NATIONAL QUALITY FORUM

+ + + + +

PATIENT SAFETY MEASURES
STEERING COMMITTEE

+ + + + +

THURSDAY
OCTOBER 28, 2010

+ + + + +

The Steering Committee met at the National Quality Forum, Suite 600 South, 601 13th Street, N.W., Washington, D.C., at 9:00 a.m., William A. Conway and Lisa J. Thiemann, Co-Chairs, presiding.

PRESENT:

WILLIAM A. CONWAY, MD, Co-Chair, Henry Ford Health System
LISA J. THIEMANN, CRNA, MNA, Co-Chair, American Association of Nurse Anesthetists
JAN ALLISON, RN, Surgical Care Affiliates
ROBERT BUNTING, JR., MSA, CPHRM, CPHQ, MT, WellPoint
DONALD KENNERLY, MD, PhD, Baylor Health Care System
CLIFTON KNIGHT, MD, Community Hospital of Indiana, Inc.
STEPHEN T. LAWLESS, MD, MBA, Nemours Foundation
ALAN LEVINE, Consumers Advancing Patient Safety
STEPHEN E. MUETHING, MD, Cincinnati Children's Hospital Medical Center
JANET NAGAMINE, MD, RN, Society of Hospital Medicine
PAUL NAGY, PhD, University of Maryland School of Medicine

PRESENT (continued):

DAVID P. NAU, PhD, RPh, CPHQ, Pharmacy Quality Alliance
PAUL R. SIERZENSKI, MD, Christiana Care Health System
DANIEL SOLOMON, MD, Brigham and Women's Hospital*
IONA THRAEN, MSW, Utah Department of Health
DAVID E. TURNER, MD, PhD, MPH, Monsanto

NQF STAFF:

PETER ANGOOD, MD
HEIDI BOSSLEY, MSN, MBA
ANDREW LYZENGA
ELISA MUNTHALI
LINDSEY TIGHE
JESSICA WEBER

ALSO PRESENT:

KAY SCHWEBKE, MD, MPH, Ingenix*

*Participating via telephone

C-O-N-T-E-N-T-S

P-R-O-C-E-E-D-I-N-G-S

9:04 a.m.

CO-CHAIR THIEMANN: Welcome everybody, and thank you for coming. And thank you for your work thus far associated with the Patient Safety Measures Steering Committee work.

We just wanted to go around, do brief introductions of ourselves; one or two statements about our past in the sense of what represents what specialty, what's your expertise and so forth.

And so I guess I will start. I am Lisa Thiemann, I'm Senior Director, Professional Practice with the American Association of Nurse Anesthetists. Been a CRNA, a certified registered nurse anesthetist for almost 15 years with a specialty in pediatrics and some past program administration for nurse anesthesia programs.

CO-CHAIR CONWAY: And welcome also. And thank you for spending all last weekend

scoring this large amount of measures. We appreciate everybody's effort there.

I'm Bill Conway. I'm the Senior Vice President and Chief Quality Officer for the Henry Ford Health System. My clinical background is pulmonary and critical care medicine. I have five daughters and know a lot about shoes.

MS. MUNTHALI: Elisa Munthali, NQF staff.

MS. WEBER: Jessica Weber, NQF staff.

MS. TIGHE: Lindsey Tighe, NQF staff.

MS. BOSSLEY: Heidi Bossley, Managing Director for Consensus Development Process here at NQF.

DR. NAGAMINE: Good morning. Janet Nagamine. I'm a hospitalist at Kaiser Santa Clara in California, Patient Safety Officer -- or used to be, actually, and former Quality Chief. I'm here representing Society of

Hospital Medicine.

MR. BUNTING: Good morning. I'm Bob Bunting. I work for WellPoint. My medical training is as a medical technologist laboratory science. I've got more years than I'd like to count for quality and patient safety. My current role is a clinical research manager.

DR. LAWLESS: I'm Steve Lawless. I'm with Nemours Foundation. I'm a pediatric intensivist, but for the last four years I've been the Vice President of Quality and Safety for Nemours. We're a large pediatric multi-specialty group in Delaware and Florida. My venue is patient safety, environmental safety, infection control, risk management, peer review and other things as assigned.

DR. KENNERLY: I'm Don Kennerly, and I'm an internist by training, and I've been with the Baylor Health Care System in Dallas for the last ten years. I've served in various quality and safety roles and currently

serve as the corporate Vice President for Patient Safety across the network of hospitals.

MS. THRAEN: Hi. My name is Iona Thraen. And I'm the Patient Safety Director for the Utah Department of Health, so I'm the lowly MSW Public Health, we're from the government, we're here to help you. And I brought the paper redundancy when all the systems go down, so I just want you to know that. Thank you.

MR. LEVINE: My name is Alan Levine. I'm the Consumer Advocate. I volunteer for Consumers Advancing Patient Safety, a stakeholder group on the Consumer Council.

Formerly I was an employee of the federal government. I retired in 2008 from the Department of Health and Human Services where I worked for the Inspector General's Office and did -- coordinated a $3 million study on Medicare adverse events.

MS. ALLISON: Hi. I'm Jan Allison. I've been an RN for 30 years. And I'm a Certified Health Care Safety Professional. I've been in the ambulatory surgery center industry for 25 years, and I work for Surgical Care Affiliates. We're a company that owns approximately 130 surgery centers across the country and growing.

MR. LYZENGA: I'm Andrew Lyzenga, NQF staff.

DR. SIERZENSKI: I'm Paul Sierzenski. I'm an emergency physician at Christiana Care Health System. Still practice clinically. I direct emergency trauma and critical care ultrasound for that institution. And I sit on our college's Quality Performance Committee.

DR. NAGY: I'm Paul Nagy. I'm trained as a diagnostic medical physicist. And I serve as a quality informatician at the University of Maryland. I direct the quality efforts and have been doing Six Sigma Lean for

about the past 15 years.

DR. MUETHING: Good morning, all. I'm Steve Muething. I'm a pediatrician. I'm a Safety Officer at Cincinnati Children's. And I'm representing NACHRI, the National Association of Children's Hospitals and CHCA, which is the Child Health Corporation of America.

DR. NAU: Good morning. I'm David Nau, Senior Director with the Pharmacy Quality Alliance. I'm a pharmacist with a Ph.D. in Health Services Research. I have split my time between academia and running a research team for a health plan. Been with PQA for the past year.

DR. KNIGHT: Hi. My name's Cliff Knight. I'm a family physician from Indianapolis, Indiana. I'm representing the American Academy of Family Physicians. And I'm Chief Medical Officer of Community Health Network, a five hospital system in Indianapolis.

DR. TURNER: Yes. Good morning. David Turner. I'm an occupational medicine physician. I'm been in this capacity probably for about 20 plus years; 10 years I've been in corporate roles. I'm currently with Monsanto Company. My responsibilities there really are twofold. I work towards a global health policy in terms of preventive medicine, and I'm also working very closely with our benefits team in terms of developing a package that really supports preventive health issues.

DR. ANGOOD: Good morning. My name is Peter Angood. I'm the senior advisor for patient safety at NQF. A surgeon and critical care guy from background and spent several years at the Joint Commission as well. And I've been working with Heidi to oversee this project overall. And we appreciate your attendance and your efforts on our behalfs. Thank you.

CO-CHAIR THIEMANN: And I think we're going to turn it over to Elisa Munthali

and Heidi Bossley for an overview.

MS. MUNTHALI: Thank you very much. Before we go through the slides, there are a couple of housekeeping items that we wanted to bring to your attention, especially for those of you that are participating here at NQF. We wanted to tell you first about the bathrooms, which are very important. I think they're around the corner and by the front door. And also for those of you who are here, just please help yourself to the food in the adjacent room.

We wanted to remind you that this is an open meeting. It's open to NQF members and to the public. And they'll have opportunity to provide comment at specific points during the agenda.

We've also invited measure developers who will participate either here in person or via teleconference or webcast. And they'll be here to introduce their measures and to provide any clarity to questions that

you may have.

You may have noticed that we have a court reporter who is transcribing this meeting and is taking notes. And the audio recording will be posted to the website as well as the transcription. And NQF staff will be putting together a meeting summary to accompany those materials.

And so we will go ahead with the presentation. And maybe at this time, operator, do you know if Dr. Solomon has joined the call? She hasn't answered. She said she'll notify us once he does.

We do have one Committee member who will be participating via teleconference today and tomorrow, Dr. Dan Solomon.

NQF provided some of this information to the Steering Committee during an orientation call that we had. And you've received much of this information before, but we thought it was important to reiterate.

NQF is a private, nonprofit

voluntary consensus standard setting organization with a membership of over 400 organizations. Our members are organized into eight very distinct stakeholder councils that include consumer groups, health plans, health professionals, purchasers, public and community agencies, quality improvement organizations and suppliers.

Our Board members mirror the diversity of stakeholders that are interested in our mission with a deliberate but slight over representation of consumers and purchasers.

Our Board established three standing committees to help guide their work. Those include the Consensus Standards Approval Committee, which we refer to as the CSAC, and they consider all candidate standards or practices and make recommendations like you will do today for NQF endorsement to the Board.

The National Priorities Partnership

is a 32 member organization collaborative that assesses high impact priorities and goals and takes collective action to address those goals. And the Leadership Network provides guidance on NQF's education, research and recognition programs.

I'd like to talk a little bit about developing consensus. We apply a very specific process that we call the consensus development process, also known as the CDP, to gain consensus about which measures or practices should be National Voluntary Consensus Standards. As I previously mentioned, our membership is open, and it is a diverse representation from a full spectrum of health care stakeholders, including private and public organizations.

And now we'd like to show you a visual schematic of the consensus development process, which we call the CDP. And this schematic shows the important steps of the entire process including our current step,

which is the Steering Committee's review. Following this, the Committee will make recommendations, and those recommendations will be turned into a draft report that the NQF staff will put together. The report will be available on our website for our public and member comment for a 30 day period.

Following that, the Steering Committee will address any comments that are brought forth during that period, and then the report will be posted on the NQF website for member voting for 30 days.

After that, the report moves on to the CSAC, and they'll consider your recommendations, and they'll make recommendations to the Board. The Board would then ratify those recommendations and then a 30 day appeals process follows.

And we'd like to talk a little bit now specifically about the patient safety measures project. It's funded by the Department of Health and Human Services. And

there are two main goals.

The first one is to identify and evaluate and endorse additional measures that are suitable for public reporting and quality improvement that specifically address health care associated infections, medication safety, and other safety measures. And then the second goal is to identify gaps in existing patient safety measures and to recommend potential measures to fill those gaps.

We wanted to give you an overview of the project and how we got to where we are today. There are three technical advisory panels, or what we call TAPs, that were formed to address medication safety, health care associated infections, perinatal care. And they were formed to assist the Committee with their work. And they all met between August and September of 2010.

The Health Care Associated measures, which I will refer to as the HAI measures, are on a separate expedited CDP

timeline. The draft report is now available on the public website for 30 days. And the deadline ends on November 8th.

For the perinatal measures, the measures that we received, we received them from one steward. After the initial review by the TAP, the steward withdrew those measures to concentrate more on measure development. And they hope to resubmit those measures for an NQF endorsement maintenance project on perinatal care that starts in spring 2011.

So today what the Committee will do is review 13 medication safety measures and six additional safety measures. I would like to note that two of the HIV medication measures are pending the TAP's evaluation, and so the Committee will review those separately, probably within the next few weeks. I think you've received emails from Lindsey and Jessica with those meeting dates.

And so now we just wanted to bring your attention to the timeline for both of

these phases of the patient safety measures project. The first is related to the HAI measures and the second the medication safety and additional measures.

We wanted to just kind of alert you to many of the meetings that are coming up. And we do apologize. We have several meetings that we're planning in the next few weeks. And we're trying to get a lot of work done before the holiday season.

They include a follow-up meeting from today's meeting, evaluation of the HIV measures that I mentioned before, and review of the comments that will come from the HAI draft report.

We wanted to talk a little bit about your role collectively as a Steering Committee. You will act as a proxy to the NQF multi-stakeholder membership for the Patient Safety Measures Project. You will continue to work with the NQF staff to achieve the goals that I mentioned earlier. And most

importantly, you're here to evaluate the candidate standards and evaluate them against our formal measure evaluation criteria.

In addition to that, you're making recommendations to our membership for endorsement. And you'll respond to comments that are submitted during the comment period. And your Co-Chairs, Dr. Conway and Ms. Thiemann, will represent you as a Steering Committee at the CSAC meeting. And you'll also respond to directives that the CSAC puts forth.

You also have roles as individual members. And we've assigned all of you primary and/or secondary discussion lead responsibilities for measures based on your experience and expertise. And prior to this meeting you did conduct a pre-meeting evaluation online. And really that's not your final recommendation, but it's to help us to facilitate today's discussion.

And now I'll turn it over to Heidi

Bossley, who will go over NQF endorsement criteria.

MS. BOSSLEY: I think this is where Elisa and I are going to be fighting for the microphone because it's attached.

You all, I know, have been looking at the criteria, but I think it's helpful just before a full two days of looking through many measures to go through the criteria again. So I'm not going to go very quickly, but -- or I am going to go quickly. But if you have a question, stop me.

So we have new criteria that were approved by the Board in August of 2008. We continually take a look at that criteria and, in fact, in January there will be a new set that's a little bit more robust and explicit on the importance and the scientific acceptability components. But, again, that's not yet effective until January. So you are operating under the August 2008 criteria.

But what we did back then was take

a look at how could we strengthen and clarify our criteria looking for that stronger link to the national priorities and also higher level performance measures; getting more proximal to the outcome if you're looking at process measures. And again, looking truly our ultimate goal to outcome measures.

Want to see that we have some greater measure harmonization. It's not helpful to have two measures out there that are almost the same, but not quite. So, again, trying to push toward that.

And then, as I said, the last two. So if you could go to the next slide. Okay.

Conditions for consideration. This is something that we as staff do. We make sure we have agreements. The forms are complete. You're not looking at completely blank forms; that type of information.

Then the four criteria are importance, scientific acceptability, usability, and feasibility.

So just again briefly. Importance when you looking at this, this is your must pass criteria. So today I think and tomorrow you're going to have probably quite a bit of discussion on do these measures really meet importance. And there's three components to it. They don't have to meet all three, but the closer they can the better it will be.

Does it have a high impact area? And I think a lot of the ones that you'll look at may not be your typical high impact in the way of very broad across the population, but within a specific specialty or a condition, it may actually be a high impact area. So that is one other way to look at it as opposed to, say, looking at a typical diabetes measure.

Is there a gap in performance? We're also looking for variation. So look to see if that information is provided. And then absolutely critical, is there evidence to support that measure focus?

You want to try to capture the

measures and put forward the measures that, again, are more proximal, closer to the outcome. So I think that's, again, where the evidence component comes in and you all will probably want to spend some time talking about it.

Now to the next slide.

Scientific acceptability. This where I think the fun begins. You get to look at the specifications. Is it precisely specified? Is this something that really multiple groups could take what they're provided at the ultimate end of the day and try to implement? Can they take that information and somewhat uniformly implement it?

Is there some form of testing reliability and validity? This is where in the new criteria, that updated criteria that comes out, you'll see a little bit more explicit because it's not quite as crisp as I think we'd like it to be. That's why it's

being updated. But here I think you have some measures that have been tested, some have not. So I think it's something that you will spend some time, and staff will be happy to guide you on that as well.

Exclusions if there are, is there justification and are they reasonable? Risk adjustment. I don't think you're really looking at any today. You already did those measures. Those are out for comment.

And then looking to see do you have information on identifying differences in performance? If they use multiple data sources, have they tested and compared those two and shown that you can compare the results or not? And then again, disparities being a key element that we look at across all of our measures, have they started to look at stratifying? That's not a requirement, but it's always helpful to see if they've looked at it.

The third criteria is usability.

So we're looking for how much has this been used to date. Some measures when they really come in the first time have not been used very much, and that's okay. But, again, we're looking to see does it at least demonstrate that it's useful for public reporting and quality improvement? It is harmonized? And then if there are existing measures, is there something that just makes this measure rise to the top that you really think it's worth putting forward?

Feasibility. This is one that I think we're going to continue to look at to work on the criteria, make it a little bit more crisper. But we're looking again at can that data be collected somewhat easily? Is it generated during care processes as much as possible? As there electronic sources? We're really moving more toward and trying to see where we can get measures to electronic health records, electronic clinical data.

And then has the developer been

able to look at unintended consequences, anything with inaccuracies, et cetera?

I don't know that you have any competing measures today, but you did already look at some. But if you do, we'll walk you through the process. It's something, again, an area where we're finding more and more as we're going into more of a maintenance cycle looking at existing endorsed measures, plus new. We're finding we're getting quite a few competing measures coming through, and we're working on additional guidance. But if it comes up today, we'll work with you on that.

Time limited. So this is where it's fun. We have a new policy. Time limited was created given the environment and the emphasis on public reporting and pay for performance programs that were out there, and this need for a larger amount of measures. And so the NQF Board really did a look at that a few years ago and say, yes, we need to find a way for those measures where we feel are

sound in every other way but the testing aspect, can we put those through and require that they be tested within a certain amount of time.

We find that we needed to take a look at that again. And so we have new criteria. You are not necessarily held to this criteria because it did occur at and about the same time measures were submitted. So measure developers were not notified of the new change until after they submitted it. But, again, I think it's worth you all being aware of what the new criteria is, which is there is no other currently endorsed measure on that topic. Again, if that one that's endorsed is tested, and this other one is not -- I don't know that you would be able to say it's a superior measure. So that is our thinking with that piece of the criteria.

Is there a critical timeline that must be met? Is there a legislative mandate? Again, that's not something that would

necessarily apply for you all today.

And then is the measure not complex? So if a measure is a composite, looks at outcomes or requires risk adjustment, we're not sure that that's ready to go for prime time out there for everyone without having some type of testing. So that's the thinking behind that.

What we are working with stewards on, any measures that come out of any project from now on, we're trying to get that testing within 12 months. It used to be 24 months, and we're finding that's too long to have a measure out there and not have it tested. So, again, we're working with everyone to see if we can get it within 12 month time frame.

I think I'll stop there, and let's see if anyone has questions on the criteria.

MR. LEVINE: Yes. There was a slide that mentioned consistent with national priorities.

MS. BOSSLEY: Yes.

MR. LEVINE: Is that NQF priorities or some other priorities?

MS. BOSSLEY: Good question.

So when it was first created it was for the National Priorities Partnership priority areas, which are over use, safety being one of them looking at -- I should have them memorized, but I don't. But there are six of them.

DR. ANGOOD: I got them.

MS. BOSSLEY: I knew Peter would have them. Good.

DR. ANGOOD: I've been there all the time.

So it's patients and families, population health, safety, continuum of care, appropriate end of life care or palliative care, and then efficiency. Those are the six areas. And now we've added two more in the recent Board meeting, and that is access to care as well as appropriate infrastructural support in order to make all of those things

happen.

