of these measures had the same ongoing question, and I need to know how we wanted to discuss it. So, I'm fine with moving forward. Maybe presenting to the group, the old sensitivity and specificity testing would be helpful for us to move forward.

From a -- OK, go ahead.

(Erica McNamara): Oh, no, I'll just say and thank you for that discussion. We kind of went by the guidance but we didn't have to redo the reliability unless we've already done testing and so that's why we didn't submit anything new. But if you -- you know, so I …

(Crosstalk)

Karen Fields: Yes, the algorithm -- using the algorithm and looking at the old way we always evaluated these things. There's a -- we have to figure out how to bridge that gap, and that's the only reason I bring it up so.

So, from a feasibility standpoint, obviously, it's -- there's some feasibility concerns always with tumor registry data because it still largely requires hand abstraction, which can be burdensome and fraught with potential errors. However, all of us understand the strengths and limitations of tumor registry data, and I still think that this is a feasible measure. And I would rate it as moderate.

Usability, I think that it has high usability. I think that there are -- it is being publicly reported. It is frequently used by many programs internally for quality reporting. My group uses it as well, so I feel that it's highly usable and important to report measure. So I would rate both of those moderate and high and defer for any other discussion.

Amber Sterling: Great. Thank you so much, Karen.

Ben, if you have anything to add about feasibility or usability or use?

Benjamin Movsas: No, I agree with that assessment. So, yes, I feel comfortable with that.

Amber Sterling: OK. Are there any other committee members that would like to add anything on measure 0219 or ask any questions of the developer for 0219?

All right. Hearing none, we will go ahead and we will proceed forward. We're going to be looking at 0220. This is adjuvant hormonal therapy once again from the American College of Surgeons.

Our lead discussants on this are Karen Fields and Beverly Reigle. I don't know if Karen or Beverly wants to kick us off, but if you could go ahead and do a short summary of the measure.

Karen Fields: I'm going to let Beverly go first.

Amber Sterling: That is great.

(Crosstalk)

Beverly Reigle: Karen, I hope you'll jump in. Well, the title of this is Adjuvant Hormonal Therapy and the description is it is the number and percent of female patients who are 18 years of age or older at the first diagnosis of epithelial type breast cancer with the stage T1cN0M0 (NCRO) no nodes, no metastases. Primary tumor is ER/PR-positive with either tamoxifen or third generation aromatase inhibitors administered within one year or, in parenthesis, 365 days of diagnosis. It is a process measure. This is the third time that it is being reviewed. The first was 2007, again in 220 -- I mean, 2012.

The evidence has been excellent in terms of the level of evidence. It's systematic reviews with randomized controlled trials. And the systematic reviews were specifically meta analyses demonstrating a 25 percent reduction in risk of distinct cancer recurrence and death. This is obviously supported by National Conference of Cancer Network and is in their algorithm of care.

Should I continue or is that enough of description?

Amber Sterling: Yes, I think that's …

(Crosstalk)

(Erica McNamara): Karen, do you want to add, I mean, at this point, I didn't know if you wanted me to go into the rest about the improvement and …

(Crosstalk)

Amber Sterling: We can pause there for a second and see if Karen or any other committee members have anything to add to the summary.

Karen Fields: No, the evidence is some of the strongest evidence we have for adjuvant therapy. And I -- I don't think it needs to be re-reviewed or discussed so …

Amber Sterling: OK, great.

Karen Fields: I think I agree.

Amber Sterling: Great. Fantastic. If nobody else has anything to add, then we can move forward to opportunity for improvement and disparities.

Beverly Reigle: Actually, from 2008 to 2012, that did demonstrate improvement in the performance rating, but there are still disparities and primarily in age, ethnicity, insurance status.

So, as it seems to me with most of the instruments or a lot of these, obviously, higher in the non-Hispanic White, lower in Hispanic and non-Hispanic Black.

So I don't know if you -- anybody wants to add to that aspect in terms of the improvement.

Do developers have anything they want to add, this newer information? I didn't see any new evidence either in terms of support of this. I think the current evidence is, you know, have been very -- has been very supportive in terms of the actual tool.

(Erica McNamara): Yes, we would -- we would agree and just -- that's why we didn't provide any new evidence. But as with the previous measure, we do still see disparities in this, and we feel like that makes it even more important to obtain even though there is still room for improvement in the overall compliance as well.

(Crosstalk)

Karen Fields: I think this is …

Beverly Reigle: … Karen or anyone else have …

Karen Fields: Yes, I think this is -- I think there is ongoing evidence that this isn't important to measure. And I think sometimes it's also difficult to interpret and get this one down at some of the centers that try to use this data, so continued measuring of this data to really meet the letter of how to report this data, I think, is important.

Amber Sterling: OK, great. Thanks so much. Are there any other committee members besides the lead discussants that want to add anything on opportunity for improvement or disparities?

Jette Hogenmiller: This is Jette Hogenmiller, and I just wanted to what extent do we access data that gives us an idea of what are the impediments doing this case, you know, a lot from -- we got the 92 percent. Do we collect data that really provides us insight to what are the shortcomings to -- you know, what are the impediments, either measuring it or treatment adhering 100 percent over 90?

