

Innovations Consulting Group, LLC

**Initiative on Methods, Measurement, and
Pain Assessment in Clinical Trials**

IMPACT-XVII

**Assessment of Physical Function in
Analgesic Clinical Trials**

**Thursday, April 18, 2014
Westin Georgetown
Washington, DC**

Carole Ludwig
Transcription Services
141 East Third Street #3E
New York, New York 10009
Phone: (212) 420-0771
Fax: (212) 420-6007

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
Q&A and Panel Discussion: Social Participation, Work Participation And Caregiver Burden Outcome Measures <i>Moderator: Philip Mease, MD</i>	3
Interpreting the Clinical Importance of Improvements in Patient-related and "Objective" Assessments of Physical Function <i>Daniel Clauw, MD</i>	53
Q&A and Group Discussion: Patient-reported Versus "Objective" Assessments of Physical Function <i>Moderators: Daniel Clauw, MD; Lee Simon, MD, and David Williams, PhD</i>	69
Group Discussion I: Considerations for the Assessment of Physical Function in Analgesic Clinical Trials Using Patient-Reported Outcome Measures <i>Moderators: Bob Dworkin, PhD; Philip Mease, MD; and Dennis Turk, PhD</i>	100
Group Discussion II: Considerations for the Assessment of Physical Function in Analgesic Clinical Trials Using Clinician, Observer, and Laboratory Outcome Measures <i>Moderators: Bob Dworkin, PhD; Philip Mease, MD; and Dennis Turk, PhD</i>	155

**Q&A and Panel Discussion: Social Participation, Work
Participation, and Caregiver Burden Outcome Measures**

PHILIP MEASE: Good morning, could we come to order, please, and could the panel members join us on the stage, that's Laurie, Monique, John, Ashley, Shannon and David. This is great. So what we're being tasked to do is to remember the discussions from the end of yesterday afternoon, the presentations by David, Monique, and John, and then also just have more free ranging discussion about some of the additional topics that were brought up

I thought since Shannon -- well, actually, Laurie really didn't have a chance to present yesterday either. But I've asked Shannon Smith, since she did not have a chance to present yesterday, just to in a sense introduce herself and also to give a few comments about some of the work that she's been doing with Bob Dworkin related to, if you will, instructing or training patients to do questions and questionnaires, just because I think it's a side area of interest that affects all of us. And then Laurie, if you could be, since you didn't have a chance to present formally yesterday, if you can be thinking about a few comments that you might make to introduce yourself, then everybody else --

LAURIE BURKE: That could be dangerous.

MEASE: Shannon, do you want to go ahead?

SHANNON SMITH: Sure. So I hear a lot of perception in the kinds of measures that we've been talking about, you know, the social participation, the work participation --

MEASE: Shannon, could you describe first what you do?

SMITH: Okay, so I am at the University of Rochester and we have done a lot of systematic reviews of methodological issues in the literature. So looking at how adverse events are reported, comparing what is showing up in a manuscript versus what is on clinicaltrials.gov, so are there discrepancies between what people say they're doing, probably prespecified, to what they're putting in the publications that they have from these clinical trials? Also looking at pain intensity ratings that are reported in the manuscripts, so how well are they described, are they telling us how frequently the participants were filling out those questionnaires, were they asking them to rate their average, their least, their worst pain...so a lot of methodological systematic reviews.

And then the other bigger project that I alluded to a little bit yesterday was looking at a project that we are doing on training participants to be better raters of

their pain intensity. So if you think about it, pain intensity is the key measure in an analgesic clinical trial, and participants get no training about what they're supposed to be doing. They're asked to rate their pain and they're given a 0 to 10 NRS or VRS, or a VAS, they're given some measure to rate their pain and they're not given any instruction about how to use that.

And so we have developed a pain rating training system to help people to understand what they're supposed to do and to give them more of an absolute scale so we ask them to, okay, so on the 0 to 10 scale, what is a 10, what will that actually be like, can you think of an example of what the worst possible pain imaginable would be. And we really push them to not just think about their own experience but really to think about what the worst experience that anyone could have.

And then we also asked them to do the same for a one. So what is a one, what would really be the most minimal pain, but still pain for you, and we've heard really great feedback from participants in the cognitive interviews that we've done. We hear them saying things like it really helped me to anchor myself, you know, I really had to think like, okay, well I would have said before that my pain today was an 8 but then I thought, well

it probably isn't an 8 if my 10 is, you know, burning alive. And so it really helps them to anchor and give a better, more accurate description of their pain intensity each day.