MS. BOSSLEY: Right.

MS. THRAEN: So just for clarification because I struggled a little bit with some of these measures that were more quality focused, I thought, than safety focused. So even though this is called Patient Safety Steering Committee, are we operating more broadly as you just described?

DR. ANGOOD: Some of the measures have been -- it's been a struggle for us to find whether they should fit in safety, per se, or whether they're appropriate in other areas. I think the better answer is look at it specifically along the criteria that Heidi just reviewed. If you think they are not meeting muster, then move them on. If you think they're reasonable but they don't quite fit safety, let's talk about it at that moment.

We have a whole variety of other groupings of measures projects, and we do

shuffle the deck from time-to-time. But, you know, think about it as an end user looking in, you're really just looking to see what's NQF endorsed measure, you don't necessarily care what bucket it's in. But we've got these measures in this grouping for us. It was sort of the best proximate area for some of these.

DR. LAWLESS: A question for you just clarifying on the testing.

MS. BOSSLEY: Yes.

DR. LAWLESS: During the discussions, because that opens up a nice Pandora's box. And so the clarification do we use testing in our evaluation, or do we just kind of sublimely know it's there?

MS. BOSSLEY: So whatever testing you find there, you should evaluate. If there is none, then we'll look at the time limited potential. Yes. But if you do see testing, we fully want you to evaluate whether you think it's adequate testing, has the conclusions really drawn to the point where

you think it's a well, precisely specified measure that could be used. Yes.

And going back to the priorities, just one thing I wanted to let everyone know. The Secretary because of the new ACA law is looking at national priorities. So we anticipate that there will be a new set, hopefully and in line with what we have now with the NPP priorities, but those will become the new priorities as a part of this criteria. So we're actively looking to see what comes out of the work, whomever does that work, with the priorities.

MS. MUNTHALI: Thank you.

Before we go into the evaluation, we wanted to let you know we're having some technical problems with the phone. We're on? Okay. Great. So we can go on with that.

But before doing that, we wanted to let you know that on those thumb drives that all of the Steering Committee has, the briefing materials that you received last

week. So all of the materials, including the measure submission forms, evaluation forms, the agenda, the measure assignments for reviewers, that's on here. So you can upload those on the computer.

And we've included this information that was also included in the briefing materials. These are just some talking points. Essentially just make sure so that those that are participating by teleconference and by the webcast that you identify the measure that you're presenting by the ID and the title as exampled on the screen. And make sure that you cover all of the evaluation criteria as Heidi alluded to earlier.

And we can advance onto talking about voting. And we are very excited to be using for the first time hand-held voting devices. This is the first project that is using them. So we ask that you bear with us if there are any technical problems that may arise as being the first to use it.

We think it's pretty easy to use. I don't have one to show you right now. They were in front of me, but I don't know where they are. But they're very small, very light. There are only nine -- actually there should be zero to nine on the keypad numbers on there. And the ones that you should be concerned about are numbers 1 through 4, and I'll tell you why in the next slide. Those correspond to endorsement recommendations that you may have.

So when entering your response, you make sure that you select the number, then you hit send. And in this example if you wanted to recommend a measure for endorsement, you would hit one, yes, I recommend this measure as written, and then you hit send.

Likewise, if you don't want to recommend the measure you would hit no, which is represented by the three on the key card, and you would hit send.

You can modify your response. You

hit the caution symbol on the device, then you select the corresponding number that you want to select. And then you hit send.

But we must warn you that you can't change your response once you hit send. So, you know, take your time. Make sure you use it correctly. So if you feel like you've made a mistake, you can correct it as long as you don't hit send.

So if you select one and you wanted to select two that you recommend with modifications pending the developer's modifications to the measure, then you can hit two -- you can hit the caution sign, hit two and then hit send.

And as soon as you hit send, and as soon as everybody hits send, the results will be displayed on screen, and they'll also be announced by the co-chairs so that those that are participating by the webcast and also by teleconference would know the result. And I think that's it in terms of our slide

presentation.

Dr. Solomon? Dr. Solomon or Dr. Diamond?

Operator, are you there?

DR. SOLOMON: This is Dan Solomon.

Actually, I was trying to speak during your presentation, and I guess it was on mute or something. But I'm not quite sure how we're supposed to be voting from -- via teleconference. And I actually don't have the WebEx information. So I don't have that in front of me either. I just have the printed briefing material.

MS. MUNTHALI: Okay. That's a good question.

What we're going to do is just ask you for your vote. That's the only way that we can do it. And Heidi Bossley will send you the materials. So you will be receiving an email from her very shortly.

DR. SOLOMON: Okay.

MS. MUNTHALI: Operator?

OPERATOR: Yes, I'm here.

MS. MUNTHALI: Would you mind leaving the lines open at all times?

OPERATOR: Yes, all lines are open now.

MS. MUNTHALI: Thank you so much.

Okay. So I will turn it over to your Co-Chairs.

CO-CHAIR THIEMANN: At this time we'd like to move into evaluation of the individual performance measures. And at this time we'd like to ask for general overview comments by the measure developer for the four that are slated for consideration at this time, which would be PSM-017, 018, 019 and 020.

So if we have the measure developer Ingenix.

DR. SCHWEBKE: Yes. Hi. This is Kay Schwebke.

CO-CHAIR THIEMANN: Great. Terrific. Thank you, Ms. Schwebke.

If you'd like to go ahead and provide an overview regarding these four performance measures from the measure developer's perspective, we'd appreciate that.

DR. SCHWEBKE: Yes, happy to do so. So we have four measures here. And I apologize. In previous meetings we haven't been asked to give overviews, but I will do the best to provide that for you.

The first measure is measure PSM-017-10. This identifies patients with rheumatoid arthritis who are taking one of three specific medications, methotrexate, sulfasalazine or leflunomide that had serum ALT or AST monitoring in the last three report months.

So the way the measure is built is we identify people using the condition confirmation that are specified and the denominator details, identifying individuals that are two years or older who have continuously enrolled in medical benefits and

pharmacy benefits or have been identified through a disease registry. And we look to see whether or not they're taking one of these medications.

The patient has to be "actively" taking one of these medications, and that's defined as the following. There's a filled prescription for one of the medications within the last 90 days -- sorry, within the last 120 days. And in addition as we look back over the 12 month report period the prescription has had a number of days filled that's greater than 90 days. And the purpose of that is to make sure that not only has the patient recently filled the medication, but they've also been taking it for a prolonged period of time.

If those are true, then people are placed in the denominator. And then really the intervention of interest is to identify whether or not that serum ALT or AST test was obtained within those last three report

months. Because we have many people, many customers who have claims that go beyond the last three months of the report period, if we have data that extends 90 days after the report period, then we also accept that laboratory test to achieve numerator compliance.

The compliance for this measure in our testing database of over 15 million members was 66 percent. And so we believe that there is a reasonable gap in care here that can be addressed.

The remainder -- there's a few bits of evidence that have supported this measure. One from the pharmaceutical manufacturer, a second from the American College of Rheumatology. And in the ACR 2008 Recommendations they're actually quite clear. Their specific guideline recommendation is that for individuals who have been on chronic therapy with one of these medications that specific medications should be obtained, and

that includes a complete blood cell count, chemistry panel, determination of creatinine levels. And I state that now only because as we walk through some of the upcoming measures it's really kind of based on the same logic and the same literature support.

Now just let me know if you want me to stop here or if you want me to continue to move through the other three measures.

CO-CHAIR THIEMANN: No. I think that's a nice overview. And thank you very much. I think as we go through each individual measure having you available to answer questions from the specific Steering Committee members would probably be the best way to proceed at this time. So thank you very much.

DR. SCHWEBKE: Okay. Yes, you're very welcome.

CO-CHAIR THIEMANN: At this time, we'd like to -- the primary discussion leader for the first performance measure up for

consideration, PSM-017-10, Dr. Kennerly and secondary discussion leader Dr. Solomon.

Dr. Kennerly, would you like to provide the introduction for this performance measure, please?

DR. SOLOMON: Was that -- I'm sorry, were you asking Dr. Solomon or -

CO-CHAIR THIEMANN: I have primary discussion leader Dr. Kennerly.

DR. KENNERLY: No, I don't think I was assigned this.

CO-CHAIR THIEMANN: Dr. Solomon, would you care to then -- were you listed as primary discussion leader then?

DR. SOLOMON: I'm happy to discuss it. I actually can hear you well, but I couldn't anything that Dr. Kennerly was saying.

CO-CHAIR THIEMANN: Okay.

DR. SOLOMON: So I don't know if the microphones can be replaced.

CO-CHAIR THIEMANN: Thank you.

DR. SOLOMON: As a rheumatologist, these are very familiar measures. And as the past chair of the Quality of Care Committee at the American College of Rheumatology we've spent long periods discussing these measures as part of the recommendations that Dr. Schwebke discussed as far as the ACR's 2008 Recommendations regarding monitoring.

And as she noted, these sorts of recommendations are part of the manufacturer's discussion as well. They've been recommended by the ACR based on really an expert process without a lot of evidence. The total of the evidence is really a variety of case series that looked at people who had toxicities related to these medicines and attempted to develop some sort of monitoring guidelines which might stave off those toxicities. But there's really never been any formal epidemiologic studies or trials that would support these specific measures.

Having said that again, there's

broad agreement amongst experts that these are reasonable measures. The exact frequencies of the monitoring is debated even amongst the rheumatology community. There's people who want these to be done less frequently, and people note in large cohorts that are recently published that people who get these tests done less frequently seem to have similarly low rates of toxicities. Again, there's really been no very formal comparison though of different monitoring frequencies. And so we're in a bit of a data vacuum.

There is broad agreement that there should be monitoring, but the precise monitoring frequency I think is where people still debate the issue.

What else can I say?

CO-CHAIR THIEMANN: Dr. Solomon, if you would care to also expound on scientific acceptability, feasibility, usability as well associated from your perspective of this measure, we would appreciate that.

DR. SOLOMON: Sure. Again, the scientific acceptability I just discussed. I mean, there's weak evidence, but there's broad agreement that monitoring should be done. As far as the exact monitoring frequency, there's really not broad agreement whether it should be done every eight weeks, every 12 weeks, or every, you know, six months.

The feasibility, I mean these sorts of lab tests are generally easy to identify using administrative claims data. And they're difficult to find in electronic medical records, obviously, because people often get labs done outside of a health system. And so I think that if the administrative claims data are used, people believe that they are complete capture of the data. I don't know if that's ever been carefully tested, but I think there have been some tests of that just to suggest that it is a valid and reliable source of information.

So, you know, feasibility and

usability, I think, are commented on by the administrative claims access to these data. And I've commented about the scientific acceptability already.

What else can I tell you?

CO-CHAIR THIEMANN: No, I think that's terrific. Thank you for the overview.

And, Dr. Kennerly, anything additional to add to Dr. Solomon's comments?

DR. KENNERLY: No. I don't think so. Aside from, again, I apologize if I was supposed to be doing something here, but I didn't have that on my to do list. But I agree with the discussion that he's presented.

CO-CHAIR THIEMANN: All right. Terrific.

I'd like to open it up to the rest of the Steering Committee Members at this point. If we proceed through, importance to measure I think would be the first critical threshold for the measure to consider. And so I'd like to open it up to the rest of the

Steering Committee members concerning opinions.

MS. THRAEN: Can I get a clarification first? On the reference, and this is probably from the developer, in the textual information they reference the discussions with American Gastroenterological Association. And I didn't quite understand what they were saying there, whether or not they were -- it says, whereas the measure did not describe any similar combined work with the ACR, the measure developer stated there was pre-existing relationship with the AGA leading to a greater effort to work together. But that doesn't tell me concretely where AGA is related to this particular measure, and could the developer comment on that?

DR. SCHWEBKE: Can you just clarify? I'm sorry, it's a little bit difficult to hear some of the members on the phone. Are you referring specifically to a comment in the measure application or the

document that I sent back to answer some specific questions that the previous Committee had?

MS. THRAEN: Actually, I'm commenting on the review notes that are in our documents. Let me repeat that. It said that --

CO-CHAIR THIEMANN: Before you proceed, would you mind specifying exactly where in the document you're looking at so all Steering Committee members as well as the measure developer might be able to focus?

MS. THRAEN: Fine. Summary Table of TAP Ratings of Subcriterion Comments, page 8. It says, the TAP noted that the measure, and describes the measure, referred to discussions with the AGA, whereas this measure did not describe any similar combined work with ACR, rheumatology, the measure developer stated there was a pre-existing relationship with the AGA leading to greater effort to work together between the organizations. However,

the measure developer also noted that this measure specifications are consistent with ACR guidelines.

So I guess I'm confused on whether or not this is a measure that's applicable to both AGA and ACR or it was determined that it wasn't. I just didn't understand the language.

DR. SCHWEBKE: The measure specifically is designed for people with rheumatoid arthritis. And actually after we had developed this measure we actually had approached ACR with the interest in asking them to review the measure to make sure that they were comfortable with the measure logic time frames, et cetera. And at that point their recommendation was for us to really focus on their ACR 2008 Recommendations along with some earlier recommendations that they had published I believe in 2006.

Now with that being said, we also appreciated along with other measures that

we've developed not for rheumatoid arthritis but for another condition, inflammatory bowel disease, that we were seeing some similar medications that are being used for IBD, inflammatory bowel disease. And we wanted, whenever possible, to be consistent when it made sense to have monitoring recommendations for drugs used to treat RA to be consistent whenever possible for drugs that were being used to treat IBD.

So that reference to AGA is more in the spirit of our attempts at harmonization and actually very strong collaboration that we have had with AGA where AGA has actually reviewed all of our GI measures and have given us feedback that we've brought back to our external consultant panel to try to achieve harmonization with medications used by most specialists.

So, I appreciate the confusion. And, hopefully, that provides some clarification.

MS. THRAEN: So does this measure in its current form achieve that?

DR. SCHWEBKE: It this form it does. I think as we talk about some of the upcoming measures there are definitely some differences. And some of it, I think, gets to the earlier comment that the literature is not always clear around the timing, the frequency at which some of these tests should be done. And sometimes because there's no clear evidence-based study that's defining that for us, we are turning to national experts to help with that definition. And I will say that sometimes there is disagreement between our rheumatology specialists and our GI specialists. And when we've seen that discrepancy, we've tried to err on the side of being a little bit more conservative and in allowing for a longer time frame.

But I think that actually is not such an issue here. It may come up with some of the other measures that we're going to

discuss today.

MS. THRAEN: Thank you.

DR. LAWLESS: This is Steve.

CO-CHAIR THIEMANN: One thing I'd like to suggest which I think we didn't say earlier. If you'd like to make comments, one of the things I think some of us have found in past is to turn your ID card, table card up so that we know who wants to speak. Great. Terrific. Thank you very much. Go ahead.

DR. LAWLESS: Yes. This is Steve Lawless.

I'm curious about who is reporting. I see the data sources from a multitude of data sources, but I'm not sure as a safety measure who is reporting and then what are they reporting back to. So it's a good process, but there's no --

DR. SCHWEBKE: Yes. Good question.

So basically this data is coming from multiple payers. So this is a national database of over 15 million members.

Important to keep in mind that most of the members in this database are -- patients in this database are commercially insured. It is a very geographically diverse database. It's not coming from a single payer, it's actually coming from multiple payers.

It's derived from customers we have where we have shared tools with payers, payers who have purchased certain products. And sometimes as part of that contractual agreement in a de-identified way they have then contributed their data to this large database that we can use for a variety of benchmark purposes.

And so basically these are administrative claims, including LOINC codes which actually had been a very rich source of making sure that particularly if a diagnostic test is done, that we're not only looking for a CPT code, but we're also potentially including a LOINC code as another data source option. That information is coming in through

payers. Providers don't need to be submitting anything, this is coming in through paid claims is another way of saying it.

CO-CHAIR THIEMANN: Dr. Nau?

DR. NAU: Well, with regards to evaluating importance of any of these measures that we're considering, I think the challenge is that it's context specific or really relevant to what you're trying to accomplish. And so I think if we're taking the perspective of evaluating importance based on the need to create a national public report on the most important health care quality issues, I might say this was fairly low importance relative to some other issues. But if we're looking at this from the perspective of some focus quality improvement efforts for patients with rheumatoid arthritis, I'd say it's important to include this within that measure set.

So, I think that's the challenge of not really knowing the perspective to take

when evaluating the importance of some of these. I tend to take the narrower perspective there of saying that if we're trying to do some quality improvement around rheumatoid arthritis and safe medication use in that population, then I think this would be an important measure to include.

So, I think that's where different people around the table might be taking different perspectives. And so I think that's where would it be useful to have a brief discussion just about what perspective we should be taking or if we should just have our own perspectives on that?

CO-CHAIR THIEMANN: I think that's a terrific point. At this point why don't we hold that for a second and we'll engage in that conversation I think in a moment. I'd like to hear what the other three individuals who have lifted their cards.

Dr. Turner, I think you were next, and then we'll come back to Dr. Nau's

question.

DR. TURNER: Which I agree, I think that's an excellent question that we need to have a frank discussion around.

My question just relates to some of the commentary that was provided by the Technical Assistance Committee when they were speaking relative to the three separate drugs that the sponsor has offered with this measure.

And I guess I would like just a little bit more commentary in terms of the relevance to this type of testing and the frequency if it should be considered to be comparable among the three drugs, or if there might be some specific differences that could suggest that the measure is more appropriate for one or the other?

DR. SOLOMON: This is Dan Solomon.

I'd like to give you some feedback on that. And Chris, you probably have some information from the GI perspective.

You know, most of this information comes from methotrexate because it's been in use for rheumatoid arthritis for the longest period, the leflunomide being about a ten year old drug, and sulfasalazine not as widely used and not as widely studied.

And so the methotrexate data is the richest in the cases of toxicity and the formulation of the monitoring frequency is really based on methotrexate data.

Primarily there are some data around sulfasalazine, leflunomide is much more sparse and I think that people in the rheumatology community have attempted to simplify this for practitioners by making the monitoring similar across drugs.

DR. TURNER: Yes, and if I could just, with your permission, ask a follow-up question. Not in the same context, but just based upon the commentary provided by Ingenix now and in earlier response to the question, I'm wondering about your database. It sounds

to me that it's quite a robust database. And I'm wondering if one is looking at national reporting of this measure across multiple payers and multiple commercial plans if simple administrative claims data is going to be sufficient to capture this measure in totality or if one is really going to require more sophistication as probably is present within Ingenix capabilities?

DR. SCHWEBKE: That's a great question. I think that laboratory tests are actually one of the data sources where administrative claims does extremely well.