Amber Sterling: I think if the developers want to respond to that question, that would be appropriate.

(Erica McNamara): I don't think it ...

(Off-Mic)

(Erica McNamara): So we don't -- in the data that we collect, we don't actually have specific reasons why hormone wasn't received or why we're not seeing complete compliance with this measure.

Some of our speculation has to deal with based on where the patients are, where we see disparities sort of in the last and Hispanic and other races is based on maybe just not having the follow-up information for those patients, but we don't actually know. That's just one of our speculations.

Beverly Reigle: In working with some of these hospitals that are in these areas, what they have found has been that they have throwaway zones and addresses that are hard to contact with, so that has been a problem. And one of the hospitals I've been worked with has been trying so very hard to improve this collection and trying to help them not feel threatened to give follow-up information, how they can be reached.

Karen Fields: Perfect. Just speaking from practical experience, sometimes the handoffs between rad onc, and med onc, and chemotherapy transitions, and then adding hormonal therapy make a lot of -- create a lot of challenges for making sure that the patients get started on their hormonal therapy as well just to -- there's a lot of therapies for some of the patients that have to be well-coordinated.

Jette Hogenmiller: Well, and I think in general as we're looking at different measures that the question we have to ask in terms of process because is there something we can also capture to provide insights so we can elevate that measure. I mean, should that be part of a discussion. So just kind of throwing out there for thoughts.

Amber Sterling: I think these are some really great discussion points, these are definitely some things that we may want to keep in mind as we move forward to the in-person

meeting when we're talking about, you know, gaps and measurement and sort of some areas or maybe our measures aren't quite meeting your expectations at this point. So thank you so much for those comments.

All right. I think that we can go ahead and move forward onto reliability and relatively testing. If Karen and Beverly want to continue on.

Beverly Reigle: I have a question first on the specification just real …

Amber Sterling: Sure.

Beverly Reigle: … quickly. And I think it's just my researcher brain.

I was looking at the numerator/denominator, so the -- mainly in terms of the inclusion and exclusion. I found it interesting that exclusion criteria and perhaps this is the way it's done for these types of measures is exactly the opposite of the inclusion. Typically, your inclusion -- any exclusion would be anything that in that particular group of people that you've already included, anything that you actually would want to exclude, so it wouldn't be the opposite.

Is this typical to have this? I mean, it seems like it's in all the reports this way so I just wondered.

(Erica McNamara): I think that's sort of a legacy that we continued on for when we first submitted that we wanted to make sure that, you know, as far as how we -- you look at the flow it doesn't say men or women. It's just something that we provided for extra clarification within the exclusion documentation for the NQF.

Beverly Reigle: And that's typical for NQF …

(Erica McNamara): I …

Beverly Reigle: … that …

(Erica McNamara): … I think it's typical for us.

Beverly Reigle: Oh, oh. Well, actually it's in -- I've noticed it in the other reports that we reviewed, and that's why I was asking about that, so I don't want to take a lot of time on that. I just was -- it's more of a learning curve for me so …

(Erica McNamara): Oh, OK, then -- I'm sorry, I can't speak to that, but we just kind of continued on what has been done in previous submissions.

Beverly Reigle: All right. Well, we'll go on to reliability and validity.

I think I can say the same thing. It's very much the same as our previous report, Karen, if I can refer to you in that using the algorithm that we were given, it does end up to be insufficient. And yet it seems to me to be very, very important. I think it went into really looking at the validity of the measure.

And personally I thought it was a very important value. I think I wasn't following very well the algorithm that it was -- I mean, I understand that there were only two elements in this particular one. It was timing up hormone therapy and it's recommended but not received or the two elements it appeared to be looking at or testing and all elements were to be tested, so therefore it was considered insufficient both for reliability and validity. So in a way, it sort of goes back to what we were talking about earlier.

I don't know if you want to add more things to that. I didn't really see anything other than the algorithm. It doesn't seem to be the source other than that aspect, so I don't -- Karen, I don't know if you want to add something to that specifics.

Karen Fields: I'll just comment to the reviewers that it's -- this one, to me, when we're trying to evaluate this data and report it is the recommended but not received, and then the timing of when a patient receives it before or after the 365 days always create some confusion. We've tried to use this internally and then among some of our external sites that we collaborate with and there's always a challenge in that, so more -- testing more variables might continue to help make sure that this looks like a valid measure that everybody interprets in a similar way.

I think there's -- people are getting used to this measure using it more and more and understanding how to report this measure, but I just always find this measure harder to interpret among our sites. And maybe I just have a -- have the collection of sites that has more trouble interpreting, so I don't know if the measures are thinking that I am -- if they have that same kind of experience with this measure.

Amber Sterling: Great. Thank you so much for that feedback, Karen. I think that's something that's really important for the developer to hear. That's certainly something that they can take back to the drawing board with them. And I doubt that you're the only one that has trouble with, you know, any particular measure.

(Erica McNamara): Yes, and I think just -- Amber, can I make a quick comment?

Amber Sterling: Please.