And so I think that that is related to what we're talking about here because a lot of what these measures that we're talking about involve is perception, and pain intensity is about perception, right? And so if we're training people to be better raters of their pain intensity and teaching them to think through the day about what their pain experience is and not just kind of have this abstract view of, oh, my pain intensity is a 4, it's pretty much a 4 every day, I'm just going to rate it a 4, if we're teaching them to sort of couch their perception in reality more I think we could do the same sort of thing with physical function, with work participation, with social functioning to have them think more objectively about their experiences rather than just having sort of a view of, okay, well, things aren't going great, I'm just going to think, you know, I sort of have this schema that my life is not that wonderful and maybe not appreciate the things that they have in their day to day experiences.

MEASE: So one of the issues that that brings us is that I know at least at our Center the coordinators are

very, very, very careful and concerned about not coaching patients. And so to comment about the line between teaching or training and coaching, especially when it comes to having the requisite threshold of intensity to get into a trial, for example.

SMITH: We do the training before they start the study so we give them the information about, or we give them guidelines about what they should, so we say come up with what an example of a pain of 10 should be, come up with an example of what a pain of 1 would be, but then we don't say to them that doesn't sound right, that's not an accurate, that's not what a 10 should be.

We will ask follow-up questions, so is that the worst possible pain that anyone could have, and if then they say, well, no, then, you know, they are kind of prompted to think more carefully about what that worst possible pain would be. But we don't give any feedback that they're wrong in any way. We let them come up with the anchors, and then used that to help guide their pain considerations every day and then no further.

MEASE: And do you have any comments about differences in ethnicity, gender, socioeconomic status, et cetera, in the training process that you've had?

SMITH: So we have had some people who have told

us, you know, we'll ask the question at the end of the training so, so, you know, does this seem like it's going to be something that will help you and some people will say, no, I already really kind of knew how to rate my pain intensity. But we do hear a lot of, and I don't know how it's tied to SES or anything like that, race, or anything like that, but then there are a lot of people who say, you know, it gave me some guidance about what I should be doing day to day to rate my pain. It gave me that framework so that I wasn't thinking like, oh, my pain is an 8 even though it is probably more like a 6.

VIBEKE STRAND: Phil?

MEASE: Yes, Vibeke.

STRAND: Well I'm just very curious because I think that in general, patients who have been experienced with pain have already had what they would say is the worst pain. And what you're asking them now, an abstract comment of what is the worst pain you could possibly have. So to me that seems, I mean we've done a lot of emails back and forth and I think what you're doing is really good, but I'm just curious, you are now asking them to abstract it and take it out of their experience and put it into some other context.

SMITH: Right, we wanted it to be an absolute

scale, so we didn't want, so my worst pain that I've ever had is childbirth, but if tomorrow something worse happens to me, now that's my new 10. We didn't want their to be this range that that worst pain intensity could change if something worse happens. Yes, it is abstract, but we wanted to make it something that really shouldn't ever be worse than that, that should be the worst possible pain.

STRAND: But you're not a chronic pain patient, right?

SMITH: Yeah, that's true.

STRAND: And I think a lot of chronic pain patients are that because they have had a really bad experience. And so I'm wondering whether they can actually really go beyond that to abstract it to something worse.

SMITH: If they say to us that their worst pain is some experience that they have had, and we say, okay, is that the worst possible pain that anyone could have and they say yes, then that's fine. They can use their own personal experience. We just want them to think, we want them to really carefully consider whether or not that is the worst possible pain that could ever be experienced.

BOB DWORKIN: Vibeke, this is based on Charlie Cleland's brief pain inventory where a 10 is the worst possible pain I can imagine. And so we figured that is

arguably the most well validated pain intensity measure and so we based our 10 on Charlie's 10.

STRAND: No, no, no, I hear you, that's the worst pain I can imagine, but you are asking something different. You are asking what is the worst thing that someone could have. You got abstracted beyond what I could have. That's why I'm asking.

SMITH: And that's true, but we also, if they come up with something that is their own experience, that's a perfectly legitimate example as long as they feel like that's the worst possible pain imaginable.

MEASE: Lee.

LEE SIMON: So just help me out to understand, I understand the academic free-floating sense of what you're doing, what I don't understand is how you are going to apply this in the context of clinical trial work. I mean the problem that I have is I can tell you I have to be put to sleep to clean my teeth, my wife gets drilled without Novocaine. The thought of that I could collapse right here as I think about that (laughter).

So what is it that you're going to do with this information that you're culling from these patients in the context, you are not going to create any kind of absolute scale that is applicable to everyone, so what are you doing

with this? And it's interesting, but what are you going to do with it practically?

SMITH: So the idea is that each person then has a scale that is representative for them. That is the absolute range of what a 1 could be, the most minimal pain to the most absolutely worst possible pain imaginable.

SIMON: In each circumstance? Meaning if they go into a clinical trial you believe that that criterion that they've applied to them would be applicable to whatever experience they go into?

SMITH: I mean that's why we want it to be an abstract, absolute scale. So we really want them to come up with the least and worst possible pain.