Now, you know we have done before, not with this specific measure, but we have done before a chart review process. Now if we assume the chart, the paper chart is the gold standard. We haven't done this with an EHR system. Where we've compared the output from looking at administrative claims to the chart review abstraction process and found that there were certain aspects of care where

administrative claims not only matched the chart review, but actually did better. And this was actually alluded to by I think Dr. Solomon where we found out with laboratory tests because they are sometimes done within the facility as well as sometimes done at a reference facility. Laboratory tests actually had a better capture rate than chart review.

And so I think that this is one of the areas where we can feel confident that our data collection is complete.

CO-CHAIR THIEMANN: Dr. Nagamine?

DR. NAGAMINE: Thank you. I have a question for Dr. Solomon. I'm trying to get a sense. I'm an internist, and I'm trying to get a sense of out of all of all these people who develop leukopenia or transaminitis, what is the incidence of harm? Like, I see a lot of patients who chronically have low white cell counts or high LFTs, but how many of these people die? Do we know the incidence of -- morbidity/mortality rates on these?

DR. SOLOMON: I think that's a great question, and it gets back to the earlier comment about what's the broad public health importance of these measures. As a rheumatologist who sees lots of rheumatoid arthritis and uses these drugs often, it seems very important to me. But I think what your question is asking for is what's the prevalence of the real harm that's caused by these drugs and you know, that secondarily would monitoring in an enhanced way or making it a quality measure really improve outcomes.

DR. NAGAMINE: Right. Right.

DR. SOLOMON: And I don't think -- we don't know the answer, you know in fact to your question. I mean people do die of hepatotoxicity. You know, there's cases reported through MedWatch and there's cases that aren't reported to MedWatch. But these people do die and have significant harm from these issues. It's rare.

And, you know Dr. Schwebke

mentioned 66 percent screening. I think, you know if we drill down this data my bet is that 66 percent of people do it within these frequency intervals, but it's probably 80 percent that do it at some point. Let's say let's make the frequency interval six months. I think the proportion that comply is much higher. And I don't know that for sure, but I've looked at these sorts of data at our institution and we have similar 60 to 70 percent are in this range. But if we loosen it to within six months, it goes up to 80 percent. And, you know, we see -- we have 3,000 rheumatoids at Brigham and Women's Hospital and it's been a long time since we've seen a death from any of these because we do reasonable monitoring and if people have abnormalities, we change dosing. So you might say well that's evidence of success of the monitoring or you might say it's not such a big issue. And I just don't know because we haven't really done the appropriate studies to

determine whether the lack of bad outcomes is because of the monitoring frequency or whether it's just because it's not so common.

DR. NAGAMINE: Thank you.

DR. SOLOMON: Did I answer your question?

DR. NAGAMINE: Yes, it does.

And along those lines I had a question for you as well in terms of the interval. You know, when you initiate a drug you monitor them frequently and then you taper off after they've been on it for a while and shown to be stable. So if someone has been on these drugs for years and never had a bump in their LFT or a bump in their white count, why would you continue to do it Q3 months because --

DR. SOLOMON: Because it's a debated point, honestly. And there are data that these are idiosyncratic reactions that could happen anytime. And I could find you case series of people that have these after

many years of drug. And so I don't think we can be absolutely certain that three years without a problem means never a problem.

Having said that, it's probably the case that the prevalence of problems does go down over time.

DR. NAGAMINE: Okay. Thank you.

DR. SOLOMON: But it doesn't probably go to zero.

DR. NAGAMINE: Okay.

CO-CHAIR THIEMANN: Steve Lawless?

DR. LAWLESS: I'm a little bit confused, and maybe you can help me and the two reviewers. Is the drug, is there anything different from being on methotrexate on taking the drug and having this recommendation being versus on being on methotrexate and having rheumatoid arthritis? So, are we selecting out a population here because is it the drug that you're monitoring, is the population and the drug you're monitoring? So is there something more prevalent that someone with

cancer on methotrexate is not being monitored with the same recommendation?

DR. SOLOMON: I mean I can answer for RA and Chris can answer for GI.

I mean, we've studied this dosage of methotrexate most intensively in rheumatoids or outpatient once weekly methotrexate is used for certain indications, RA being one of the primary indications. And so we have a lot of data around that. And there's support in the rheumatology community around these monitoring -- doing monitoring, and again as I said the exact frequency I think we could probably find people on many sides of the argument. But people believe that it should be done at some frequency.

I think in the cancer, it doesn't apply at all for cancer where the dosing is tenfold and patients have a completely different set of issues.

So, I think it's an interesting question. I think that the measure pertains

to a group of people who use the drug and who we've studied. But I don't know how it would apply to other groups.

I mean, with IBD and such, you know there's obviously other liver issues and so the toxicities may be may be more accentuated. And the same thing goes for psoriatrics who take methotrexate because they have a higher incidence of metabolic syndrome and stiata hepatitis, et cetera.

So, I think that it's safe to stick with a drug and an indication where we do have some data, not perfect data, but I don't know that it wouldn't apply to some of these other conditions as well.

Chris?

CO-CHAIR THIEMANN: Dr. Solomon, Dr. Kowdley is not present. So from a GI perspective, we wouldn't be able to get their perspective.

DR. SOLOMON: -- has commented on this in our drug safety working group that the

data are slightly different in the IBD population with greater toxicities.

CO-CHAIR THIEMANN: Dr. Muething?

DR. MUETHING: Mine's a follow-up question. I think I just really want to make sure this is clear in my mind is that it sounds like we have strong evidence that there's variation and frequency of screening. But I just want to make sure I understand correctly. We don't have published evidence that improved screening improves outcomes?

DR. SOLOMON: Boy, I'm not aware of any evidence. I don't know that I've systematically searched the literature for that this year, but I probably did it about two years ago, and I didn't see anything. And I'm not aware that there's evidence that outcomes are improved based on frequency. Again, that's kind of an expert-based opinion without any sort of prospective data.

DR. MUETHING: Right. Thank you.

CO-CHAIR THIEMANN: Dr. Schwebke,

from the perspective of is there an exemplar within the Ingenix database that Ingenix has been able to demonstrate changes in patient outcomes based on examination of its database?

I know on another measure during maintenance you had altered one of the measurement time periods and saw an increase in compliance. Is there anything that Ingenix has done to follow that trail to demonstrate improved patient outcomes using its database?

DR. SCHWEBKE: We have not specifically looked at this measure to see if there's a difference between the population who had monitoring and the population that did not have monitoring. It would be a little bit challenging because the patient population that doesn't receive monitoring might be a different population, and that might be difficult to define and really identify with clarity using claims data. But we have not specifically looked at our data to address that question.

CO-CHAIR THIEMANN: I have an additional follow-up question. In the performance measure application opportunity for improvement, in this measure and in other measures you've indicated that through endorsement of this performance measure would improve medication compliance. Do you have anything in your experience in working with Ingenix that would actually demonstrate the patients, although they have a filled prescription, that they actually take the medication realizing that labs do reflect that, but that if there is an actual comparison that you've done?

DR. SCHWEBKE: Actually, that comment was not intended with that purpose in mind. That comment was more along the lines that if someone is having a problem, let's say they're on methotrexate and they've now developed an anemia with side effects, that the nature of side effects is often a driver to people not taking their medications are

prescribed. And that monitoring might identify reversible side effects that could be addressed through various means, like dose reduction that could then improve medication adherence. So that was kind of really the intent behind that statement.

CO-CHAIR THIEMANN: Thank you.

Dr. Nau?

DR. NAU: Just to follow-up the question you just asked. The answer is that administrative claims data for prescription fills are a pretty good proxy for actual use of the medication by the patient, and studies have borne that out. So I think that's where the requirement in the measure that the patient be actively on the medication I think can be relatively accurately inferred from the claims data.

CO-CHAIR THIEMANN: I don't see any other cards flipped at this point.

I think it's the will of the Committee to return back to Dr. Nau's original

question, earlier question about what is our focus and how we would interpret importance to measure, whether it's on a broader perspective or a more narrow perspective specific to the individual population. So, I'd like to open it up to that discussion there from the Steering Committee.

DR. LAWLESS: This is Steve Lawless.

Let me second that. I think that is an excellent question. When I saw looking at the various measures, I saw bundles. And so one I saw bundles and disease -- and the burden of reporting. And then I got to thinking, does this open up a Pandora's box that does NQF want to use these kind of measures as a way for people to justify the testing of the measures.

And I don't mean ill-intent, but if a measure has a 66 percent compliance rate in a group that's most wedded to this, I think the intent is either a research focus

eventually for people or a way to try to create sticks rather than coming from curiosity about whether this works or not.

And so, I have to think from a disease standpoint is are we worried about the drugs, are we worried about the population, or are we worried about a specific element that is more of a research focus?

CO-CHAIR THIEMANN: Iona, I think you were next.

MS. THRAEN: I support what has just been said. Also, I had a couple of struggles.

One, it struck me with several of the Ingenix specifically that -- and I'm getting a doctorate in medical informatics, so I'm sort of speaking out of both sides of my mouth when I say this and I apologize for that. That just because we can doesn't mean we should. And in some of these instances some of these indicators I didn't see the clinical evidence to drive the need. I saw

the technical infrastructure that could make it happen, which is great, but I felt like there needed to be stronger clinical evidence that that should be the focus and the infrastructure secondary to that clinical rationale in terms of accepting or endorsing or not endorsing. And so that was my struggle.

So, I got really excited about the fact that Ingenix could do all this work. But then when I read further on the technical comments, which is why I raised the question of AGA versus ACR, you know are the clinical societies really supporting this as a need and either an opportunity for improvement. And then I sit in government so I always think anything that gets approved here or gets endorsed here at this level, Medicare, Medicaid and Public Health is going to adopt. And so I'm thinking the accountability of this side question. And as I was going through that I also felt, and I'm not a

clinician, but I felt invasion of privacy in the sense that at the level of detail of monitoring some of these practices, I really saw an invasion into the patient and clinician relationship. I mean, almost down to the point of -- and I know I'm speaking out loud here, I probably shouldn't be doing that. But this notion that there was an invasion in the practice relationship; now maybe that's what we should be doing theoretically is monitoring that practice relationship. But my mother who had rheumatoid arthritis who was on these drugs for many years, when she was first put them was advised that there was risks associated with them. And then, and I know this is idiosyncratic to me, but you know she had that knowledge and they worked out the monitoring relationship based on her experiences.

And so, I really struggled with this notion of safety versus quality improvement versus public accountability. And

I didn't see that it qualified as a safety issue in many of the cases, of the individual cases, it was more quality improvement. And then public accountability then is sort of waiting to see what everybody else is going to recommend, and then we're going to adopt them and put them out there for public review.

So, I struggled with this whole set in general.

CO-CHAIR THIEMANN: Mr. Levine?

MR. LEVINE: If I recall correctly, overuse is a national partnership priority. I don't know the costs of these tests, but certainly if we consider within the context of an overuse paradigm, certainly the frequency becomes an issue. I just want to mention that. Maybe that's in line with public accountability.

CO-CHAIR THIEMANN: Dr. Nau?

DR. NAU: Well, and maybe I can direct my question to Dr. Angood or others on the NQF staff to speak to this issue relative

to the other projects that you've done. Can you tell us if you've given more specific directive to other groups in terms of what perspective to take, or is there a sort of precedence here of what perspective we should be taking when considering importance of these?

DR. ANGOOD: Well, this is a topic almost bordering on ethical discussion type of thing. And I don't think we'll come to an answer today. I'll ask Heidi to make some comment as well.

But as NQF as evolved, it is looking for how to refine its approaches and continue to get toward quote, best in class measures that are out there. However, within the NQF staff we don't have the depth of expertise for every measure to be able to provide the scientific expertise on whether or not that's the right type of measure, et cetera. So, that's why we utilize Steering Committees and TAPs to provide us that

scientific expertise.

Now as an organization do we focus on the disease, do we focus on the patient, do we focus on the broader public health components? Well, it's kind of all of the above, isn't it? And it's difficult, therefore, to make these judgments. So that's why the guiding principles of the criteria for accepting a measure are there. To try and keep you focused in on the merits of that particular measure most specifically for who is going to be utilizing it most frequently and does it meet those criteria.

If we stepped back and started doing public health, and is this the right thing and get into all those others, it gets really kind of muddy and murky. So I would encourage you to just stay focus as best possible on those criteria.

But, Heidi, do you want to add some other comments?

MS. BOSSLEY: Sometimes it helps

just to do --

DR. SCHWEBKE: If I could just answer, there was also a question about precedence. And there actually is a precedent specifically for some RA medication monitoring measures just endorsed earlier this year as part of the Enriched Administrative Claims Project. There were two or three measures specifically in the RA population, specifically for people on specific RA medications looking for monitoring of various lab parameters including transaminitis. The difference is that those measures were focused on individuals who are just starting these medications. And our measures are focused on people who are chronically taking these medications.

So, people are looking for precedents as they kind of struggle with this difficult issue are there precedents there.

MS. BOSSLEY: So I would just add sometimes I find you hit a point where you're

just not sure where you are. And I think we just need to do probably a poll, and maybe do the subcriteria under importance, because I think that's what you're struggling with. And let's see if you think it conditionally, partially, minimally meets it. And then I think just do a vote on whether you think it passes importance.

To me it always comes down to, does this measure inform consumers? Because that's ultimately what we're looking for. And does it meet the criteria in importance. And that I think is your immediate question that you all need to probably just vote on, and let's see where you are and go from there.

CO-CHAIR THIEMANN: Dr. Nagamine?

DR. NAGAMINE: I just wanted to give one other perspective on the context question.

If our objective is to inform consumers, you know it's sort of a numbers game and sort of an epidemiologic, or you

could take that approach as well.

On the one hand you could say if rheumatologists wanted to do better, certainly this would provide some guidelines to do better. And if you take consumers in general, what we're looking at is about 2.1 million I think have RA, if that's correct. And so, you know that gives me some context. But the impact and the safety question; high volume, high risk are other things that I think about in a safety measure. And there's some volume, but I'm not sure what the risks to not doing a CBC and an LFT Q3 months.

DR. ANGOOD: A useful basic model that I often fall back on is just that; the risk severe and the volume of that severity. So is it three people but high risk, or is it 10 million people but low risk? And you sort of construct that in your own mind as to what's the meaningfulness. And you know, you may not be a rheumatologist, but you can sort of get some sense for any of these, and other

measures. You know, the severity of risk and the volume that it impacts.

CO-CHAIR THIEMANN: Dr. Conway?

CO-CHAIR CONWAY: Yes. I had a similar but a little bit different reaction in going through all these this weekend. And it just struck me that the real opportunity is have an integrated approach to the monitoring of immune modulating drugs in inflammatory disease. And what we've been served up because of the methodology here is this kind of fragmented collection of proposals. And I was frustrated because it would be great to turn all of this over to some pharmaceutical or think tank organization to put this together in a more logical way. And what disturbed me was we've got a bundling of drugs that are completely different medications. And instead of timing intervals that sometimes don't make sense and there's different specialty societies in disagreement, and the whole area it looks fertile. I think this is

an area that we could probably reduce variation and standardize our approach, I'd say as a profession and a nation. But it requires a different approach to this than the way we've been served up all these things.

It looks to me like this just isn't really for prime time, and that crosses about six of these categories.

CO-CHAIR THIEMANN: Dr. Lawless, I see your name tag going up.

DR. LAWLESS: I'm going to have to ask NQF because you made a distinction about this. These are entitled Patient Safety Measures. But then you imply population. And it means a lot different from people as a priority and everything else. Are we evaluating these as a population safety trend or a patient safety measure from your perspective?

DR. ANGOOD: Well, again, I think that's quite honestly difficult to answer. We're hearing so far in the discussion some

sort of pros and cons to these measures, not just the one we've talked about but the clustering. And our primary discussion and points internally has been to put them in the patient safety cluster. Will they make safer care for those patients with these diseases?

MS. BOSSLEY: I would just add though, and Kay maybe you can remind me because I don't have the measure up specifically. These are intended to be reported out, though, at the individual clinician level and then roll up, but not specifically at the population level. So, I think the focus starts very narrow on individual practitioners. Does that make --

DR. LAWLESS: Well, no. Because you just said it rolls into physician-specific on the reporting, and that's not what I heard before. So, how would this link back to the physician?

MS. BOSSLEY: So, this measure as it stands right now, and Kay, correct me if

I'm wrong because I may be wrong. But all of the previous measures that Ingenix had put forward are intended to be reported out at the individual clinician level. It can then be rolled up into group practice and everything else. So they are intended, it's more I would say a patient safety focus as opposed to reporting out at the population level. Does that makes sense?

DR. SCHWEBKE: Most of that is true. The unit of analysis is the patient. And we do have many customers who use this measure as part of care management disease management where they're directly interacting with patients and making sure that they're connected with care. But then we have about 40 percent of our customer base is using them to look at a quality performance either linking to providers, to clinics, to regions to see if there's areas where there's variation, to see if there's areas where perhaps they need to address certain quality

issues.

So just reporting out at the level of what's happening to the member and then depending on kind of how the customer needs to use that information, be it interacting with patients, giving information to patients or trying to measure performance at the level of the provider or rolling it up, as you mentioned; all that flexibility is there.

CO-CHAIR THIEMANN: And Dr. Schwebke, this is just as a follow-up to the public reporting component. In the application it was my understanding that Ingenix does not have any information associated with the use of this measure or some of the other ones in public reporting initiatives. So my question is what is your perspective about the true applicability of this measure for public reporting since that is one of the elements?

DR. SCHWEBKE: Well, my sense is that you're right, we have customers using our

tool. We don't exactly know which measures they're using or how they're using it. We actually are in the process now of trying to gather that information so we can submit that in the future.

But my sense from talking to various customers is they are finding this useful to give to share information back to providers so that providers can see how they're performing compared to others.

And I also do know that they are sometimes used to try to identify the quality of care that providers may be providing.

CO-CHAIR THIEMANN: And I think Dr. Kennerly was next.

DR. KENNERLY: I wanted to see if we could maybe integrate some of what Dr. Conway and Dr. Muething have both articulated in terms of the notion that if what we are really asked to do here is to be creating ways to judge the sufficiency of practice, I guess the question then is do we have evidence that

if a physician fails to follow this pattern, that they are not meeting standard of care? And I think really what we've heard is maybe, but I don't know that we have a sense of what risk if we establish this as a standard of care, that we really would have the sense that someone is practicing out of the bounds of sensible medicine. Because of the lack of testing I think and of looking at outcomes of those who have failed to have that level of follow-up at this point.

So, I think it's a fairly harsh criticism, if you will, to be able to -- you know, I mean again from the quality improvement perspective maybe, but I guess I just feel like this as a group of them I think are not as persuasive with regard to making individual judgments about a physician. And that certain of their patients may fall out for a variety of reasons. And I think I just have some concern that we're in a sense, permitting judgment about something that has

relatively modest evidence of benefit to the patient population that we're focusing on.

CO-CHAIR THIEMANN: Forgive me, because I don't know whose card went up next. But I'm going to go Dr. Nagamine, please.