(Erica McNamara): OK. So I think this is one of the measures that we really see a lot of improvement in the collection for programs that are using our rapid reporting system because it does kind of allow programs to start looking for this information in a sooner time period and have a notifications that hormone therapy is needed not just so that the patient will be able to receive the treatment, which we obviously is the most important, but the registry can capture it. But we do know that hormone therapy capture is not necessarily completed within the cancer program, so it's a little more -- it's not as readily available to all programs as a surgical procedure.

So, thank you. That's very helpful to us.

(Crosstalk)

Beverly Reigle: And in terms …

Karen Fields: Well, would you use it just to …

Beverly Reigle: … of the statistical test, I may have misread this, but it looks like the -- other kinds of statistical tests were -- are perhaps encouraged like the kappa. Is that -- am I reading that correctly? Staff, this is I think are the NQF staff.

Melissa Mariñelarena: Hi, this is Melissa from NQF. Yes, that is correct. That is encouraged. We say that percent agreement is not sufficient. It's just not enough. Sometimes -- again, that's a judgment call. Sometimes if we have at least the percent agreement for all of the data elements, then we would have a little bit more information to go by, again not -- because we don't use the measure, we appeal and so hearing that information from Karen is very helpful. But, you know, when we looked at this and reviewing it, we saw percent agreement of 84 and 79 for two data elements, you know, based off of that. That's what we had. There was no kappa scores. Just -- there just wasn't enough information to come up with a rating for it. That's all.

If we had kappas on, at least some of the critical data elements that we talk about and then some of these -- I mean, it would be difficult to have scores for -- if we would ask for kappa for everything, there is one, two, three, four -- you know, five data elements in the numerator, 10 in the denominator. And really the exclusions are opposite of the denominator. You know, we wouldn't necessarily expect that.

Beverly Reigle: Right. So that's very helpful. Thank you very much.

So, did you want to go on to the feasibility? It really sounds very much like our previous one in terms of have the data collection through electronic records, but also manual paper extractions also. So, it seems -- I mean, it seems to be working. I don't -- and our committee comments are pretty positive in terms of stating it. It's rated as moderate so by one -- by one person.

Karen, I don't know if you want to add to that.

Karen Fields: Well, I keep the -- I would keep the moderate rating just because it's hand-collected data that -- and until we can move to all electronic data …

Beverly Reigle: Right.

Karen Fields: … it still poses a little bit of a feasibility gap. I don't think we should stop measuring it at all.

Beverly Reigle: I agree with that.

Amber Sterling: Great. Are there any other committee members that want to add anything about feasibility before we move on to usability and use?

OK. If not, we will go ahead and move forward to usability and use.

Beverly Reigle: Again, I think it's very much the same and certainly being used by several entities for accountability and performance improvement, CoC, different ones. So I think it's highly used and, you know, I don't see any -- nothing from the committee came back that was a negative. I don't know if anyone wants to add to that.

Amber Sterling: Great.

Karen Fields: No, I agree. It's very usable metric and I personally use it in our own institution for performance improvement strategies and we instituted a call back system because of this measure, so I think it's valuable.

Amber Sterling: Great. Thank you so much. Great, we're through (two) measures. Is there anybody that wants to add anything about this particular measure, 0220?

All right. If not, we will move on to 0559, the measure title is, combination chemotherapy is recommended or administered within four months, 120 days, of diagnosis for women under 70 with AJCC T1cN0MO or stage 1B through III hormone receptor negative breast cancer.

Again, this is American College of Surgeons' measure. And our lead discussions are Beverly Reigle and Benjamin Movsas. So Beverly or Benjamin, I will turn it over to you to give us a brief summary of the measure. And I just want to mention that that means you don't have to go into evidence yet, just do a brief summary and then we'll hit evidence coming up.

Beverly Reigle: I will turn it over to Ben if that's OK.

Benjamin Movsas: Oh, sure. Beverly, are you -- what's your specialty?

Beverly Reigle: I'm a faculty at the College of Nursing and breast cancer is …

Benjamin Movsas: OK, sure.

(Crosstalk)

Beverly Reigle: … area, I'm the director …

(Crosstalk)

Benjamin Movsas: OK, great. OK, so I'm happy to start it off. OK, so this measure is combination chemo is recommended or administered within four months of diagnosis for women under 70 with AJCC T1cN0MO or stage 1B through III hormone receptor negative breast cancer.

And you can see listed there the numerator and denominator statements. And this is again one that is a maintenance of a process evaluation and the level of evidence. We said we're going to wait on that but basically it's very high. So anything -- any questions so far about this one? Do you want me to go over the evidence next?

Amber Sterling: Hold on just one second. Beverly or anyone else have anything to add about the summary?

Beverly Reigle: No, I mean I think that's exactly what it is.

Amber Sterling: OK, perfect, then go ahead and …

Benjamin Movsas: OK. So the level of evidence is category 1, it's very high. There is many randomized trials as noted that showed approximately 1/3 reduction in the risk of distant cancer recurrence or death.

So this is obviously a process that is very important. And the developer has attested that the underlying evidence of the measure is not changing for the last NQF endorsement. So I think we know there's high level of evidence.

Do you want me to move on to the performance gap analysis?

Amber Sterling: Sure, that sounds great. Go ahead.

Benjamin Movsas: OK. Please interrupt me at any point if anybody has -- I'm from New York originally so I talk fast. So if anybody has anything to add or interrupt, please do so at any time.

I think the gap shows that overall there has been improvement from 2008, 85 percent up to 89.4 percent in 2013, so clearly, we're moving in the right direction which is good, but again, like similar to other measures that we've looked at when you look at disparity, there's still a gap based upon disparities as shown by race or ethnicity, as shown by insurance status, which to me suggests that this is still something where we have room -- quite a bit of room for improvement. The comments were again from the committee that this is high evidence.

There was a question about why HER2-positive is not clearly stated in the numerator since that is part of the NCCN guidelines to this population. I thought the answer was because we're collecting whether it's HER2-positive or negative, so it's kind of either way. Is that the answer to that question or did somebody else have more insight on that?

Amber Sterling: (Erica), if you, guys, want to answer, that would be really helpful.

(Erica McNamara): Yes. So we don't look at HER2 status for what our (nano) patients should receive the multi adjuvant chemotherapy. We do assess the receipt of immunotherapies are HER2-targeted therapy. And I apologize if that wasn't clearly -- completely clearly explained, but, no, you were correct.

Benjamin Movsas: OK. Is there any other question from the committee on -- committee members on that issue? Otherwise, I think that from my perspective, you know, it's still a very important measure to continue to analyze because of its high level of evidence and the fact that there is performance gap particularly related to disparities.

So next, I guess we should move on to the reliability validity discussion unless there is any other questions?

Amber Sterling: Beverly, did you have anything that you wanted to add before we move on to reliability?

Beverly Reigle: No, my question was answered, so.

Amber Sterling: OK, great. Then you can proceed.

Benjamin Movsas: OK. So very similar issues to what we have addressed in the past and that is that overall, let's get to the page I want to be looking at here. So I think that overall comments, and I would agree that the data is routinely used during patient care and should be available in the medical record. Oh, wait, I think I skipped the feasibility, hold on a second.

Amber Sterling: Yes, I think you jumped ahead. We're looking at reliability.

Benjamin Movsas: I jumped ahead of the -- yes.

Amber Sterling: That's validity.

Benjamin Movsas: So -- oh, here we go. I see it now. The specifications, I would agree, are clearly defined and appropriate with no particular concerns. Somebody made a comment about, should the numerator elements -- it says, should recommend, replaced, considered where histologies were deleted and not supported by evidence. So shall we discuss that for a moment?

Amber Sterling: Yes, so let's discuss that.

Benjamin Movsas: OK. So the numerator before it said considered and now it says recommended, is there a particular significance, are wee saying because it's stronger like we feel more confident about it, so we're saying recommended or am I reading it to this too much?

(Erica McNamara): Is that a question for us as the developer?

Benjamin Movsas: Yes, I think so, yes.

(Erica McNamara): So we clearly changed that based on the registry codes that are used, so the registry codes that are used to say recommended.

Benjamin Movsas: OK. I mean, how comfortable is that but I just …

(Crosstalk)

(Erica McNamara): … it was consistent.

Benjamin Movsas: OK. And where histologies were deleted, I think that's consistent with what we've done before. Somebody wrote as a comment not supported by evidence, so is there anybody that is concerned about that? That wasn't my comment but does anybody want to discuss that? If not, I think …

Beverly Reigle: It's not my comment. I don't …

Benjamin Movsas: OK, fine.

Beverly Reigle: At least not worded that way. I did wonder about the deletions but I didn't say it wasn't supported.

Benjamin Movsas: OK, OK. So I mean to me, it was just kind of like similar to what we've done before, so it's consistent where histologies were deleted previously maybe because we don't have as much evidence for those situations.

Beverly Reigle: Right.

Benjamin Movsas: So I'm OK with that, but if there's any other thoughts, let me know.

Karen Fields: I would say that it's consistent with what the clinical trials are used to make this a high level evidence and I don't think that many of these histologies were addressed, so I think it's appropriate to exclude them.

Benjamin Movsas: Yes. And I agree. So moving on to reliability testing, we have the same issues before that it's not necessarily sufficiently reliable but, you know, I think overall, and the same thing for validity, but having said that, my own viewpoint is, and we've discussed this before, is that when you look at the big picture in terms of where we are and the importance of this and our understanding of the reliability and validity, I think that it's still important and

reasonable to continue to measure this. So -- but I open that up for discussion to other people agree, disagree, any other thoughts on that.

Amber Sterling: Great. Thanks so much, Ben. Beverly, did you have anything that you wanted to add …

Beverly Reigle: No.

Amber Sterling: … to the validity testing or …

Beverly Reigle: Not at all.

Amber Sterling: OK, great. I mean these are similar conversations to the other two measures, so I think …

Benjamin Movsas: Right.

Amber Sterling: … we can go ahead and proceed onto feasibility.

Benjamin Movsas: OK, great. So in terms of feasibility, I agree that the data is retained that we only collect same issues related to the burden of collecting the information but I think that over time, these, you know, does get a little bit easier with more data being abstracted from electronic sources, et cetera. So I think it is overall feasible.

And in terms of usability, that basically overall I feel that this is -- has been publicly reported, it's used in multiple settings, has a large impact. And I think that I would overall say that it has a high usability. And so that's kind of my overall summary. And any other comments or thoughts about that?