DR. NAGAMINE: Along those lines I was going to circle back to Iona's comments earlier. So if her mother and her rheumatologist agreed that she didn't want to drive in for Q3 months CBCs, would her doctor be dinged for not doing them, and could her doctor be dinged by the insurer saying you don't meet our standards. You're not practicing within the recommended guidelines, and so therefore you're not part of our group. And could she lose her doctor that way?

So, I think that's the downside. I mean, I'm not saying that we should not diligently monitor patients. My sister-in-law has severe RA. But I think clinical practice guidelines and national standards are a little different because of that piece of it. And

they do potentially set you up for that downside.

CO-CHAIR THIEMANN: Mr. Bunting?

DR. SOLOMON: I'd like to comment here. You know, when I head of the Quality Care Committee at the DCR we had these same conversations about two or three years ago. And a bunch of rheumatologists decided that these quality measures were worth putting in place, but we worried about all of the same issues about is it affecting enough patients, is it dinging doctors, is it dinging patients, is it unfair that we said -- you know we got to set a bar and it's a middle bar in our minds for how to treat RA. I mean, it doesn't tell you if they're getting good RA care, it just tells you something that you can measure.

But I'll stop there.

CO-CHAIR THIEMANN: Dr. Solomon, just as a follow-up, since ACR reached an expert opinion consensus on these guidelines are you aware of any pilot testing that may be

done within various rheumatologist practices associated with these guidelines so that there might be in some future time some data which demonstrates that monitoring on a certain prescribed timeline with these medications improves patient outcomes?

DR. SOLOMON: I haven't been head of the Quality Care for the last year, so it may be that that's happening. I know there were discussions about having a research agenda to move our process measures to valid outcomes measures. So, it may be that's happening, I just don't know right off.

CO-CHAIR THIEMANN: Thank you.

Mr. Bunting?

MR. BUNTING: My comment is not just about this measure, it applies to all. But since we're starting with this one, I think what I'm hearing is what I wrestled with over the last couple of days when I completed the survey. And that is if you're strictly interpreting the NQF rules, or the

regulations, the guidelines, the definitions then we try to evaluate whether something is completely met, partially met, minimally met, and we assigned it to those buckets. But then you're asked, you know do you recommend this measure.

I followed strictly those things. If it met, or partially met or minimally met I recommended it. But I think what you're hearing now and what is evident based on my analysis of the Excel database that we have a privilege of seeing today, is you have a large number of people who are saying it met the criteria, but then they're voting no. And I think that's what I wrestled with over the last couple of days is that it meets, but I'm not enthusiastic about it.

I can understand the benefit, not just on this measure but some of the other measures, but if I were a physician or in charge of an office practice, how much time would I invest in this, would this be the

thing that I want my organization to pursue? For a lot of these measures the answer would be no. I think it meets the criteria. I think it's measurable. I think it has value. But then the question is how much value does it have and how many resources am I going to allow for this type of measure.

CO-CHAIR THIEMANN: Mr. Levine?

MR. LEVINE: Yes. I was curious whether there's any data in terms of other countries? Whether there's any kind of standard in terms of practices? I mean I don't know how international the rheumatology community is, but I'm just kind of curious.

CO-CHAIR THIEMANN: Dr. Solomon, would you have any comments on that?

DR. SOLOMON: I'm just thinking of studies that I've seen about monitoring. And honestly, the vast majority come from large U.S.-based cohorts. I don't think the rest of the world is so wrapped up in this. But a lot of the data comes from administrative claims

data sets that we can get our hands on in the U.S. and some other cohorts.

And I'm just thinking right off if I can recall large cohorts of non-U.S. I can't think of any right off. But that's not a systematic review of literature, that's just what one person can remember.

CO-CHAIR THIEMANN: Dr. Nau?

DR. NAU: Well, I think we're having a really good discussion on this, and I think it's worthwhile. Because I think this really gets to the fundamental issue of what this whole Committee is trying to accomplish and what NQF endorsement means.

And I think that Heidi brought up the issue of consideration of consumer reporting. And I would say that if we use that criterion, then all of these measures are dead in the water, as are most of the already endorsed NQF measures because none of them are really perfectly suitable for direct public reporting that could be interpreted and used

directly by a consumer.

So, I think the issue here is trying to find the right balance point of what's going to be useful for improvement and could some of these, perhaps, be rolled into some overall assessment that maybe could be helpful to evaluating overall safety of care for patients with the relevant disease. And so I think that's where it's tricky to find the right balance point of how much is enough. And then the importance issue, part and feasibility issues largely become contact specific. You know, some things may be very easy for one organization to use, it may be difficult for others, it may be useful for rolling up at a physician level but some may not.

So, I think that it's tricky. And I think what Heidi was trying to suggest earlier is maybe we just move forward acknowledging we've got these different perspectives and potential different

utilization here. And so I think probably we should move on now that we've kind of got a sense of the different perspectives people take. But, you know I think each of us is going to have our own impression of -- you know, and how we vote based upon the context from which we come from and the world in which each of us functions.

But I think it was a good discussion. And it's been helpful for me to kind of appreciate the different perspectives of the different Committee members.

CO-CHAIR THIEMANN: And on that point, I was going to circle back around to Heidi's recommendation and start to look at the various subcriterion, and work through that. Although the TAP has already previously weighed in on those areas, the Steering Committee members were also asked to evaluate all these. So I think we go through. Let's work with 1a Demonstrated High Impact Aspect of Healthcare associated with this performance

measure. And see if we want to do a straw as to where people fall out on this at this point.

I actually would like to wait for Dr. Conway to come back.

MS. THRAEN: Could you just review the number system for the voting again? I didn't take that down. I'm sorry.

CO-CHAIR THIEMANN: Actually, for the individuals I don't believe we're going to do the keypad for the individuals. So criterion we're only going to use the keypad for the actual endorse, not endorse or endorse with modifications or abstaining. So when we actually work through all of the four criteria, we'll go ahead and then take a vote for whether or not the Steering Committee makes a recommendation for endorsement. But we'll just do hands poll for the individual items.

So, for section 1a Demonstrated High Impact Aspect of Healthcare for

importance to measure and report. I'm going to jump in since we've had such good discussion associated with importance to see if we have any individuals supporting that the performance measure completely met this subcriterion.

I see a puzzled look.

DR. NAU: Well, are you asking whether we think overall it met that category or whether we're rating it as completely, partially, minimally?

CO-CHAIR THIEMANN: I was actually doing each sub. I'm happy to do the overall if people feel that we're ready to do the overall importance. But in some ways I thought that there was some need still to actually interpret the high impact possibly and how the individuals on the Steering Committee may interpret that definition, and how NQF defines high impact. So that's why I was gravitating towards the sub first and moving through each of those. And then we'll

do a collective as to whether the Steering Committee feels that the measure developer has demonstrated importance in the overall category. Okay? Okay.

So, is there any needed additional discussion on this one 1a subcriterion, the summary of evidence of high impact for this performance measure within healthcare, or does the group feel that we could move on to going ahead and raising hands on whether or not the performance measure completely, partially, minimally or not at all met that subcriterion?

Okay to take a poll? Okay.

Any individuals who for the Steering Committee who feel that the performance measure completely met and demonstrated that there's a high impact aspect of healthcare for this performance measure? I'm not hearing any or seeing any.

Does the group feel that the performance measure partially met? I'm seeing one, two, three, four, five, six.

DR. SOLOMON: And I'm raising my hand.

CO-CHAIR THIEMANN: Seven. Great, I was going to ask Dr. Solomon since I can't put a visual on you.

And minimally? One, two, three, four, five, six, seven, eight, nine.

And not at all? Not seeing or hearing anyone.

Moving on to Opportunity for Improvement. How does the group feel? That the performance measure completely met the burden to demonstrate opportunity for improvement? Seeing none, no hands and not hearing Dr. Solomon, that's a zero.

Partially met? One, two, three, four, five, six, seven.

DR. SOLOMON: And me.

CO-CHAIR THIEMANN: And Dr. Solomon. Great. Terrific. I was pausing to see.

Minimally? One, two, three, four,

five, six, seven, I believe. And I think that's a total for present. Not at all? One. Sorry. Dr. Lawless.

The Outcome of Evidence to Support Measure Focus, 1c. Completely met? I'm seeing zero and not hearing Dr. Solomon, so zero.

Partially met? For Outcome of Evidence to Support Measure Focus partially met, anyone?

Minimally met?

DR. SOLOMON: I'm saying minimal.

CO-CHAIR THIEMANN: Minimal? Okay. One, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, I believe.

And not at all? Two? Okay. Thank you.

Elisa keep me on track for totals.

And then so now we are evaluating whether overall the Steering Committee feels that this measure is important to -- has met

the burden for threshold to proceed on with the evaluation. Has the threshold criterion been met by the Steering Committee? If you're answering yes in support, please raise your hand. I have two hands.

And if you're answering no, please raise your hand. One, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve.

And Dr. Solomon?

DR. SOLOMON: I would say yes.

CO-CHAIR THIEMANN: You would say yes. Okay. So I think that increased to three with yes.

And any abstaining? One. Thank you. I didn't see that as an option, so I didn't ask it. Okay.

So, given that the majority -- is it the consensus of the Steering Committee then that this performance measure PSM-017-10 did not meet the burden to pass the threshold for importance to report; to measure and report? Sorry. I believe that is the take on

numbers. Okay. Great.

MS. BOSSLEY: So just so you all know what has occurred and if you're on the phone. This measure now will not move forward. You won't vote on any of the other criteria, and your recommendation is to not to recommend for endorsement.

CO-CHAIR CONWAY: If we can move on to PSM-018 titled Patients with rheumatoid arthritis taking methotrexate or sulfasalazine that had a serum creatinine in the last 6 months.

Lisa is the primary reviewer for this.

CO-CHAIR THIEMANN: Thanks, Dr. Conway.

The performance measure PSM-018-10 titled "Patient with Rheumatoid Arthritis Taking Methotrexate or Sulfasalazine Had a Serum Creatinine in the Last 6 Months Reported". This measure has a lot of the same similar characteristics to the measure that we

just spent an extensive amount of time discussing. And so I'm not sure that it's really necessary to give a full report on this one, such as Dr. Solomon did.

The Technical Advisory Panel did indicate that there was minimal evidence for importance. They did describe the Ingenix reliability testing internal to its own database, which were consistent with my evaluations of that.

And then talked about the use of the expert consensus guidelines and so forth.

So, I'm going to be very brief on that given the past discussion that we just had, unless anyone of the Steering Committee has specific questions regarding the performance measure from my presentation. Otherwise, I think we should open it up to questions to the performance measure developer, if any.

CO-CHAIR CONWAY: Or do the secondary reviewers have something to say, Dr.

Solomon or Kennerly?

DR. SOLOMON: No, I have nothing.

CO-CHAIR CONWAY: Okay. How about the measure proposer? Are they still on the phone?

Should we move on to then voting on the importance of the measure to report, we'll do it by the three sections?

Excuse me, go ahead.

DR. NAGAMINE: I have one question. I'm sorry.

CO-CHAIR CONWAY: Okay.

DR. NAGAMINE: I have a question for Dr. Solomon about the incidents of renal failure on these drugs. From what little I know about these drugs, creatinine is less of an issue than LFTs, is that why the interval is six months?

DR. SOLOMON: The renal failure is very uncommon. I think it's really more the fact that if the creatinine clearance is changing, that the dosing should be reduced.

DR. NAGAMINE: Got it.

DR. SOLOMON: And that the value is that every six months because that's unlikely to change rapidly --

DR. NAGAMINE: Okay.

DR. SOLOMON: -- unless there's some other illness.

DR. NAGAMINE: Okay. Thank you.

CO-CHAIR CONWAY: Are there any other questions?

Okay, let's take a -- oh, sorry.

DR. MUETHING: I apologize. This is another clarifying question following yours, and thank you for asking about that. Because I don't know about the incidents of problems with this with these drugs. So just to be clear, so is it if I'm the physician or the provider caring for a patient and prescribing these three drugs, if I do not know the creatinine clearance am I potentially causing trouble for this patient in my prescribing habits?

DR. SOLOMON: Yes.

DR. MUETHING: And then the six month issue is that was drawn because some reasonable belief that it can change over six months and that time period is a reasonable time period that I should be aware of the most recent creatinine clearance?

DR. SOLOMON: Yes.

DR. MUETHING: This feels different than the last one, in that it feels like I should know this if I'm going to be prescribing these three drugs.

CO-CHAIR CONWAY: Other questions?

DR. NAU: Sure. And I guess the issue with safety here is perhaps twofold for monitoring the creatinine. One is, does the methotrexate create renal impairment, and also does a change in creatinine function then effect the clearance of the drug and thus create other toxicities as a result of the renal impairment. So, I think there's potentially twofold reasons for the monitoring

of the creatinine. I guess then we could debate over the frequency and whether that's the right frequency as in this measure. But I think there are multiple reasons that creatinine monitoring would make sense. It's just a matter of how important it is within the overall evaluation of care.

DR. LAWLESS: Since the measure is over age 2 -- is the population -- creatinine in most children is not a sensitive measure of the renal function. And the change in creatinine takes a long -- the renal function can decrease can significantly before the creatinine even changes. And I worry about creatinine as an indicator in someone who has got a chronic disease and also has a low muscle mass because the creatinine is also not a good indicator of renal function.

So, I think it's well intended, but it's not sensitive enough to pick up what they're intending to do.

CO-CHAIR CONWAY: Any other

questions?

DR. SCHWEBKE: Well, I certainly appreciate that comment. And I think that we've all begun to appreciate the limitations of the serum creatinine. But I think we need to keep in mind that KDOQI, who also recognizes the limitations of the serum creatinine also recognizes the need of monitoring the serum creatinine to calculate the GFR. And so all of the GFR is absolutely a better indicator of renal clearance. You still need that serum creatinine to calculate that value.

DR. LAWLESS: But you also need a urine creatinine, too. But I'm just saying that that is a measure in itself, the creatinine, just as a sensitive measure for that is not what is really considered a particularly good gold standard for a lot of the population you're dealing with.

It's something, I admit that. But it's not --

DR. SCHWEBKE: You got the urine creatinine to calculate the GFR. And actually KDOQI has been really clear about that. That it is absolutely appropriate to take a serum creatinine and to use that information along with the age of the patient, the gender et cetera to calculate the GFR without the need for a urine creatinine.

DR. LAWLESS: I'm sorry. I feel like a Tea Partier, and I apologize.

I also have a nephrology background.

If someone's urinary creatinine is not of a certain level, the creatinine clearance is not a good calculation. So I just -- you need it as a verification, especially in someone with a chronic disease.

So, I'm sorry, I'll get off my horse here for a second.

CO-CHAIR CONWAY: Other questions? Okay. Shall we get a straw vote of where the Committee stands.

First looking at the impact of this measure, how many feel the criteria were completely met? Okay.

How many feel that they were partially met? Okay. Three -- six, seven.

Minimally met?

And not at all met?

Dr. Solomon, how about you?

DR. SOLOMON: Partially.

CO-CHAIR CONWAY: Partially. Okay, we'll add that.

In looking at whether there's a gap that's been demonstrated in the measure that was submitted, how many feel that that evidence was completely met? Okay. None.

Partially met? Three.

Minimally met? One, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve.

Dr. Solomon?

DR. SOLOMON: Partial.

CO-CHAIR CONWAY: Partial. Okay.

Is that everybody? Very good.

And was a relationship to outcomes demonstrated in the measure that was submitted? How many felt that that was completely met? None.

Partially met? None.

Minimally met? One, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve.

And not at all? There were three.

And Dr. Solomon?

DR. SOLOMON: Minimal.

CO-CHAIR CONWAY: Minimal. Okay.

Now on the overall status of this measure, whether this is important to measure and report. This will be a yes or no vote. How many would vote yes on that? Okay. And how many would be no?

Dr. Solomon?

DR. SOLOMON: I would say yes.

CO-CHAIR CONWAY: Okay. We have, it looks like 12. What's the total here?

Looks like 12 noes and one yes, and a couple abstaining.

CO-CHAIR THIEMANN: Yes, two.

CO-CHAIR CONWAY: Okay. All right. It looks like that does not meet the criteria of importance to measure and report. Any disagreement with that among the committee members? Okay.

Very good. We'll move onto the next measure and pass back to Lisa.

CO-CHAIR THIEMANN: Okay. I did just want to say one additional comment based on comments around the table for the past two measures. Clearly there's some desire to reach to somehow measure this population. But that at this point in time I got the sense that the Steering Committee just didn't feel that these measures in the way that they were specified were going to get at what maybe was the original intent of the performance measure developers. And so from that perspective, I think that it's important to acknowledge that.

That there's still that desire to really look at this population and demonstrate methods for quality improvement.

So, the issue still does need to be looked at, and I think we would encourage Ingenix and the performance measures and the specialty societies to try and maybe come together in continuing to foster that issue.

MS. BOSSLEY: And I would add that when we write this report we won't just say you didn't recommend it. We actually do provide some information. So part of this will be, you know, and we'll look to you to help us draft exactly where you think these measures should go. Like what would you like to see the next time.

CO-CHAIR THIEMANN: And I think at the end of the day that's possibly some of the discussion that we'll have in wrap up/closeup of the day's activities.

So moving on to PSM-019-10 and primary discussion leader Dr. Lawless and

secondary discussion leader Dr. Solomon.

Dr. Lawless?

DR. LAWLESS: Yes. I'm really adding on to what we've been talking about. I just had a couple of extra comments. When I did the primary review that I thought in terms of this measure, and probably it is very applicable, also the other measures, the reporting burden I was struck by. It's very informatics database driven in terms of the coding of which patients, which exclusions. And so I thought the reporting burden, particularly with that line, was a little bit high. So that if someone wanted, who was not part of the registry, wanted to look at the applicability to their patients because, you know not available on reporting or data not available is a lot of times an indicator for people in public reporting that that person has something to hide. And I'm worried about the persons who said I just can't get this data, mu patients aren't part of a registry.

So that was there.

The 66 -- and it was addressed actually in the measure nicely, the 66 percent compliance rate. But again, no evidence of the outcome with it. Do those patients make any difference or not, did they follow them.

I also looked a little bit at -- and I didn't know how to work with it or not in terms of the importance or not, that it excluded patients who weren't on continuous benefits. I thought in an underlying way it was going to be an over-reporting, maybe people can't afford it. And so it implied already about the over use that someone else had brought up. That there was a cost associated with this, and who was going to take the burden of cost with this. And so those are the concerns I had there.

Again, and most of the other comments were very similar -- most of my other feelings were similar to the other measures in terms of what we've already discussed.

DR. SCHWEBKE: Let me jump in here and just address two of the concerns you raised because I think it's probably actual a clear understanding of how administrative claims work.

Now the first is excluding people of underlying benefits is critical because otherwise we have a problem called data incompleteness. And so what I mean by that is let's say that a member only had benefits the last two months of the reporting period. And if you're looking for that intervention and it's not there, it may not be there because the test wasn't done or the intervention wasn't completed, or it's also possible it was done but it's not captured because that individual didn't have benefits.