Amber Sterling: Great, thank you so much. Beverly, did you have anything you wanted to add?

Beverly Reigle: No, again, I agree with what was said.

Amber Sterling: OK, great. Anybody -- other -- any of the other committee members have anything that they want to add about 0559? If not, I will turn it over to Melissa to discuss 0223.

Melissa Mariñelarena: OK. And the lead discussants on 0223 are Crawford Clay and Brent Braveman. And again, these are very similar measures also developed by the American College of Surgeons, 0223, adjuvant chemotherapy is recommended or administered within four months or 120 days of diagnosis for patients under the age of 80 with AJCC-III lymph node positive colon cancer.

This is a process measure and this was first endorsed in 2007, last came to us in 2012, so it's up for maintenance with us again. And I will hand it over to either Crawford or Brent to give us a summary of the measure.

Brent Braveman: Sure, this is Brent. I'll go ahead and go through if that's OK, Crawford. This is my first time doing this, so folks, please jump in and ask any questions or give me a redirection if I'm skipping over something.

So it is -- this is a maintenance measure facility level. The evidence includes systematic reviews and the evidence was last presented in 2012 including category 1 level evidence related to the NCCN Practice Guidelines. And so the overall level of evidence is moderate. And the developer test, there's been no changes in the evidence since the last time that the measure was evaluated. Anything to add on evidence?

Melissa Mariñelarena: OK.

Crawford Clay: This is Crawford. This may be a good time for me to jump in, excuse me. But I just read recently that the guidelines have been changed from 140 days down to -- I can't remember if it's 30 or 90 days. Does that impact this at all or anybody know anything about that?

Melissa Mariñelarena: This is Melissa from NQF. Does anybody know about a change in these guidelines?

Crawford Clay: OK.

Melissa Mariñelarena: It doesn't sound like it.

Crawford Clay: All right.

Melissa Mariñelarena: So we'll definitely be following up on that on our end.

Crawford Clay: Yes, OK.

Melissa Mariñelarena: Do you have -- do you have the source where you read that?

Crawford Clay: I'll see if I can come up with it and I'll send it to you all.

Melissa Mariñelarena: OK, that is great, thank you.

Crawford Clay: All right.

(Erica McNamara): You know, that would be helpful. And also because -- this is (Erica) from the College of Surgeons. Sometimes we find that the guidelines are written from data surgery and we will get from date of diagnosis. So we aren't aware of any changes …

(Crosstalk)

Crawford Clay: Yes.

(Erica McNamara): I'm just speculating that that might be too different.

Crawford Clay: Yes.

(Erica McNamara): But I don't know.

Melissa Mariñelarena: Yes, probably, if you send …

Crawford Clay: Excuse me.

Melissa Mariñelarena: If you send it to us, we'll …

(Crosstalk)

Melissa Mariñelarena: Yes, we'll share it with the committee and with (Erica) and her colleagues.

Crawford Clay: OK.

Melissa Mariñelarena: Thank you. OK, Brent, do you want to go on and talk about gap and opportunity for improvement?

Brent Braveman: Sure. So the performance rate has improved in 2012 versus 2008. So the last reported performance rate was 86.5. Recent performance data was not provided. And disparities were noted in race and ethnicity as well as in age. And so the preliminary rating for opportunity improvement is moderate.

Melissa Mariñelarena: Great. Thank you. In this -- in this -- when we talked to (Erica) and her colleagues as well, they don't have more recent data because of the same issues with registry data -- collecting the more recent data. So 2012 was the most recent complete data that they had.

Any other questions, comments? (Erica), was that you?

(Erica McNamara): It was, sorry, I was debating on whether or not I should make an additional comment, sorry for that. We can give some information. We do see trends very similar to this when we look at -- as I've noted before, one of our other systems that's the Rapid Quality Reporting System which does have more current data, but it doesn't include all CoC programs so that's why we didn't include it in our application. So if you'd like at the in-person meeting, we can speak to some of the disparities that we still see there.

Melissa Mariñelarena: Great. Thank you. That would be helpful. And the reason we highlighted this is because in our new maintenance process, looking at performance and trends in performance is something that we focus on now in maintenance because we don't focus so much on evidence.

If the evidence hasn't changed, we're not going to revoke it, we don't need to discuss it, but performance over time is something that we're highlighting. So that was why we specifically covered that and we had talked to (Erica) and her colleagues before because looking at data from 2012 seems a little dated but, you know, they gave us the explanation why.

OK, we can move on if there are no other questions or comments. Brent, we can discuss reliability -- sorry, the specifications.

Brent Braveman: OK. So like similar measures we've discussed, this is collected through registry and through paper medical records. And you see that the numerator, denominator and exclusions they are clearly listed. And we felt the specifications were clear and there were no questions there.

The data was -- updated data was provided from 2012 and that was -- the performance rate was provided earlier. And as in other -- the measures we've just discussed recently, there is the same issue in terms of how we need to evaluate the reliability. And that only -- it only assessed percent agreement for two data only. So it's marked as insufficient, but we need to use our judgment in terms of the reliability.

Melissa Mariñelarena: Great, thank you. Does anybody have any comments, questions?