So, you know, the whole purpose o making sure that you have people in your measurement period with benefits is critical because you're counting on administrative claims coming through that will only come

through if that members has benefits or enrollment. And if you don't exclude those individuals, you're going to basically have misclassification, identifying people that have incomplete data.

The second thing is the burden of reporting is actually extremely low of administrative claims. That's actually probably one of the clear benefits of measures that use administrative claims is nobody needs to submit anything, no provider needs to be submitting, identifying your patient population or indicating that labs were done. That is all done through the processing of claims.

So with measures like this, the burden of reporting is low. I think where the confusion might occur is that in the denominator population it's been noted that these registry is a potential way to get into the denominator for this measure, but that's optional. And we include that only because we

have customers who do have disease registries, so they at least have that opportunity to move that population into measure if they desire so. But most people on this measure are not identified through disease registry. They're identified through the administrative claims. So the burden of reporting is extremely low.

DR. LAWLESS: And actually, thank you. That helped clarify a lot of it for me.

To enter the database, to get enrolled in the database is there either a cost or does IRB approval or anything. The database itself is captured, how would a patient know they're in that database?

DR. SCHWEBKE: The patient isn't aware. The health plan, the health plan is contributing to use identified data into the database as part of their contractual agreement.

DR. LAWLESS: Okay. Thank you.

CO-CHAIR THIEMANN: Dr. Solomon, anything additional? Any additional comments?

DR. SOLOMON: No.

DR. ANGOOD: I'll open it up to comments from the Steering Committee. There's no comments from the Steering Committee. It looks as if we are possibly ready to go into whether or not importance to measure has been met.

So, using the same process that we just recently did for the previous two, looking at section 1a Demonstrated High Impact Within Healthcare, does the group feel that the performance measure completely met that? Seeing zero.

That the performance measure partially met that? One, two, three, four, I believe.

And minimally met that? One, two, three, four, five, six, seven, eight, nine, ten.

And Dr. Solomon?

DR. SOLOMON: Partial.

CO-CHAIR THIEMANN: Partial.

Any abstaining? Zero. Okay.

Opportunity for Improvement. Completely met?

Partially met? Two.

Minimally met? One, two, three, four, five, six, seven, eight, nine, ten, eleven.

Dr. Solomon?

DR. SOLOMON: Partial.

CO-CHAIR THIEMANN: Partial.

Abstaining? One.

And for Outcome of Evidence supporting the measure. Completely? Zero.

Partially? I see zero.

Minimally? Okay.

Dr. Solomon?

DR. SOLOMON: Minimally.

CO-CHAIR THIEMANN: Minimally.

Abstaining, or no at all. Sorry. Not at all. I forgot not at all. Three.

Keep me on target.

Any abstaining? Now we'll go

abstaining. Zero.

All right. So the numbers.

And for the overall, does the Steering Committee feel that the performance measure met the threshold for importance to measure and report? Yes? I see zero. No?

Dr. Solomon?

DR. SOLOMON: Yes.

CO-CHAIR THIEMANN: Yes. Okay.

Any abstaining? Zero. Great.

So I believe that the numbers show that the performance measure will be not considered further at this point.

Moving on to performance measure PSM-020-10 for, I believe Dr. Kowdley is not here, so Dr. Knight I think will be stepping up for a primary discussion leader.

DR. KNIGHT: Right. Thank you.

CO-CHAIR THIEMANN: Thank you.

DR. KNIGHT: You know, this has overlapped with what we've already talked about. The differences are instead of

rheumatoid arthritis here, the focus here is inflammatory bowel disease. Methotrexate is included as it was with the previous ones we've talked about, but in this case we're also looking at azathioprine and mercaptopurine.

The incidence here is five to ten percent liver toxicity, which is felt to be reversible with stopping the medication.

The compliance is about 38 percent. And the difference here, they group methotrexate and azathioprine, mercaptopurine but some of the recommendations are fairly varied from the standpoint of consensus expert opinion on how often this should be reviewed. Perhaps one to three months on methotrexate; perhaps annually with the others. And so this recommended measure here is for a six month reporting period. So there's some significant difference between the first one we looked at looking ALTs, AST which was recommending every three months instead of six months. So

there's some certain differences there, but I think in general a lot of overlap again with what we've already talked about and the same sort of principles.

And I guess what I didn't see was, again, the real strong evidence seemed to be based much more on consensus expert opinion and difference was noted between rheumatologists and the gastroenterologists.

CO-CHAIR THIEMANN: Any additional comments from Steering Committee members? Dr. Nau?

DR. NAU: Well, I just wanted to ask the person from Ingenix to elaborate on the different monitoring threshold of every six months versus every three months within the patients who had RA and the rationale for those differences?

DR. SCHWEBKE: Yes, happy to do so. As the primary you had just mentioned there's a lot more inconsistency here between the sources that have recommended monitoring. And

when this measure was actually initially built to be consistent with our RA monitoring measure, we actually used a three month report period. We then set a consensus process working with AGA and a subcommittee that AGA had convened. And based on their input as national experts, they encouraged us to use a more conservative threshold of six months. That was then the final reason for us to change it from a three month to a six month intervention period.

CO-CHAIR THIEMANN: Dr. Lawless, I see your name card.

DR. LAWLESS: Just a question. Why age 12 was chosen? Because the inflammatory bowel disease goes down to younger, and I'm just curious.

DR. SCHWEBKE: Yes. That's an area were we felt have great data as far as what's the age at which we think most people are going to be diagnosed with IBD and placed on medication therapy. There's really little

literature.

So we did a couple of things. One, we looked at our database to see if we could identify when these individuals seemed to be presented and perhaps identify a population with IBD. So we based this threshold on our database as well as that of discussing the 12 year threshold with the AGA subcommittee.

DR. LAWLESS: And a follow-up question to that, because I know -- and I'm just speaking from the pediatrics world, there are two major inflammatory bowel disease registry groups. Are they included in support of the measure?

DR. SCHWEBKE: They were not and I don't know -- is that the disease registry through AGA?

DR. LAWLESS: Yes, the disease registry through AGA, and then there's also the Improved Car Now Network.

DR. SCHWEBKE: Yes. So actually that disease registry was built and launched

after we had developed this measure. And actually, we have met with the AGA subcommittee as that disease registry was being built.

CO-CHAIR THIEMANN: Dr. Nau, do you have another question or -- okay. I just wanted to make sure.

Mr. Bunting?

MR. BUNTING: There's a comment on page 22 that says it is difficult to understand how if the measure has been available since 2006 and used by other organizations, that there is not better reliability data related to this particular measure.

So, does this measure exist? And if so, why are we looking at it. And if it does exist, why do we not already have data? Is that a question for NQF or a question for Ingenix?

DR. SCHWEBKE: As far as the compliance rate -

MR. BUNTING: The comment says that there's not better reliability data. So what data does Ingenix have on this measure?

DR. SCHWEBKE: Well the data that we have is looking at our 15 million benchmark to determine the compliance. What we don't have is we don't have a direct chart review versus our administrative claims to be absolutely sure that we're measuring without a measure.

The other thing is this: We have repeatability in that we have looked at this measure in a variety of databases, but they tend to be kind of subsets of the same database. I'm not sure that's fair to say that's true repeatability. So, we at least have a large dataset where we have calculated compliance We have done a chart comparison review on other measures that identified that data collection for lab results is actually quite reliable with a low burden with administrative claims. And we know that there

is clearly a performance step.

CO-CHAIR THIEMANN: Any questions for Ingenix, or any additional comments from the Steering Committee members at this time?

MS. THRAEN: Okay. I'm going to have to ask you to repeat something you said in that response. Back to the point that you talked about you looked at it in other databases in relationship to this measure. Could you repeat what you said about that?

DR. SCHWEBKE: Yes. So basically we have several steps of testing, there are three main steps of testing.

We start off my identifying in this situation a 1,000 members that have inflammatory bowel disease. And we calculate their compliance. And then we actually go in and look at a random number of members with IBD who both passed and failed this measure. And then we look through the claims to make sure that we haven't missed something with our logic and we truthfully are capturing those

people that haven't inflammatory bowel disease who seem to have complete data based on their enrollment eligibility and have or have not had the -- things actually match what we're seeing on the output.

But remember, that is looking at the results based on dividing into the details of administrative claims. We're not going back to a chart or an EHR to confirm that. Okay. So that's step only one.

The second step is that we take a normal number database and we look at the same features that mainly are at this point focusing on compliance.

And then the third step is we're looking at a 15 million member database. And those populations overlap a little bit, so I don't think it's fair to say that these are three separate populations That they're kind of subsets. You know, the 1,000 member is kind of a subset of the 1 million member and it's kind of the subset of the 15 million

member. There's a lot of overlap, it's not complete. And that's why I think it's more fair to say that they are very similar databases.

Does that make sense?

MS. THRAEN: Yes. I just didn't understand the reference. Thank you.

DR. SCHWEBKE: You're welcome. I just don't want to be misleading and give you the impression that we have three distinct databases that would truthfully, you know I think be an indication of repeatability and reliability.

CO-CHAIR THIEMANN: Dr. Schwebke, a real quick question about the compliance, reporting compliance for this proposed performance measure is 38 percent. And with the gastroenterologist's opinion of this, I'm curious as to why the compliance rate isn't higher if they tend to be the individuals managing the patients?

DR. SCHWEBKE: We discussed that,

and there were some thoughts that came up. They believe that this was a true difference here. They supported this measure.

They believed that a lot of people who had been placed on medications and disappear to some extent. Have their medications renewed, maybe go down to their primary and the primary doesn't realize that the monitoring has been indicated. So they believed that this was real and, in fact, they were concerned enough about this that they believed that all of the compliance measures on our IBD measure list were measures that could actually be used for educational purposes not only for their specialty group, but maybe even more primary care practitioners who are also involved with the care of these individuals.

CO-CHAIR THIEMANN: Thank you.

Dr. Kennerly?

DR. KENNERLY: I think you, the Ingenix folks, have a unique opportunity here

a large dataset to be able to be looking at what the clinical outcomes are of patients who fail to meet these monitoring criteria. And it would seem that over a period of time, again with some hope that these patients would have continuous enrollment for a lengthy of time, you might be able to characterize those who failed to meet either three month monitoring or six month monitoring, or perhaps, heaven forbid, annual monitoring as they move perhaps from practice-to-practice or indeed from location-to-location.

And to look to see in the claims data whether or not there appear to be complications associated with failure to monitor. And I wonder if you might comment on whether either: (1) You have any of that data or plans to use what you have in order to begin to generate some observations that help with regard to the benefit or failure to monitor?

DR. SCHWEBKE: We've actually

talked about a variety of our measures, many of which are process measures like many NQF endorsed measures to try to establish ways that we can use our database to datalink, process outcomes, process measures to true outcomes.

It's challenging for a variety of reasons, one of which is that members often don't stay in the same health plan. And that's critical because a lot of these outcomes we might not see for a long period of time. As a member changes insurance, which unfortunately happens often, and a health plan typically only has on an average 24 to maybe 36 month about a patient, and you don't have often that time frame that you need looking at administrative claims data alone to answer that question and to really feel confident that you have the right answer without having a lot of member drop off. So that's just one of multiple limitations.

Moving forward we can say to think

about this, because we appreciate the importance of being able to identify, hopefully, what's important as far as an outcome. And again I say we, like many, are trying to aggressively look at how we can start pulling in more granular data like EHR data and other clinical data that might give us a survey information, may give us longer abilities to look at true outcomes.

Now assuming that if a member moves in their health plan that hopefully at least stay with the same provider. That may not be the case. But we certainly do continue to look at ways as new data become available to maybe answer some of the hard questions like this.

CO-CHAIR THIEMANN: Any additional comments, questions?

I believe we're ready to assess importance to measure and report.

Looking at section 1a High Impact does the group feel that the performance measure developer has completely met the

burden for demonstration of High Impact to Healthcare completely? Zero.

Partially?

Minimally?

Dr. Solomon?

DR. SOLOMON: Partially.

CO-CHAIR THIEMANN: Partially.

And not at all? Zero.

For 1b has the performance measure developer demonstrated an opportunity for improvement on this proposed measure? Completely?

Partially?

Minimally?

Not at all?

Dr. Solomon?

DR. SOLOMON: Minimally.

DR. ANGOOD:

CO-CHAIR THIEMANN: Minimally.

And for evidence supporting the proposed performance measure. Has the measure developer completely met that? Zero.

Partially? Zero.

Minimally?

Not at all?

And Dr. Solomon?

DR. SOLOMON: Minimally.

CO-CHAIR THIEMANN: And so is it the will of the Steering Committee that the measure developer has met the burden for importance to measure? Yes? I see a two and a half. We'll commit to three. So we have a three.

And no? Any abstaining?

Dr. Solomon?

DR. SOLOMON: No.

CO-CHAIR THIEMANN: Thank you.

So I believe the majority of the Steering Committee, the measure failed to demonstrate importance to measure and report. So we'll be moving on to, I believe, asking actually whether or not the NQF members or there any public comments concerning the four measures that were just considered?

MS. MUNTHALI: Operator, can you open the lines? I think they're open, but we just wanted to make sure.

CO-CHAIR THIEMANN: Iona?

MS. THRAEN: Yes. I have something that's just dawned on me, and I apologize for this. I've been operating under the assumption that many of these measures are the practitioner that's been involved with these measures are specialists, which my operating assumption is that specialists who are specialists in a particular area are practicing fairly narrowly and are kind of up to date, et cetera. It's an operating assumption.

What's the likelihood that some of these areas are going to be managed by generalists or family practitioners? Because I see there's a discrepancy in voting going on right now, it seems, that some more of the generalists are saying yes, we could use that kind of support in terms of the frequency of

monitoring, et cetera. And the specialists are pretty comfortable with managing it individually, you know not using sort of standard. So I guess I have to ask that question in terms -- and I'd like to get feedback from those who would see themselves in that role of managing these kinds of patients on an outpatient basis after a consultation or something, but that they're the ones who are actually doing the ongoing maintenance of the patients. What are your thoughts about that?

DR. KNIGHT: No, I think that's a great question. And that's, as I looked at this from a generalist standpoint as a family physician, I've looked at these and what does the weight of something being endorsed by the National Quality Forum, what does that do as a proponent of a measure? And I guess the thing I continue to struggle with, though, is the evidence and the cost benefit, and what's the expense of all the testing if we really don't

know in the long run that that's really affecting the benefit that we're looking for.

So, for example, on the last one I did vote that I felt that it should be added. And that was more because of that gap that was exposed there of only 38 percent compliance and that maybe there was a greater impact to that one than with some of the other ones.

But I think your point is well taken that there are going to be generalists around the country that personally I would refer to these patients and have them managed by a rheumatologist or a gastroenterologist. But I know that there are significant numbers of primary care providers around the country who may not have that luxury of a specialists available that would look to guidelines, recommendations from organizations like the National Quality Forum. So, you know, I think that's a great question, and then it all boils down to the cost versus the potential benefits

as we make these decisions.

And that's where I've struggled. There's evidence that really shows that there's a significant opportunity here.

CO-CHAIR THIEMANN: Dr. Kennerly, I believe you have name card up?

DR. KENNERLY: Indeed. I think, first of all, just in personal I'd like to thank Ingenix for submitting these. Because, obviously, I think they're trying to fill a perceived gap, and I think perhaps a real gap.

And I think the other thing that perhaps raises for me and the Committee in being new to this group is the degree to which we serve in a role of more actively trying to be filling the gaps. Meaning that we as a group as opposed to the community of metric builders who are going to look at theirs and submit them, and right they should, but I wonder if part of being more passive than that from the group's perspective, you know winds up with then less in the way of a message from

this group, if you will, to say here are some areas that we think would benefit from metric development and perhaps encouraging those sort of not just a broad call, but a more specific call that might look out at the priorities themselves and try to see if we could perhaps as a group be thinking about how we might have conversations that might help to shape what we received. So that in effect we don't necessarily just say "Gosh, send us what you have," and have good folks be spending time on doing that. But trying to sort of create some guiding principles, perhaps.

CO-CHAIR THIEMANN: Dr. Nagamine?

DR. NAGAMINE: I'm not a outpatient doc. I'm an inpatient general internist, hospitalist. But as a practicing physician what I look to are clinical practice guidelines which are evidence-based. Fortunately, I work for Kaiser and we have extensive research on what is the evidence and what are the standards out there. And so I

have that to look to. But that's different to me than an NQF endorsed safety or quality measure. It has different implications.

So, I think that that might be one way to look at it. Are you looking for guidance on what is standard of care on one level versus the accountability and insurer perspective, which has implications for exactly what you described: This understanding between your mom and her doc that she didn't want to come in for testing maybe as frequently as the guidelines say.

I think there needs to be some room for that. But where you go into a different bucket is when you have evidence that says if you don't do this, people will die or will be severely harmed; that's the category that I think I would want to be focused on. You know, the big stuff, the stuff that really matters, the stuff that really makes a difference. And we know that because there's evidence. Because we all know there's enough

that we could do, enough that we should do, but in this day and age of resources, we have to pick and we have to prioritize. And if there's not good evidence, it's hard to justify making something a national measure.

CO-CHAIR THIEMANN: Any other comments?

MR. LEVINE: Yes. I'm just wondering, you clarified that in your mind, at least, there's a distinction between NQF addressed measures and practice guidelines, perhaps, put out by the Agency for Healthcare Research and Quality, or some specialty organization. But I'm wondering if the public appreciates that.

And I mean my own view as a patient advocate and consumer, if NQF endorses something, I would see that as a clinical and practice guidelines. And maybe lawyers would too on both sides of the tort fence.

DR. NAGAMINE: And I guess I'm speaking from someone who practices as well as

has been a quality chief. And as a quality chief I have to look at where the resources go, and whether we do a failure modes effects analysis on a known risk or whether we collect data to report. Those are the choices.

And, you know, I know there's plenty of work to do. And so I just think that standards are well intended, but on the sharp end and locally in hospitals you have many competing priorities. And they're really all important ones. And so it's really important and critical that we can differentiate the stuff that kills people from the stuff that would be nice to do.

CO-CHAIR THIEMANN: Dr. Nau?

DR. NAU: Yes. I guess this gets back to the fundamental question I raised earlier of perspective. And I guess that's where what does NQF endorsement mean. Does it mean that these are things that everyone in the nation should be measuring and should be publicly reporting versus if a particular

entity, a particular group wanted to focus on trying to evaluate quality or safety of medication use in patients with RA, what would you look to? And in that case I would say some of these are irrelevant and important to look at if you're concerned about safe use of medications in patients with RA and IBD.