Karen Fields: The only comment I'll make is having used this measure in the past, it's always a challenge to make sure that the users understand, you're talking about date of diagnosis, which start with that very first colonoscopy, not at the time of surgery. So I know that that's an ongoing challenge for making sure that the data is aligned.

So seeing any of your old data about the sensitivity and specificity of this task to help add to reliability wouldn't make me feel happier about it.

(Erica McNamara): That is something that – and we actually have the breakdown by the number of days, what our compliance risk we saw originally, so hopefully that will help.

Melissa Mariñelarena: OK, we can pass that along. Thank you. Any other questions, comments? OK, go ahead, Brent.

Brent Braveman: Let's see. So I think we've also addressed validity testing at the same time given our approach to reliability. Is that right?

Melissa Mariñelarena: Yes, that's correct.

Brent Braveman: OK. And so then in terms of feasibility, similar to other measures we've discussed, data maybe -- be able to be collected through the medical record

and EHRs. But since it relies on our registry, there may be some level of burden, and so feasibility is likely rated as moderate.

Melissa Mariñelarena: Great. Thank you. Crawford, do you have anything you would like to add?

Crawford Clay: No, I think he did a good job.

Melissa Mariñelarena: Great. Thank you.

Crawford Clay: Sure.

Melissa Mariñelarena: OK.

Brent Braveman: And it's being -- the measure is being publicly reported and through several sources, CoC, quality oncology practice initiatives, the PPS-exempt, cancer or hospital quality reporting. And so the preliminary rating for usability and use is high.

Melissa Mariñelarena: Great. Thank you

Brent Braveman: We have seen an increase in improvement of performance and since 2008 from 81.7 to 86.5.

Melissa Mariñelarena: Great. Crawford, is there anything that you would like to add?

Crawford Clay: No, it seems like, in terms of usability, pretty usable thing.

Melissa Mariñelarena: OK. Thank you. Are there any other -- any other questions or comments from the committee? OK. We are doing great on time. If there are no other questions or comments on this measure, we will move on to our last measure, right? OK, our last measure for the day.

This is 0225, this is also from the American College of Surgeons titled, "At least 12 regional lymph nodes are removed and pathologically examined for resected colon cancer." Our lead discussions are Brent Braveman and Jette Hogenmiller, and which one of you would like to take the -- take the lead on this?

Jette Hogenmiller: This is Jette. As the new member, I would love it if Brent would – do I know he did the last measure, I just have a few comments ultimately. Would you mind doing that, Brent?

Brent Braveman: Yes, I'm happy to do what I seem to do OK on the first one, so I'll give it a try.

Jette Hogenmiller: Good job.

Melissa Mariñelarena: Thank you. Go ahead.

Brent Braveman: OK. So this is a facility level measure. Most recent endorsement was in 2012, so it's a maintenance measure. And in terms of preliminary analysis for the evidence, there is -- was a systematic review.

And so the level of evidence was 2A in terms of evidence through the NCCN Practice Guidelines. There is an additional evidence including the systematic review, but that review noted that there's lack of consensus related to the minimal number of lymph nodes necessary to be examined.

And studies looking at registry data have shown that the proportion of patients within a hospital undergo an adequate lymph node examination may not be associated with the survival benefit at the hospital level. The developer tested that there was no changes in the evidence, and so based on the guidance from the evidence algorithm, the evidence is rated as low.

Melissa Mariñelarena: Right. So this is Melissa from NQF. So this, there were some questions at the bottom that we've post to the committee. And so this is asking if you're aware of some higher level evidence. And we don't have to have a full discussion now, but during the meeting, something to consider would be if you feel that there is a need to repeat the discussion and if you would want to revote on evidence.

So some of the possibilities would be the committee is fine with the evidence that is presented this (task) before and you would just -- we would be fine and we would move on. If not, we can have a discussion on the evidence that was

presented before you and then the committee could revote on this evidence and then depending on the results of that, it would either pass or not pass. So those are just some options, we don't necessarily have to discuss them right now, just some things to think about before we go to the in-person meeting.

Jette Hogenmiller: This is Jette. There might be some value in that because looking in this measure and its low level of, you know, the 2B, I went to validity to look for some items. So there may be some opportunity to have a smaller discussion about that to better inform this measure.

Melissa Mariñelarena: OK, yes. And, you know, we …

Karen Fields: Hi, this is -- OK.

Melissa Mariñelarena: The rating is also – because this was a concern back in 2012. So that also contributed to the -- to the low rating.

Karen Fields: Right. This is Karen Fields. I was going to remind you that the last time we evaluated this data, we wanted to see more evidence for the importance of this data. It seemed it's more like since there is such a wide range and there is a lot of controversy around it. It seemed more there is some arbitrariness applied to the 12-node minimum.

And as everyone knows, some of this is related to the quality of the surgeon, a lot of this is related to the quality of the pathology. And so we would -- I think that we need to continue to have our discussion both on the quality of the evidence and I would love it if the developers could cite any more rationale and evidence for us to help us guide us with that.

Crawford Clay: I agree.

Brent Braveman: Yes, there was an interest …

Karen Fields: Thank you.

Brent Braveman: … in the papers here saying that the developer had noted in 2012 that this would -- the evidence would be a moving target. And so it would -- it would

seem that there should be aware as to whether there's been movement since 2012 in terms of the evidence.