So I guess I'm thinking of it from the context of if we're interested in that issue, which measures would we turn to versus are these the most important measures in the world to evaluate and invest your resources in, which I'm sort of making the distinction of independent of resources and priorities nationally. You know, what are the appropriate measures? If you want to invest resources in a particular area, which are the most important measures to look at?

And so I think that's a little bit different perspective. I think from either standpoint you could argue that some of these aren't maybe the highest priority no matter

what the perspective. But I think that perspective is an important factor I think in the differences in the ratings around the table.

MS. THRAEN: This is to the NQF folks. Right now we're in a position of making this dichotomous decision, yes or no, endorse or not endorse. And in adding a level of complexity, which I don't intend to want to do, but this idea of recommending -- I mean, a lot of work has gone into evaluating these measures. And just sort of saying no and sort of trashing them to the side is uncomfortable for me. Because there is value in what has been done, but for a different -- maybe at a different level then what we're making this decision for.

So this idea of sort of a categorization of measures that says well this one we think is strong for public accountability purposes, safety risks. This one would be a good quality improvement

measures. This one might be a good feedback measure or decision support measure, or something like that as opposed to a yes or no, they're on they're off kind of decision.

And I don't know if that falls in with your scope of work in terms of what you're having to do, in terms of the Health and Human Services. But there's just so much work and value here that I just feel badly that we're kind of trashing it.

MS. BOSSLEY: Well, I don't think you're trashing it. But I think that's my personal takeaway from that.

But this is something that NQF continues to look at as measurement evolves. And originally and still now we're looking is the measure appropriate for public reporting or quality improvement. And public reporting should also involve internal quality improvement as well.

But there are efforts underway as we speak, I mean literally now where NQF and

the Quality Alliances Steering Committee, which is from the Hospital Quality Alliance, the Ambulatory Care Quality Alliance, multiple quality alliances are looking at is there actually more of a spectrum from internal quality improvement all the way to reporting out to the public. And I think clearly there is, just as you were talking about. And I think that's what everyone struggles with.

So, what is happening now is there's a final report that is going to the Consensus Standards Approval Committee, the CSAC here, with staff recommendations on to how to begin to split them out a little bit more and start talking about measures maybe within the spectrum in the process of being used for certification or recognition programs. It's being used for accreditation for payment programs and then full on to reporting. And we'll see what the CSAC and the Board says, but it's very possible that we will head more toward developers telling us

where they are in that scheme and that spectrum, and then evaluating whether again is there use in that measure.

We ultimately, I think, want to see measures continue to progress on that spectrum. You wouldn't see it at the initial endorsement, but you'd see it at the three year maintenance. We're not there yet, but I think we're headed there.

MS. THRAEN: And so based on that as we decline on many of these, then these would possibly be revisited as your bank of alternatives?

MS. BOSSLEY: Yes. I don't know when.

MS. THRAEN: That's fine. I get it.

MS. BOSSLEY: But, yes.

DR. ANGOOD: Well, and coupled with that, just sort of brought it out what Heidi was just describing, is that before the measures actually get to this stage, we've actually already been in dialogue with many of

the measures developers. Because staff review these measures, not just for the completeness of the submission form but whether the staff has their own concerns about is this going to pass through. And we have ongoing discussions with a lot of the measure developers through that. And when they get to this stage, then yes it's up to the Steering Committee and the TAP decisions, but most of them have already been through some dialogue.

So, as Heidi describes what I just said, we are in this interaction. It's not a yes, no or you're out of here. It's a dialogue because we're really trying to improve what's best for healthcare in the long run.

CO-CHAIR THIEMANN: Dr. Nau?

She has a smile. Mr. Lawless?

DR. LAWLESS: Just one question, and I'm really even coming from a curiosity more than anything else. The measures that we're all discussing today all come from

Ingenix. And so I'm a little bit curious, you described the process and process improvement. And they're all following the same format. So we're reading of all the measures the exact same format, same process. So getting through the review process and up through the -- did a lot of work, a lot of reviews. And it seemed like it just struck me with all the societies going on and all the push for patient safety how one particular group was successful enough to get X number of measures here this far when we're talking about that. Is the process onerous? I was wondering did they find the grail to get into the key here, or -

MS. BOSSLEY: We don't do much weeding in the way of -- you know, other than if we have a blank form, we're going to turn it down. If we don't have an agreement signed, we're going to tell them no. But beyond that, you all are the people who read through it. So what you see before you is what we received, other than the ones that

were withdrawn.

So in other projects you see a little bit more variation across the types of developers. This project just happens to be quite a few from Ingenix. You have a few from specialty societies, and so on. It's just -- this is unusual. Usually it's not just one large --

CO-CHAIR THIEMANN: I think it's also from a performance measure development, I'm sure which many people around the table understand, the length of time to develop a performance measure to even submit to NQF, and often times there needs to be some demonstration of broader consensus, not just the individual performance measure developer drafting the application and drafting the measure. That can take a couple of years to actually process the literature, do the literature, digest it, reach out, get comments and so forth. So it's a long time.

And then also I think it's also

complicated by when NQF issues a call for measures is a limited time. And so unless the performance measure developer is already at the end and has something ready in the bin to go, that can somewhat fall under the umbrella of the NQF project, that sometimes complicates what I think probably NQF sees, right?

MS. BOSSLEY: Yes. And I mean, we recognize that it's been develop for developers to know what's coming next because there hasn't been a nice schedule. We now have one related to maintenance. And it's kind of wrapped around that where we have endorsement maintenance projects. You're a pseudo one, you will do some maintenance in a little bit. You're not a full blown one.

But we have probably seven to eight topics per year in a three year cycle that we'll be going through. So cardiovascular and surgery are the first two starting, renal starts in January. And so we're hoping that that helps developers know what's coming out,

know what time frame they've got approximately. It's not going to be perfect, but it's probably better than it was.

Schedule, yes. We like to cycles, yes. We're doing cycles.

CO-CHAIR THIEMANN: This is somewhat off topic as well, but going back to that performance measure scheduling, the maintenance scheduling. And I think from a perspective of an NQF member participant that measure developers need to be aware that they can submit new measures during that performance measure maintenance phase, which probably the measure developers are aware, but maybe not necessarily the NQF members individually or as their individual associations are aware.

Any additional comments at this point?

I know we were scheduled for a 15 minute break, but then we also had a working lunch at 12:15.

All right. Do we want to take a five to ten minute break, and we'll reconvene? It's about 11:40 by my watch. Whether that's right or wrong. And so we'll reconvene at 11:50 and then start to work through at least maybe one more measure.

(Whereupon, at 11:39 a.m. the above-entitled matter went off the record and resumed at 11:52 a.m.)

CO-CHAIR CONWAY: And what we could try to do is see if we can get through the measure 21, and then break for lunch and do measure 22 during lunch. And if the discussion of 21 goes past 12:30 maybe we'll interrupt in the middle of that one and have lunch finished. How would that be as a plan? Is that okay? We have up PSM-021-10: Adult patients with multiple sclerosis taking interferon having a serum ALT or AST test in the last 12 months.

And our primary reviewer is Janet Nagamine. But before we take this section,

maybe Kay Schwebke from Ingenix would like to say some introductory comments about the whole measure set for MS.

Are you on the phone?

DR. SCHWEBKE: (Off microphone).

CO-CHAIR CONWAY: Kay, hang on. You're not coming through very well. You're breaking up, maybe try not using a speakerphone.

Hello, Kay?

DR. SCHWEBKE: Can you hear me a little better now?

CO-CHAIR CONWAY: A little better.

DR. SCHWEBKE: Well enough that you can hear me?

CO-CHAIR CONWAY: That's better

DR. SCHWEBKE: Okay. So, the two multiple sclerosis measures: (1) Both focused on individual -

CO-CHAIR CONWAY: I can't understand this. Kay -- Kay -- Kay, why don't you work on the phone on your side and we'll

move on to hearing from Janet and see if you can fix your phone problems. You continue to keep breaking up.

DR. SCHWEBKE: Okay.

CO-CHAIR CONWAY: Janet was the primary reviewer.

DR. NAGAMINE: So just a brief recap, again this is MS patients, adult MS patients on interferon and a serum ALT/AST in the last 12 months.

So do you want me to jump into importance or -- okay.

So in review of the TAP Committee's report that we have here, in terms of the impact gap and relation to outcomes, it looks like it was either minimally or partially that they voted. So, in the end they did vote that it met criteria.

Some of the comments that they made was that there may not be validity. It's based on consensus recommendations, so there's not strong evidence that doing this would

impact the outcome.

The compliance rate is 63.4 percent. There was a question from one of the TAP reviewers if the current recommendations call for monitoring every three to six months, why are they looking at yearly monitoring? And the differences, again, between RA and IBD in the incident or the intervals of measuring.

The other comment that they made is why AAFP would weigh in on this as opposed to the neurology specialists group, who manage MS.

So those were the TAP sort of reports.

And Bob and I are one and two reviewers on this, and we had a discussion and we had a discussion that is sort of similar to what we've been discussing this morning. And more specific to MS, you know back to that fundamental question of high volume, high risk. MS effects approximately 400,000 people in the U.S. Of the 400,000, approximately 30

percent, is my understanding, have relapsing remitting MS, which is the population that would qualify for interferon.

So the numbers here, that's about 120,000 people, of which five to 14 percent develop -- oh, I'm jumping to the white cell count. That's leukopenia. But for LFTs and liver enzymes I believe it's like 23 to 39 percent develop grade 1 transaminitis. So that's an LFT up to 2.5 times normal.

And for interferon you can prescribe it up to two times normal. So that's not a contraindication to start INF is your LFT is elevated two times above normal.

And grade 3, which is the really severe transaminitis is 1 to 2 percent of that 120,000 who would be on this drug.

So, those are sort of the numbers to give you some perspective of the people we're talking about.

Bob, please.

MR. BUNTING: Well, as she said, we

had the opportunity to meet yesterday and discuss this, so she covered it very succinctly. But just to emphasize, you're looking at 1,000 or 2,000 people if you really want to look at the grade 3, and you get back to that cost benefit analysis: How any resources do you want to develop to this for minimal gain? So, obviously this measure mirrors the previous measures that we've discussed.

CO-CHAIR CONWAY: Okay. We're open for questions, discussion. David, go ahead.

DR. NAU: Well, I guess then it sounds as though we're suggesting that because MS isn't very common, that it's not important to bother looking at this. I don't know if that's what you're implying, but I think if that is the case and the consensus view of the Committee here, that we don't bother to look at anything that's extremely common, then we might as well just not look at any of these measures for MS, IDD and so forth.

So, I guess that's why I'm trying to put a context, you know what we're trying to get to in terms of assessing importance. Because to me if your assessing, you know safe medication use with interferons, it seems as though you would be remiss not to be at least yearly monitoring liver function and so forth.

So, I guess once against that perspective issue, we've hammered here for hours. But I guess that's where I'm kind of lost because if we're suggesting rare diseases don't need safe monitoring and medications, then let's just go home now.

DR. NAGAMINE: Can I clarify that? That was one piece of context. But I also didn't get into the evidence piece. There's not strong evidence that monitoring the CBC yearly would effect mortality or outcomes. And I did speak to my rheumatology colleagues about this particular drugs, and their thoughts about that. And again, differentiating between clinical practice

guideline, patient variability and they all tell me, you know when I first start these patients we check it like every month, and then we go to every three months, and then we go to every six months.

And so if you say that Q12 is the standard and you're catching somebody who is in a different phase of the treatment, you know, you might ding somebody who is monitoring but perhaps less frequently. I don't know, but 12 months is certainly a fair interval.

DR. NAU: Well, and let me just respond to that, too.

Some of these issues seem to be the scientific validity of the measure in terms of what the interval --

DR. NAGAMINE: And that's what the test.

DR. NAU: Should be versus the importance.

DR. NAGAMINE: Right.

DR. NAU: And so I guess that's where we -- I think most of are kind of creating this gestalt of overall impression based upon multiple criteria that we're factoring into our impression of importance. And that's where it's tough for me to even keep those very separate. And so I guess if we're really just trying to figure out importance, you know once again I guess it's all a matter of perspective. But really I guess if we think that low utilization rates, low overall incidents of adverse events, I guess that's where we're trying to figure out how those factor into importance.

DR. NAGAMINE: The other point that the rheumatologists made was that the bad stuff that happens is acute and would not necessarily be prevented by outpatient monitoring on regular intervals.

CO-CHAIR CONWAY: How about going clockwise? Bob and Steve and David, and then Lisa.

MR. BUNTING: Thank you.

And to answer your question, I don't think that we want to dismiss any measure just because its population is small. That was not the crux of my comment. That was just part of it.

I think we have to look at the totality of the evidence. And if we knew that we did XYZ we could prevent the adverse outcome, I think we would probably vote to do XYZ. With this, I'm not sure of the benefit of it.

So, if you could prove to me or if anybody could prove to me that if you did this, you would prevent the acute event, then I think we would support that. I just don't think the evidence is there, regardless of the number of people effected by it.

CO-CHAIR CONWAY: Steve?

DR. LAWLESS: Yes. And also to clarify. It could be the rarest disease out there and I'd be fully supportive of it. It's

not the disease incidence, it's the model which we're working through. Because I look at the NQF as very important. That if we put this through as an potential model with its flaws, I think the credibility will be lost. And I think for other diseases -- so I'm looking for a medical safety management way of doing this that can be a model for other disease states or medicines to be used.

So, I think the importance is not -- I'm not looking at it as a disease. It could be RA, it could be some weird thing. It's the methodology and the evidence so then other people then would reproduce from it.

CO-CHAIR THIEMANN: I wanted to thank Dr. Nau for making that point and bringing up that distinction. Because it's a very important distinction. That just merely the sheer numbers of an individual suffering from a given disease doesn't necessarily indicate importance or not importance.

But I also wanted to thank Dr.

Nagamine for bring forward some of those statistics. Because in my opinion that was what in my opinion was missing from what is truly the incidence of this. And looking at what real people are we looking at potentially impacting here; not just the actual number of individuals diagnosed with MS.

So, from that perspective, and I think w have to take it step, by step, by step as NQF has laid out looking at each individual. Is it high impact? Looking at the opportunity for improvement, and then looking at is there evidentiary support for the outcomes linking those. And so I think that goes back to what Dr. Nau was talking about, looking at it based on that element versus just the disease issue.

CO-CHAIR CONWAY: Go ahead.

MS. THRAEN: I was going to look it up, but just for clarity's sake is this under the medications coming from the medication safety group, this indicator?

CO-CHAIR THIEMANN: Measure?

CO-CHAIR CONWAY: It's from Ingenix.

MS. THRAEN: Yes. But the TAP, was the TAP the medication safety group? So this is a medication safety question?

CO-CHAIR CONWAY: Yes. Yes. Yes.

MS. THRAEN: I just needed to clarify that.

CO-CHAIR CONWAY: Other questions? Yes?

DR. KENNERLY: One thing I think in hearing the response I think to a question that I've asked the Ingenix folks earlier was if you begin to start looking at databases, claims databases largely from payer groups and you begin to get issues associated with migration of patients in and out of those databases, I wonder if there's some caution here also around a 12 month interval when in fact you would have to have somebody in fairly substantial continuous enrollment to be

certain that you did not have it done just before you enrolled, or perhaps just after you left in terms of looking at these kinds of measures.

CO-CHAIR CONWAY: Are there any Committee members on the phone that has comments or questions?

DR. SCHWEBKE: In response to that, so you can hear me better, I switched phones.

CO-CHAIR CONWAY: Is that Dr. Schwebke?

DR. SCHWEBKE: Yes.

In response to that, that's actually why we require eligibility over the entire 1 month report period. And do also give credit if there's three months of additional data that comes in after the end of the report period.

CO-CHAIR CONWAY: Do you have any overview of comments on both of these measures now that you've got a well working phone?

DR. SCHWEBKE: Well, you know I

think that the struggles that you're going to address with these are very similar to the other measures that we've discussed this morning. You know, these measures are based on expert opinion. And when these medications go through the FDA process, all these individuals are monitored.

And then the only thing I would add is that the one to two percent grade 3 level adverse event, which is an ALT greater than 5 or higher, I can actually upgrade that information and the manufacturer has actually now published that up to ten percent of individuals on interferon for multiple sclerosis have grade 3 events. But I appreciate the challenge, and that is linking, you know does monitoring make a difference? You know, the challenge of course is we're probably never going to have studies that are going to really look at that. You know, I think that measures like this are going to always be based on expert opinion. So it's

going to challenge of us deciding, you know is the relative value high enough given the absence of any multiple sclerosis measures at this point that are NQF endorsed that would warrant endorsing a measure of this nature.

CO-CHAIR CONWAY: Okay. Thank you.

David?

DR. TURNER: I guess I'd like to actually address the question to Ingenix. I'm just thinking about compliance rate and then trying to reference that relative to the indication for the drug in MS. And if I understood Dr. Nagamine's comments about that this would be in the relapsing percentage maybe 120,000 patients that would actually have an indication for this drug, then was the compliance actually assessed amongst that group? And I guess the follow-up question to that is are they coded differently within the claims data so that one that would be trying to identify compliance within this population would actually be able to assess that?

DR. SCHWEBKE: I am looking at that question right now. My recollection is there is only one ICD-9 code for multiple sclerosis, and we should have that. Yes, that's right, 340.

So the current ICD-9 coding system does not distinguish between the different types of multiple sclerosis. I honestly don't recall with the ICD if we're going to see that granularity. And so we don't know what the specific sub-type of multiple sclerosis is. All we can say is that we've identified them as having multiple sclerosis and they've been taking the interferon recently for a duration greater than three months.

CO-CHAIR CONWAY: Any other questions or discussion?

Should we move on to grading the importance of the measure? There's no heads either nodding or disagreement, so I guess we'll move on.

This isn't a whole lot then

telephone conference call. I know, I know, it's not easy.

All right. Let's take a look at the straw vote on the impact of the weight of evidence demonstrating impact of this measure.

So all those grading that as completely demonstrated, please raise your hand? Okay. There are none.

Partially demonstrated? Looks like there's three.

Minimally demonstrated? There's eleven -- 12.

And do we have anyone on the phone?

DR. SOLOMON: Yes, we do. Partial.

CO-CHAIR CONWAY: Partial. Okay. I think that's the whole group.

How about the weight of evidence on demonstrating a gap? Anyone in favor of that being completely demonstrated? There's none.

Partially? Five. Oops, six. Could we try that again. Raise your hands high. Six. Six partial.

Minimally demonstrated? Nine.

And any not at all?

Dr. Solomon?

DR. SOLOMON: Partial.

CO-CHAIR CONWAY: Okay. And the weight of evidence relating this measure to the outcome of the condition. Those feeling that it's completely demonstrated, raise your hands. Okay.

Partially demonstrated? There are none.

Minimally demonstrated? Thirteen.

Not at all demonstrated? Two.

And Dr. Solomon?

DR. SOLOMON: Minimal.

CO-CHAIR CONWAY: Okay. Now in the overall importance to measure and report this measure, we'll be voting yes or no. How many of those thing this should receive a yes vote, please raise your hand? Two. Okay.

How about no vote? Thirteen.

And Dr. Solomon?

DR. SOLOMON: No.

CO-CHAIR CONWAY: No. Okay.

All right. Well that's another measure completed.

Should we have lunch while we work through the next measure? Okay. Then we'll move right along. That's fine. It will probably be similar.

Janet, I think you were the primary reviewer again.

DR. NAGAMINE: So this is PSM-022-10 dealing with adult patients with MS taking interferon that had a CBC in the last 12 reported months.

And review of the TAP Committee's votes, the impact, there were two that said minimally two that said partially.

There wasn't a lot o comments on this one in terms of the gap. The compliance rate for this one was 58.2 percent in relation to outcomes, most of them said partially. So there weren't a lot of comments.

I think Bob and I's discussion on this one was pretty much mirrored with the previous discussion.

MR. BUNTING: And this one probably even more so because the frequency is less defined than the one we just voted on.

CO-CHAIR CONWAY: Okay.

DR. SCHWEBKE: Actually one of the articles provided indicates a prevalence of leukopenia that is five to 14 percent. So that probably is actually a little bit higher than I think what we saw with the transaminitis.

CO-CHAIR CONWAY: Okay. Questions or comments? Okay.

Well then, let's move on to grade the importance of this measure.

Regarding the impact of the measure, those who feel it's completely demonstrated please raise your hand.

Oh, sorry. Please.

CO-CHAIR THIEMANN: A real quick

question for NQF staff. Because on 22 the Technical Advisory Panel made a request of the performance measure developer to change the time frame from 12 months to six months. And it's my understanding that Ingenix agreed to that change stating that there was evidence to support decreasing the frequency from 12 to six.

So, if we're voting on the importance to measure, are we voting on the six month or as specified originally in the original application of 12 months?

MS. BOSSLEY: It was changed.

CO-CHAIR THIEMANN: It was changed?

MS. BOSSLEY: So we should, and we will correct and have Kay go back in and update this to be six months.

CO-CHAIR THIEMANN: Okay.

MS. BOSSLEY: I'm sorry. You're right. Evaluating this based on six months as opposed to 12.

CO-CHAIR THIEMANN: Okay. Just in

case that influences an individual's assessment of the evidentiary support for outcome since the frequency of monitoring influences.

CO-CHAIR CONWAY: Okay. Looking back at the demonstration of impact of this measure, those that feel that it was partially demonstrated, please raise your hand. There's two.

Those that feel this is minimally demonstrated? Looks like 13.

And Dr. Solomon? Dr. Solomon, would you like to vote?

DR. SOLOMON: Minimal.

CO-CHAIR CONWAY: Minimal? Okay. I think that's everyone.

Taking a look at whether a gap has been demonstrated for this measure, those who feel that that was completely demonstrated please raise your hand.

Those that feel it was partially demonstrated please raise your hand? There's

five.

Those that feel it was minimally demonstrated please raise your hand? Ten.

And Dr. Solomon?

DR. SOLOMON: Partial.

CO-CHAIR CONWAY: Partial? Okay. I think that's everyone.

Okay. As far as the relationship to outcomes on how well that was demonstrated, those that feel it was completely demonstrated please raise your hand. None.

Partially demonstrated, please raise your hand. None.

Minimally demonstrated, please raise your hand. Twelve.

And not at all demonstrated? Three.

Dr. Solomon?

DR. SOLOMON: Minimal.

CO-CHAIR CONWAY: Minimal. Okay.

We'll move on to overall voting in this category, yes or no on the importance to

measure and report on this measure. Those voting yes on that, please raise your hand. There are two.

And those voting no, please raise your hand. There's 13 noes.

Dr. Solomon?

DR. SOLOMON: No.

CO-CHAIR CONWAY: And one more no.

Okay. I think we are -- oh, do we have to have public comment on it.

MS. BOSSLEY: Yes.

CO-CHAIR CONWAY: Okay. We're open for public comment. Okay. Hearing none, I think we're ready to break for lunch. Okay. Thank you.

(Whereupon, at 12:20 p.m. the meeting went off the record and resumed at 1:04 p.m.)

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

1:04 p.m.

CO-CHAIR CONWAY: We'll be looking at two measures around monitoring treatment of hepatitis C. We're beginning with patient safety measure: PSM-023-10: Patients with hepatitis C infection taking interferon that had a periodic serum ALT monitoring. And to open up this section, we could see if the measure developer has any opening comments for both of these measures.

Dr. Schwebke, are you on the phone?

DR. SCHWEBKE: Yes, I am.

CO-CHAIR CONWAY: Okay. The only thing I want to point out with these two measures is the logic is slightly different. It's similar in that it's still identifying a specific population here, individuals with hepatitis C who are taking an interferon containing medication keeping in mind that hepatitis C treatment is combination therapy with interferon and another medication called

ribavirin. But everybody is going to be on either both of these medications. Occasionally we'll treat some of these people with just interferon.

So, we're identifying individuals who are treatment.

What's unique here about the measure is that AASLD guidelines have actually been very clear with monitoring recommendations. And in fact, the 2009 AASLD guidelines have been approved not only by that organization, but also the American College of Gastroenterology and the Infectious Disease Society of America. And in those recommendations they recommend specifically a serum ALT monthly at minimum along with a CDC monthly at minimum for at least the first 12 weeks. And then there's some flexibility in that subsequent monitoring every eight to 12 weeks.

Since we can't be confident with administrative claims data where exactly an

individual may be in their treatment regiment, what's different with this measure is rather than looking for one cast within a specific period of time for compliance, we're actually looking for two or more tests for one measure the serum ALT and for the other measure the CBC. So two or more tests at least 14 days apart during the last 180 days of report period. And then we include 90 days after the end of the report period if additional claims are available.

And then allows to be sure that individuals are at least going with the more conservative time frame of monitoring at least every 8 to 12 weeks during that six months time frame. So, that's the one unique thing about the monitoring here compared to the earlier measures that we've discussed today.

The compliance for the ALT monitoring measure is 65.8 percent. And the monitoring for the kind of companion measure, the CBC was very similar, 68 percent.

CO-CHAIR CONWAY: Okay. Thank you.

Our primary discussion leader was David Nau

DR. NAU: Sure. This measure, once again, addresses the monitoring of patients taking interferons. As pointed out, most of these patients are going to be getting pegylated interferons along with ribavirin.

The clinical guidelines from AASLD do indicate that monitoring should be happening every eight to 12 weeks for patients taking these drugs. So the measure is consistent with the guidelines. And there is evidence that compliance with this parameter of the guidelines is not perfect, it's around 66 percent.

So that's essentially the key points, I guess.

CO-CHAIR CONWAY: Okay. Thanks.

And Steve, do you want to add anything to that?

DR. LAWLESS: Yes. The only couple

of things I would note is that this has the word "periodic" in it. So, obviously, it's wavering a little bit, and I agree for me as a guide or just something saying "periodic" left a little bit up to -- as a title. I mean, it's left a little bit up on the air until I know what the intent seemed to be.

I looked at the incident of this and they describe it as 1 percent of the population who are on this would get an elevation of the liver enzymes. So it's a relatively small incidence among those who are even on this that would have the rise. And I'd have to get an interpretation from a GI specialist that if I had hepatitis C and one percent of the patients had a rise in liver function tests, the issue I would have is that it the liver -- is the hepatitis C or is it the drug. And I don't know how you'd be able to distinguish this way.

So, I don't know if it's safety versus -- is the drug monitoring versus

disease state. So I was a little bit confused about how to handle that or how that interpretation would occur with this.

And I thought also -- and this is just when we talk about the other measure also -- and this may have been an oversight or not, but the denominator calculation in this one was different from the denominator calculation in the other measure, the way it's just outlined. It could be that they just wanted to shorten it in terms of specificity or it's just an oversight. But it looked like they were different, and I'm just curious why that difference is in the denominator.

DR. SCHWEBKE: If there is some difference that you're noticing, it would be helpful to know maybe which specific -- if it's the denominator time window, the denominator DTL, the denominator --

DR. LAWLESS: I think if you go to the -- yes. I'm sorry.

DR. SCHWEBKE: So that would be an

error if that's there.

DR. LAWLESS: I don't understand there may be a clarification on the TAP report I got on pages 6 through 9. Give a lot of outlines of the codes used, the clarifications, and that was not the same in the 23. So I just -- it could have just been -- the assumption was it was the same. I just didn't know if that as just an oversight.

DR. SCHWEBKE: You are correct that they should be exactly the same. We are identifying the same population here.

DR. NAU: And just to clarify, the title of the measure is more vague than the actual specifications because the measure description does indicate two serum tests in the past six months. So, that's where the -- you know, I guess I pay attention more to the description because the title overall just doesn't give you the detail there.

DR. SCHWEBKE: That comment is well taken. Periodic was used for brevity. We

have been chastised in the past with NQF in endorsed measures for having unnecessarily long measure descriptions. So, that's certainly something that we could modify and give that detail.

CO-CHAIR CONWAY: Okay. Any questions from the Committee members? Yes.

MR. BUNTING: In the report that we have it talks about the error rate, and there are different numbers tossed about; 11 percent, 2 percent, 17 percent and then a 5 percent error rate overall. Can you address the confusion caused by that paragraph? And that's addressed to Ingenix.

DR. SCHWEBKE: I think what's you're referring to is 2c.2 the Analytic Method?

MR. BUNTING: That's correct.

DR. SCHWEBKE: Is that correct, is that the section that you're looking at?

MR. BUNTING: Yes, that's correct.

DR. SCHWEBKE: Yes. So this is a

section, and the specific portion that you're referring to is just an example, and actually we talked about this earlier, where we went in and we did attempt to validate using a chart review comparison process. The results based on our administrative claims output when looking at measures versus what we were finding from a chart review.

Now, this specific measure wasn't included. It was more of a looking at where are administrative claims strong in identifying gaps in care and where might there be problems with data incompleteness where administrative data just isn't capturing the information.

So, what you're looking at is that 100 member chart review where we looked at 126 measures. And I think probably the most important thing from this was that when we look at -- it's the second bullet point, the error rate for measures that required labs for numerator compliance. That was 4 percent, and

actually the error tended to be on the side of -- the chart review was typically missing labs that had been done at outside facilities.

So, in other words, administrative claims when you're looking at labs is actually quite robust and in fact, probably is even better than going to the paper chart because you miss tasks that are often done at outside facilities.

DR. LAWLESS: Along with the error rate mentioned, is the error rate different from the numerator perspective versus the denominator perspective?

DR. SCHWEBKE: Here what we mean by error rate was that there were 14 situations out of 318 where there was not an identical match between what administrative claims told us about a lab being done and what the chart told us about a lab being done. And in those 14 cases where there was a lack of concordance, the problem actually tended to be that the chart was missing a laboratory test

that the administrative claims was able to detect.

CO-CHAIR CONWAY: Okay. Any other questions or discussion? Steve?

DR. MUETHING: Just to make sure I understand. From my reading on this summary is that they had one percent of the patients with hepatitis C had marked elevation of their ALT?

DR. LAWLESS: Yes. Let me get the exact quote that was in here. I think it was in the TAP report, page 3. Per the pharmaceutical manufacturer one percent of patients in the hepatitis C trials experienced marked elevations in ALT during treatment.

DR. MUETHING: And again, similar to the other ones, then we don't have any evidence that it was the screening that picked that up?

DR. LAWLESS: Right.

DR. MUETHING: We don't know if it was screening or some clinical change that

brought that to the -- during the trials?

DR. LAWLESS: Correct.

DR. MUETHING: Okay.

DR. LAWLESS: Or at least it was just not mentioned here.

DR. MUETHING: Exactly. Yes. Yes. Okay. Thank you.

CO-CHAIR CONWAY: Okay. Should we move on to grading the importance of the measure? Seeing no negative head nods, we'll do that.

We'll start with the evidence for the impact of this measure. Was that impact demonstrated completely? See any hands that feel it was complete.

Okay. Was the evidence partial for demonstrating that? Ten hands. Okay.

And minimal evidence for demonstrating that? Five hands.

Dr. Solomon, are you still on the phone?

OPERATOR: His line has

disconnected.

CO-CHAIR CONWAY: Okay. Thank you.

So that complete the whole group.

Evidence of a gap in performance in this measure, those feel that was demonstrated completely please raise your hand. There's none.

Those feel that was demonstrated partially, please raise your hand? Fifteen. That would be everybody.

We could move on to the relationship of this measure, the outcomes. Those that feel that was demonstrated completely, please show your hands. There are none of those.

Those that feel that it's partially demonstrated, please raise your hands?

And those who feel it was minimally demonstrated? Fourteen I think.

Is there someone who feels that it was not demonstrated at all? Or is there an abstention, or I counted wrong. It was

probably 15 voting in favor of that being minimally demonstrated.

Then we can grade the overall category. The importance of measuring and reporting on this measure. Those that feel the answer to that is yes, please raise your hand. That's one yes.

And those that feel that it's not -- the answer to that is no, please raise your hand. Those are fourteen noes, one yes.

Okay. And we have a question.

MS. THRAEN: I'm going to go back to the idea that all of these were aimed at medication safety approaches and ask the clinical people here in the room if they could give, from my understanding, give me an example of what would be a better way of measuring clinical safety other than these types that have been proposed so far in these. So, if we just use the last one as an example. Knowing that they're using claim data, is there a way of getting at the

question or a better way of getting at the safety of this particular drug that would solve a problem or identify or prevent adverse events occurring, medication safety adverse events?

I'm struggling with the fact that we're rejecting all the medication safety, proposed medication safety measures. And I'm just trying to understand better is it because these are claims data and we're just simply looking at timing of labs? Would it be better suited to have a clinical piece of information that was included with the medication that would then point to this risk for adverse events?

CO-CHAIR CONWAY: We may all have a different answer to that question. It doesn't have to do with how its being measured. From my own point of view, I think right now we don't have enough information to know about what the right time intervals are, or even if we had some agreement on what the right time

intervals were, whether drawing a lot of blood tests would have any impact on preventing a complication.

So, in my own mind I think this whole category is just not ready for prime time. We need a little bit more research. That's just my own perspective.

Others may want to answer that. It's a good question.

DR. NAU: Well, and that's where I'm trying to sort out whether this is an issue of importance or whether it's scientific acceptability of the way its specified. That's where I've been consistently sort of supporting the importance of these measures. Because I think it is important that we determine whether we're following the recommendations for how to safely monitor patients on these medications.

I don't always agree with some of the way the measures are specified, but I do think it's important to determine whether or

not we are monitoring patients on these very toxic drugs. And so there's where I've tended to fall on the side of saying it's important, but wanting to then have some discussion around whether it is the right interval or whether a different interval would be better.

So I think that's where maybe I'm sort of differing from the rest of the group in my vote around the importance issue.

DR. NAGAMINE: I struggled with that question, and I went back to the medication safety category to see what has already been done in that regard. And that helped a little bit. And when I look at these, they're broader.

Do we have a medication list in the outpatient record? Do we have documentation of allergies? Therapeutic monitoring for persistent meds? There are certain cardiac meds that are also part of the core measures in here. Drugs to be avoided in elderly. And then fall risk management.

And so those were broader and they're there. And they don't focus on one disease or one drug, but I think the impact of these are far reaching. They're really important ones.

And so when I look at if we're trying to improve medication safety, what has the biggest bang for the buck? That's sort of how I looked at it. And so if you want to ask about a specific drug, you know you could look at Warfarin, which has 31 million prescriptions a year and people die from bleeding on Warfarin. And so again, high volume, high risk was what I kept on coming back to as I struggled with these questions.

DR. LAWLESS: I made a distinguishing case in safety versus surveillance of adverse drug reactions. And the way these were tallied for me, these look more like a surveillance of adverse drug reactions and determining more of the incidence and what people reacting do with

them. So it's medication safety, but it's more quality I try to determine it, rather than a safety issue per se.

MS. THRAEN: I hate to admit this, but that's how it feels to me.

DR. MUETHING: Granted that the type of measures I'll state as an example would be much tougher to get to, I would be more intrigued by a measure of physician response to a patient with an elevated ALT. And percentage of times we fail to respond or don't make a change, or following up on Janet's comment, I would be much more intrigued by a measure of percentage time we put a patient on one of these drugs when they have a drug in their medication list that's contraindicated to use this medication.

My impression on this type of measurement is we can measure this and we can establish that it's not bad to measure this, but to generalize after a statement is that therefore you should -- all patients on this

drug should be measured in this way feels a strong statement. And that's my struggle.

CO-CHAIR CONWAY: Okay. Shall we move on to the next measure? That is PSM-024-10. This is the periodic CBC monitoring of patients on interferon with hepatitis C.

Our primary discussion leader is David Nau.

DR. NAU: And this measure is once again consistent with some of the others regarding interferon. So, I think once again it's about the ongoing monitoring of CBC in patients who are taking interferon and how have hepatitis C. And so I think I'll just leave it at that at this point.

CO-CHAIR CONWAY: Okay. Steve, you got anything to add?

DR. LAWLESS: No, nothing else.

CO-CHAIR CONWAY: Okay. Shall we move on to -- yes?

MS. THRAEN: Actually it goes back to the one before this, the age limit, the

three years of age question. Does this differ if it were under the age of three, do pediatric cases have any influence on whether or not this is good or bad or needed?

DR. LAWLESS: I didn't bring it up, and I should have actually. Is that since you're dealing with pediatrics here, and I didn't see much evidence from the pediatric world of this. So in terms of evidence brought in, consensus from any of the pediatric groups or comments, so I don't know. And I think it goes along with the evidence of putting out something that's a measure of a pediatric population, but there's no input for knowledge or anything about that. It would mean that if we were close to agreeing with it, I would probably ask for a re-vote in saying over 31.

DR. SCHWEBKE: Actually, I'll also disclose that half of my practice is treating people with hepatitis C. These drugs are actually not approved in the pediatric

population.

Now, with that being said, are there some people who treat pediatric patients? Yes. And in fact, that's one of the difficult points that we discussed with AGA when we reviewed these measures. And it was the sense that because we acknowledged these medications are not FDA approved in that population, but we also know that practitioners, once the drugs are available, will sometimes use them when they believe it's appropriate. We thought it was important to include the pediatric population.

We certainly would have no objection to changing that threshold, however

DR. LAWLESS: Let me clarify. I think your answer was a very good one. So the idea would be, this would be if you're prescribing the medicine? This does not address whether it's appropriate or it's an off-label use, or whatever else. So, if you are using it, it's not, if you're a pediatric

patient don't follow this measure?

DR. SCHWEBKE: Correct.

CO-CHAIR CONWAY: Yes.

MS. THRAEN: One more question that I don't know the answer to related to public health. So, is there any impact felt at a communicable disease level for patients receiving these kinds of treatments that's related to their lab work? I mean, I'm just asking because I don't know. That are they more susceptible for passing the disease on, if they're not -- or et cetera, that question, the public health perspective?