Karen Fields: I would think. And, you know, the main thing is this is a process measure and none of the -- most of the evidence that was presented didn't show that this was truly consistently associated with an improvement in survival or outcomes.

So ultimately, the measure the developers need to work on giving us an outcome measure that equals some kind of relationship with the quality of the dissection and the quality of the pathology review equals some sort of outcome -- demonstrable outcome.

And so -- and when I've seen things about ways that centers go about improving their counts, the surgery essentially remains the same and the pathologist comes up with new ways and data to be more specific about counting the number of nodes.

So that obviously improves the pathology reports, but there's a change therapy or is it important. So I think I'm just reflecting our discussion from 2012 so.

Beverly Reigle: This is Bev Reigle, I agree, and I -- this was probably an area that I was really -- of all the ones that we've reviewed. I had questions about this one in particular because the non-experimental data that's supported isn't even that specific, either you've got the wide ranges you mentioned. I believe that was Karen speaking, I'm not sure. So I do think more needs -- more data is needed or at least a discussion about it.

Karen Fields: Or continue to get the developers to get to some more outcome measures rather than process measures.

Beverly Reigle: I agree with you, yes.

(Crosstalk)

Beverly Reigle: … does it fully impact survival, for example, or whatever.

Female: Yes.

(Crosstalk)

(Erica McNamara): (Dr. Sigurdsson), did you have any input. We are doing -- we are working on some data that we might be able to provide by the next – so the meeting in May, but Dr. (Sigurdsson), were you on the line?

(Dr. Sigurdsson): Yes, I am. I think that in all tumor types, the -- a number of nodes that are reported is associated with improved survival, the more nodes that are examined by the pathologists, the better the result of their survival.

And it's most likely related to patients who are node-positive being, found to be node-positive because the pathological assessment of the specimen is better. So the patients who are thought to be node-negative and are found by a more diligent review of the lymph node to be node-positive then go on to get adjuvant chemotherapy which they would not have received otherwise.

So that was in colon cancer and bladder cancer and lung cancer and, you know, across the board, this has been shown to be important. And I agree, it's unlikely to be due to a change in the operation but to a better staging of the patients. And I think that is an important goal.

We and others have shown that by asking the pathologist to spend more time assessing (the) specimen and looking at the lymph nodes, do you get an improvement in staging the patients.

And in colon cancer, in particular, this change is whether or not patients will be recommended to have chemotherapy which is an additional survival benefit for those who are, you know, less diligently thought to be node-negative have become node-positive by virtue of asking the pathologist to spend more time going through the specimen.

There has been a Cochrane analysis to show this. It's quite true that the trace of 12 lymph nodes was quite arbitrary and was chosen as a reflection point in the majority of data both in the (SU) data and then the NCDD data as well as in single institutions.

The first report came from an intergroup colon study. It was designed to look at different chemotherapy regimens where it was awashed. So there were over 3,000 patients in the study and they all have the same survival. And that was where it was found that the better the pathological assessment of the specimen, the better outcome of the patient.

Now, it clearly reflects patients who are thought to be node-negative being moved into node-positive but are very good actors with the small tumor burden, but whether it's the pathologists who make the difference or us, there's a lot of variability in the lymph nodes.

You know, right-sided colon cancers have fewer lymph nodes than the left side of the colon. The elderly patients lose lymph nodes and have fewer numbers than the young patients. And there's obviously a multi factorial component to this. And while 12 is the very arbitrary nature, there's no question that survival is improved if you have a better patient who is staged more appropriately by the pathologist.

So I think it's a very important measure. It's clearly is motivated, the pathologist could look but harder at the specimen. And there's no question, it improved our ability to assign adjuvant therapy to the (appropriate) patient afterwards.

Karen Fields: This is Karen. Thank you for that review. I think it's helpful. I think though that I would still encourage us to have this discussion one more time at the cancer meeting next week. And I would tell you that I think that this is a truly important measure for places that don't -- that can work on strategies to improve the integration of their care.

And because at a community-based hospital or -- where they might not have a cancer center that's got the level of coordination between some of the services, this is a great reflection of improved coordination of care. So I think it's an important measure from that standpoint.

(Dr. Sigurdsson): All right.

Beverly Reigle: I would like to request perhaps one other piece if it's available. Perhaps the number of lymphedema that occurs, sometimes more dissection, more lymph nodes, there is certainly a correlation there with treatment-related effect and often lymphedema.

And I don't know what the percentage is in this area so I was just -- I think it might be helpful to look at it from that, certainly survival is absolutely important and -- but I think knowing something about the possibility of treatment-related comorbidities might be helpful too.

(Dr. Sigurdsson): Yes. We're happy to discuss those at the meeting, absolutely.

Melissa Mariñelarena: Thank you.

(Dr. Sigurdsson): Lymphedema is not a problem with colon cancer but obviously, the data have been shown as well in breast cancer and there's no question that is the operation is changed, the risks to the patient are changed.

Amber Sterling: Great. Thank you. OK. Brent, do we want to -- do we want to move on and discuss gap and care and opportunity for improvement?