DR. SCHWEBKE: That's a really interesting question. We don't know the answer to that as well as we do in the HIV arena. I mean, one would assume that somebody who has a lower viral level, which is going to happen when people are placed on hepatitis C treatment, one would think that they're less likely to transmit.

We don't actually have that data in

hepatitis C population. We're actually obtaining that information in the HIV population. So if we want to assume the viruses act the same, one might say yes, that's probably the case.

You know, I think that the bigger thing -- and I will say this: of the two measures, the CBC measure is actually, I think, the strongest. Because we do see some dramatic changes in the hemoglobin. And it's not only a huge safety issue, but it's not predictable. And I've had people on treatment for six months as they're heading for their 12-month treatment course where their hemoglobin has been stable, and then it drops. So it's not only a safety issue that can sometimes be unpredictable, but also it really does feed into adherence. When people want to stop it's because they don't feel well. Sometimes they don't feel well because they are severely anemic. And if we aren't monitoring for things like that, we miss both

adherence opportunities as well as safety opportunities.

MS. THRAEN: But in the proposed measure the use is administrative data. So you're administrative -- the fact that you got a lab done, a CBC and a lab done, would not tell you what those levels were, correct? It's only the clinical data that would give you that information, correct?

DR. SCHWEBKE: Correct. Correct. It would also be difficult -- getting to kind of an earlier comment, it would also be difficult with administrative claims data to know what then occurred. Oh, we could identify a blood transfusion, but we probably couldn't identify -- in fact we couldn't identify a dose reduction. Because what we're going to do in that situation is we're going to call the patient and say drop your ribavirin dose or drop your pegylated interferon dose, and we're not going to be able to detect that for some time with

administrative claims.

MR. BUNTING: Does Ingenix have the ability to look at lab values, or just the fact that the lab was performed?

DR. SCHWEBKE: That's a great question. That's where these tests called LOINC codes come in. LOINC codes are a standardized data source. Many people aren't familiar with LOINC codes. But for example, with a hemoglobin A1C, there's a CPT code that identifies that test was done. It turns out there are unique LOINC code results that actually give you the hemoglobin A1C value. Similarly, we can sometimes see that with LDL, HDL, et cetera.

So LOINC codes are that specific unit that can give you test results. It's a little bit, I think, challenging with certain labs like hemoglobins and transaminases, the LOINC codes have actually been more granular for things like LDL, hemoglobin A1C, GFR, things like that. So the data is becoming

available, but it's pretty limited at this point.

DR. SOLOMON: When you say "limited," do you mean it's limited because certain labs aren't providing the data back to Ingenix or other insurers, or is it the vagaries of how it's coded, or both?

DR. SCHWEBKE: Yes, it's both. I'd add kind of the additional dimension that you don't always have the glandular lab result for all diagnostic tests. For example, it's difficult to have a LOINC code that tells, that actually quantitates progeria.

CO-CHAIR CONWAY: Any other questions or discussion? Shall we move on to grading the importance of this measure to measure and report? We'll go through each of the three sections.

Do you feel the evidence of impact of this measure was demonstrated completely? All that think that raise their hand.

Okay. Was it demonstrated

partially? That would be nine.

Was it demonstrated minimally? And six.

And Dr. Solomon, are you on the phone?

DR. SOLOMON: Yes. Partially.

CO-CHAIR CONWAY: Partially. Okay.

DR. SOLOMON: Not partially on the phone.

(Laughter.)

CO-CHAIR CONWAY: Yes. Thank you.

How about the degree of demonstration of a gap in compliance, those who felt it was demonstrated completely please raise your hand. Okay. There are none of those.

Demonstrated partially, please raise your hand. Eight.

And those that feel it was minimally demonstrated? And that's five.

We have one more coming late for

partial.

Dr. Solomon?

DR. SOLOMON: Partial.

CO-CHAIR CONWAY: Partial. Okay. I think that's everyone.

The third question in the set is do we have a relationship to outcomes, and those that feel that was completely demonstrated, please raise your hand. There were none of those.

Those who feel it was partially demonstrated, please raise your hand. None of those also.

Those who feel it was minimally demonstrated, please raise your hand. It looks like 13.

Those who feel it was not demonstrated at all? Two.

And Dr. Solomon?

DR. SOLOMON: Minimally.

CO-CHAIR CONWAY: Minimally. Okay. Thank you.

And the overall grading of this section, the importance of this measure to measure and report it, we'll answer yes or no. Those who feel that the answer to that is yes, please raise your hand. We have two. And those who feel the answer to that is no, please raise your hand. Thirteen.

And Dr. Solomon?

DR. SOLOMON: No.

CO-CHAIR CONWAY: No. Fourteen.

Okay. Thank you.

And since you're all doing so well, working so hard, we could do a stretch break now. That's what the agenda shows. Maybe ten minutes.

I'm sorry. Are there any public comments or members on the phone line?

Okay. Hearing none, we'll take a ten-minute break.

(Whereupon, the above-entitled matter went off the record at 1:37 p.m., and resumed at 1:59 p.m.)

CO-CHAIR THIEMANN: All right. I think we're going to reconvene and see where we can go next.

We are moving on to the last two performance measures for today that are on the agenda for today. And I believe PSM-030-10. And I believe, Dr. Nau, have you agreed to step in as primary, or Dr. Knight, or tag team?

DR. NAU: Sure, we can tag off here.

CO-CHAIR THIEMANN: Okay. Sounds good.

DR. NAU: Well, measure 30 looks at patients with inflammatory bowel disease taking one of four immunomodulatory drugs that had a CBC in the last three reported months. Very similar to the measure we evaluated for patients with rheumatoid arthritis. The only other difference, really, is that with the rheumatoid arthritis patients, the ACR had made explicit recommendations regarding

monitoring of patients on these drugs. But as far as I can tell, none of the gastroenterology societies have made explicit recommendations for monitoring patients on these drugs with IBD.

Dr. Knight, did you want to add anything?

DR. KNIGHT: I don't have much to add to that, other than, on this one there was 42 percent compliance when they looked at that. And again, recommendations were based on expert opinion and it was noted that no rigorous research in appropriate screening intervals had been done.

CO-CHAIR THIEMANN: Dr. Schwebke, would you like to add anything from Ingenix's perspective on either 30 and 31 as a summary, initially?

DR. SCHWEBKE: Yes. The only thing I will add is that these were an additional two measures that were reviewed by the AGA Subcommittee who, despite the fact that they

acknowledged that they don't have any published guidelines that look at monitoring medications, they were extremely supportive of these measures.

CO-CHAIR THIEMANN: Thank you.

And any comments, questions from Steering Committee members? I'm not seeing any.

So I'm going to take that as indication that we probably should move forward to consider whether or not the group feels this measure passes threshold for importance to measure and report at this time if there are no questions or comments. Okay.

With that in mind, does the group feel that the measure developer completely met the burden for demonstrating high impact in healthcare for this performance measure; completely? I see zero.

Partially? Six. Sorry. Didn't see your hand, Dr. Nau. Just wanted to make sure.

Minimally? And this ought to be eight.

Dr. Solomon, are you still on the line?

DR. SOLOMON: Minimally.

CO-CHAIR THIEMANN: Minimally. Okay.

Any not at alls? One. Okay.

For section B, the measure developer demonstrated opportunity for improvement on this issue? Completely? I see zero.

Partially? Nine, I think, is what I have. Nine.

Minimally? Six.

And not at all? Zero.

Dr. Solomon?

DR. SOLOMON: Minimal.

CO-CHAIR THIEMANN: Minimally. Okay.

And has the performance measure demonstrated evidence for outcome? Completely?

Partially? Two.

Minimally? Eleven. Eleven.

Not at all? Two.

And Dr. Solomon?

DR. SOLOMON: Minimal.

CO-CHAIR THIEMANN: Okay. And so now we're going to do a summary vote then on whether the threshold for importance has been met by the measure developer. Does the group feel that yes, they have? Two. No? Is that everybody else? Twelve.

And abstaining? Zero.

And Dr. Solomon?

DR. SOLOMON: No.

CO-CHAIR THIEMANN: Thank you. Okay.

Moving on to PSM-031-10. Dr. Nau, I believe, you were primary discussion -- well, you were secondary discussion leader, I believe, but Dr. Kowdley is not here, correct?

DR. NAU: Sure. I'll start us off on this one.

CO-CHAIR THIEMANN: If you wouldn't mind.

DR. NAU: Okay. This measure looks at patients with IBD taking methotrexate that had a serum creatinine in the last six months.

CO-CHAIR THIEMANN: And any questions or comments from Steering Committee members on this issue?

Okay. I don't know if everyone's post-lunch, or --

So I'm seeing any comments, any hands. So I think we're moving forward to calling for votes again on 1a, 1b, 1c and ultimately for 31.

Has the measure developer demonstrated high impact for this performance measure completely? Any votes? I see none.

Partially? Six.

Minimally? Nine.

Not at all? Zero.

Dr. Solomon?

DR. SOLOMON: Minimal.

CO-CHAIR THIEMANN: Minimal. Thank you.

And demonstration of opportunity for improvement on this issue, completely? Zero.

Partially? Two.

Minimally? Eleven.

Not at all? One.

Dr. Solomon?

DR. SOLOMON: Minimal.

CO-CHAIR THIEMANN: Minimal? Okay.

And concerning evidence for outcome, performance measure completely demonstrated? Zero.

Partially? Zero.

Minimally? I see fourteen. Fourteen.

Not at all? Two.

And Dr. Solomon?

DR. SOLOMON: Minimal.

CO-CHAIR THIEMANN: Thank you.

And so taking the vote for whether the performance measure has met the threshold for importance to measure and report on this? Answer yes? Any? Zero. Okay. No? That's 15, yes.

And abstaining?

Oh, Dr. Solomon, I apologize?

DR. SOLOMON: Okay. No.

CO-CHAIR THIEMANN: I was good up until that point.

So with those two, that actually -- any NQF members, other members or public comments on these two measures for anyone else on the line? Hearing none.

Then when we are in a phase where we could start to, earlier we had talked about next steps, thinking about possibly a recap and what type of comments NQF staff may need to consider in drafting its report as to the actions and discussions here from the Steering Committee. And I know I've had some sidebar conversations, I'm sure, Dr. Conway, you have

as well about some of the difficulties and challenges in evaluating these measures and the need to still express that this is not -- that just because the Steering Committee voted that the measure developer at this point in time, it was not our opinion that the importance to measure was reflected or captured. And so maybe opening up the floor now to the Steering Committee members' opinions and thoughts about maybe important next steps or recommendations to the measure developer on this issue in drafting these measures, or some directions we'd like to see measurement in this population go.

DR. SCHWEBKE: Before you do that, can I ask a quick question about the two HIV measures?

CO-CHAIR THIEMANN: Sure.

DR. SCHWEBKE: Are those going to be discussed today, or are those still not fully reviewed?

MS. MUNTHALI: Hi, Kay. This is

Elisa.

They will be reviewed once the Technical Advisory Panel reviews them.

DR. SCHWEBKE: Okay.

MS. MUNTHALI: And then the Steering Committee will evaluate those. I think it's on November 19th.

DR. SCHWEBKE: Okay. Very good. Thank you.

MS. MUNTHALI: You're welcome.

CO-CHAIR THIEMANN: So I'd like to open it up. Ms. Thraen?

MS. THRAEN: One of the concerns, again, that I have is how these measures are intended to be used. If all of the science was here to support these measures today, and we could agree on that, then what's not being judged and again Peter said something about this being an ethical question. I'm not sure it's as much of an ethical question as it is more of a practical question in terms of what's the intent of the use of this.

So, in a large group practice when you're trying to standardize variations in practice, having feedback on the timing of a particular lab work in relationship to medications that you're prescribing, et cetera, to me that seems like a reasonable use of many of these measures using the administrative data.

In terms of using them for public accountability from a state government perspective, it doesn't seem reasonable to me.

Using them in terms of trying to tease out more clinical values in terms of practitioner response to alerts or lack of alerts, or values that are too high; that, to me, is more of a patient safety risk opportunity.

So, there's no way currently in the process for us to say, to judge, to make using explicit criteria, to make a judgment on how this measure could be used, should be used, ought to be used regardless of the science as

part of the overall consideration of its importance.

So, figuring out how to do that, I don't have any solutions to offer so I tend not to want to criticize. But figuring out how we can do that, how can we make those judgments and make those recommendations.

So maybe you used the term "not ready for prime time," I used earlier a "league approach" or a "farm team" versus the professional team approach. Rather than just a yes or a no. That there be a way of categorizing some of these measures in such a way that there's a consensus, there's a voting consensus that this would be a good quality improvement measure. Because I think that the society, our systems are looking towards an organization for a clearing house of measures. And under some circumstances under, you know if I were a specialist practitioner with my rheumatoid patients looking to NQF to say here's a measure that's gone through a vetting

process that if I were to measure something in this particular population, this would be a good way of measuring it. And I know that there's been some work on this. I would take that and move forward, and move to improve the care. That's not the same as saying that this is a national measure that everybody needs to be using.

So figuring out how to capture that spectrum, the word that was used earlier, that spectrum is something I'd like to see developed.

CO-CHAIR THIEMANN: Dr. Levine -- or Mr. Levine -- okay. I saw you move forward, so I wasn't sure.

Any other comments, thoughts. Dr. Lawless?

DR. LAWLESS: I think, actually, hopefully, I like your idea about the feedback going to them in terms of that people don't get the impression that everything was kind of wiped out here because we just don't want to

do it. I mean, it was really out of the consideration for the long term what this meant.

That said, these were more adverse drug reactions and how well you're monitoring for adverse drug reactions, which is part of medication safety but it's a sub -- it's not what's on a lot of people's top of mind when they hear medication safety.

So, I think that maybe having categories of medication errors versus medication monitoring, versus error prevention, may be a good way of outlining when you put a call for proposals out, so that there's a clarity there for them.

And I would also like to go to the point about disease. If a measure was a little bit more -- something that could be applicable across all disease states, so medication management safety and creating a template for that for people to respond to of how they would fit into a medication

management safety. This way you can insert disease and then say this is what you should be doing to prevent those errors, and you wouldn't have to worry about incidence of disease.

CO-CHAIR THIEMANN: Go ahead, Dr. Nau, and then I'll go to Mr. Levine.

DR. NAU: And I believe what the consensus of the group was that we should be evaluating importance based on sort of are these very broadly usable, applicable, important to add to some national reporting efforts. And from that standpoint, then, I think the consensus was correct in that these individual measures wouldn't meet that very high bar to say in all circumstances these are the ones you want.

I think, though, it might be useful to note that there may be some contexts or some particular entities for whom these measures would be important to measure. And so I don't think that the view of this

Committee was to condemn all potential uses of these measures forever. But to point out that they're not the types of very broad, high-impact measures that nationally we would want government agencies and so forth to be using for direct reporting on quality.

CO-CHAIR THIEMANN: Mr. Levine.

MR. LEVINE: Yes. I guess what I was thinking about was, maybe the measures should be put in different paradigms. In one case there would be the measure for public accountability and reporting, another measure or a class of measures would be for quality improvement, and then a third would be whether they'd be allowed to be used in the system, like practice guidelines or maybe we need to start thinking about making distinctions in terms of those purposes or uses.

MS. THRAEN: Baseline standards of care.

CO-CHAIR THIEMANN: Dr. Lawless?

DR. LAWLESS: And actually, maybe

just one more thing that came up, and I apologize if I don't have the full incidence.

When we had the discussion about the patients with rheumatoid arthritis and we had the discussion of your relative making the agreement with their physician about what's to do or not. A lot of time, maybe, asking the measures to one, link to outcome. And then look at the outcome as in relation to the work it takes to get that measure done. So the incidence of a car accident or getting someone with severe rheumatoid arthritis into a car, getting to a lab and the stress that creates may be higher of a disease burden than it is monitoring the CBC every six months.

So, I think thinking a little bit out of the box, that people maybe have to look at what is the impact actually of actually having to get the lab done or the stress level gets created, and is it worth the results of the test?

I don't know if that's clear enough

with it, but it struck me that in having older people I take care of, getting them into the car, getting them out there is more stressful than actually, oh great, CBC was normal. Thank you. But now you're sick because you've had to move.

CO-CHAIR THIEMANN: Dr. Nagamine?

DR. NAGAMINE: Sort of along those lines, it's the number needed to treat concept. You know how many CBCs and LFTs do we need to do before we prevent harm, I think is a critical question that I have. And I think an area of opportunity for us to define as we move forward.

I mean, I kept on trying to look for these guiding principles, and although it is written here importance to measure and report; you know, high impact, how are we defining that? I think we could be a little clearer perhaps. You know, that's possible to provide a little guidance there.

And then I just want to reiterate

that we're not saying that it's not important to create standards and to measure these things. And I just wanted to emphasize that again.

CO-CHAIR THIEMANN: Dr. Conway?

CO-CHAIR CONWAY: They may want to turn me off.

But I would just leave the NQF staff with two ideas. One is, it would be nice to see someone want to pick up this mantle and develop an interdisciplinary and holistic approach to the monitoring of high risk medications. And I don't know how to do that, but hopefully someone in NQF, you could sort that out.

The second is in areas of research. I don't have a fatalistic outlook on this even though these are infrequent complications. I think with databases today, Ingenix could be the comparison to the wild, the control state and you could compare them to, Janet is right. Kaiser has highly

standardized care in a lot of areas. So if they're in fact doing medication monitoring in a regular basis, that could be the treatment -- and you could look and see if there's any difference in the complications despite all that monitoring that some areas of Kaiser may be doing.

And there's growing pooling of databases in the country today, trying to look at clinical effectiveness, and they could be looking at complications as well.

So I think a clever health service researcher could begin to understand this a little bit better for us.

MS. THRAEN: Actually, he just said what I was going to say. This is opportunity, I think, to feed back to AHRQ and the direction of their research for comparative effectiveness.

A great example of all of these proposed measures, a great opportunity to say whether or not outcomes are really impacted

one way or the other. And then feed that back through this loop. So if there does become an opportunity for someone to check that out, to do that research and they have preliminary information that could be fed back to this so that we could revisit it, then we might make a different kind of decision based on that.

CO-CHAIR THIEMANN: Any additional comments, thoughts, closing remarks at this point for today? Knowing we'll be back here tomorrow morning.

MS. THRAEN: I want to thank the staff for all their work.

CO-CHAIR THIEMANN: Yes. Thank you very much. We really appreciate it.

MS. MUNTHALI: Thank you, everyone. We will see you back tomorrow at 9:00. And we ask that you please take your materials. This is not our office. We're renting the space. And bring them back tomorrow and bring yourselves back.

So, have a good evening.

CO-CHAIR THIEMANN: Thank you, Dr. Solomon for joining us.

(Whereupon, at 2:22 p.m., the meeting was adjourned.)