Brent Braveman: Sure. So new national trending data was provided from 2013, and the mean performance rate had increased to 89.7 percent from 81.7 percent in 2008. There was a full range in terms of min and max to zero to 100 percent but I think we've covered that. And disparities are noted and race and ethnicity and to some extent and to perhaps a little bit larger extent in (their age). And so preliminary rating for the opportunity for improvement is moderate.

Melissa Mariñelarena: Great, thank you. Again, Jette, would you like to add anything?

Jette Hogenmiller: It certainly did appear that there was movement on many fronts in terms of disparity, you know, like 5 percent so and so. It seems like overall, the movement looked good in terms of different disparities.

Melissa Mariñelarena: Great. Thank you. We can move on to reliability specifications.

Brent Braveman: OK. So the specifications are clearly defined and the measure can consistently implement it, but as in other measures that we've discussed

today, the guidance from the reliability algorithm is insufficient because of only having assessed percent agreement for two data elements.

Melissa Mariñelarena: Great. Thank you. Karen, do you have any feedback on implementation of these measures as far as the data elements and collecting information for this one, or anybody else?

Karen Fields: I don't -- we've never had trouble collecting the data elements. When I've used this measure internally or in some quality improvement projects with outside centers, we've just had to work on going through all the different people involved to make sure that everybody was optimally performing. As far as being able to record the data and the reliability of the data that's recorded, I don't think that there is issues with this measure.

Melissa Mariñelarena: Great. Thank you. All right. Jette, do you have anything to add, comments, questions?

Jette Hogenmiller: No, thank you.

Melissa Mariñelarena: OK. Go ahead, Brent.

Brent Braveman: OK, so and I think we've talked about validity in our discussion of our approach to reliability testing. And in terms of feasibility, that is also abstracted from our record either through paper medical records or EHRs. And so there is, you know, some level of burden for the data collection, and so rate -- preliminary rate and feasibility is moderate.

Melissa Mariñelarena: Great. Thank you. Any comments, questions? If not, we can move on. You're doing great.

Brent Braveman: OK. And the measure is being used for public reporting in quality improvement and is being reported through a number of mechanisms that are listed. And we've seen improvement results from 81.8 percent in 2008 to 89.7 percent in 2013. And so usability and use preliminary rated as moderate.

Melissa Mariñelarena: Great. Thank you. Any other questions, comments?

Karen Fields: No, sorry.

Melissa Mariñelarena: You can, it's OK.

Karen Fields: I was just going to say usability I think it's a usable quality metric. And again, just for the developers, I'd love to sort of hear some future plans for other kinds of strategies to get towards that better staging and or us looking at some of the qualifications of the physicians, the surgeons involved or some of those measures because we've also been able to demonstrate that -- at -- like the ADCC centers that I also represent, we've sort of topped out on those measures with the opportunity to get to a very, very high measures.

And so just moving onto that next level of trying to improve the accuracy and appropriateness of staging is something we would just want to put on the plate for the oncology surgeons.

(Dr. Sigurdsson): Definitely, and this measure talks about how communities -- small communities just do a worse job, to your point, and what can we do to really address some of those shortcomings.

Melissa Mariñelarena: Thank you, Karen. And, you know, we could have this conversation too at the meeting and this is some of the stuff that we can capture in the report because some of the conversations we want to have is, you know, where do we want to -- quality measures and cancer to go forward as we're starting with the new standing committee and looking at our portfolio since it hasn't been looked out in about four years so there's really great conversations to have.

So thank you. We are done, half an hour early. So I will pass it over to Amber to see if there's any housekeeping to take care of.

Amber Sterling: Yes, we just want to remind you that there are two additional workgroups held next week on Tuesday and Thursday at the same time, so from 1:00 to 3:00 P.M. And all of our committee members are welcome to join those calls. You're not limited to just your own workgroup call.

Additionally, as you all know, we do have the in-person meeting here on May 18th and 19th. If any of you have not yet registered or have any kind of issues

with your travel arrangements, please reach out to myself or to the Cancer EM inbox.

Additionally, we just want to make sure that everyone has submitted their measure-specific DOI forms. I think Kaitlynn is chasing a couple of people. I don't know if they're on this call but if you haven't, please go ahead and do so because Kaitlynn will be reaching out to you individually.

And at this time, I'll open it for questions from the committee or the developers before I open it to public comments. Anybody have any question?

(Erica McNamara): We don't have a question but we do want to thank you for this thorough review. It gives us a lot of information that we can go back and provide you with the Friday in-person meeting.

Amber Sterling: Absolutely. Thanks again, (Erica). Thanks for everyone at ACS being on this call. We know it's really helpful for us and for the committee.

So with that, if you could open up the lines for public comment, we will take public comment at this time.

Operator: Thank you. At this time, if you'd like to make a comment, please press star then the number 1 on your telephone keypad. We'll pause for just a moment.

And there are no public comments at this time.

Amber Sterling: OK, well, with that, we will close out this call. Thank you, all, so much for joining us today, and I will give you half an hour of your day back.

Karen Fields: Thank you.

Female: Thank you very much.

Male: Thank you.

Female: Thank you.

Female: Bye-bye.

Female: Bye-bye.

Male: Bye.

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

Amber Sterling: Bye-bye.

END