SIMON: It's interesting, because that's very different than what Nat's doing when he is trying to create a structure to select people that can reproducibly answer whatever pain they're having.

SMITH: It is very different.

SIMON: And the question is what do we need. And I just wonder if we thought about the issue of what we need. So risk based trial analysis which is now FDA, I think it's FDA speak, the idea that you can be sure that the information that you're obtaining from the patient is consistent throughout the clinical trial, the context of a

chronic pain patient becoming inured to having a new reality, which may change over time, and all of those issues and their inability to answer the questions consistently, like how much pain are you having, average pain daily versus at the log rank analysis at the end of the trial, whether there is consistency in that which there isn't. and we see this all the time, 25 to 30 percent of the patients won't give you a consistent answer during the time period of the trial, and then at the end they give you a result. And then the area under the curve is not exactly consistent.

SMITH: So I can tell you one thing, so we've been doing this study and we assign people to get training and have human assessment every night, get training and have an automated call every night to rate their pain, and get no training and get an automated call. And so it's not an intervention study, it was observational. We haven't finished looking at the analyses but one analysis that we were very interested in is they rated their least, their worst and their average pain every night.

In the trained condition where they had a human assessor, somebody calling them every night, we had significantly fewer people who were making a mistake in the order of least, average and worst. So it was like 14

people -- 11 people made mistakes out of the, I think there were like 70 in that condition.

SIMON: Well that phone call what do they do, because they're eliciting a pain response, so that's not just a spontaneous issue. What's the conversation that they have that night?

SMITH: So the assessor calls up and says, you know, is now a good time, so there's a little bit of flexibility. And then if it's a good time for the participant they ask them to just remember to think back through the day, think about the morning, think about the middle of the day, think about the evening, and then the participant is asked to rate their least, their worst and then their average pain intensity that day.

MEASE: This is interesting. I just wanted Shannon to give a little introduction to herself. Sharon, do you want to go ahead?

SMITH: I'm glad everyone is interested in this.

SHARON HERTZ: But Lee, I think that this is kind of critical to get at what you're saying. If you can, and I'm a little skeptical about whether this is achievable, but if you can develop an absolute anchor, then that should ultimately translate into potentially greater consistency because you are always in the same place. You know,

someone has pain, I have pain, oh, it's an 8, it's an 8, oh, now I really have pain, no, this must be an 8. That is going to change over time. But if they have a mindset, so that's very good. But beyond that, you know, we have, you may have heard, had a little criticism about the use of terms like moderate and severe, and that is the same issue. Because we've got people with severe pain after they get their tooth pulled and people with severe pain after they've had their abdomen operated upon.

So ideally, I mean while this is not technically an objective measure of pain, it would be great if we could standardize the conceptions underlying rating pain and then maybe the 40 to 70 entry criterion for third molar extraction will drop down to 20 to 50, and then, you know, for post hip it will be 50. So ultimately, even though the approaches sound different, maybe this could lead to that.

MEASE: Dorcas and then Bob.

DORCAS BEATON: I think that this is all going exactly to the question I had because this is what's called response shift phenomenon often, right, where you try, we know that people do shift their anchors even if we just try to ignore that it's happening. And so by training you might be able to achieve that but it sounds like your study is set up to do that and I'm wondering if you are able to re-

talk to these people about what a 10 means after you have done some of this training to see if it is stable.

SMITH: Yes, we did. So we asked them, so we did some cognitive interviews with people who had just done some training and then immediately after that asked them the questions. And so, of course, they remember what their 10 was but we also, for the people who did this 3 week study with us, 40 of the 240 participants did cognitive interviews and I don't have the data on that but we did follow up with them to make sure that they are remembering like what is a 10 for you, what is a 1 for you, and so I'll have to look at that data once we are done.

BEATON: And then there's ways of analyzing the data to see if, in fact, you're getting that correction of the recalibration.

MEASE: Bob, did you want to say something?

DWORKIN: So I also wanted to say just one thing in response to Lee about Nat's approach, as I understand Nat's approach. What he's doing is selecting individuals who, for whatever reasons, report their pain in a consistent way that matches increasing temperatures of a thermal probe.

SIMON: It's a pressure probe and/or a thermal probe.

DWORKIN: So selecting a subgroup of patients who apparently are talented at reporting their pain. Our approach is different, so what we want to do is train everybody to be better so we'd like to train patients to get to the point of the ones, the subgroup that Nat is selecting. And so we think that training people to be better (indiscernible) of the pain, you know, especially in a phase III trial is more reasonable than using QST to select a small subgroup of patients who for whatever reasons are naturally good at it. So it's two different approaches.

MEASE: Hold on, I want to give Bob Kerns, David and then Vibeke and I know Ernest had his hand up here at some point. Okay.

ROBERT KERNS: So I find this a very exciting and interesting discussion about pain levels, but I'm not sure about this particular approach. It seems to me within a, however you think about it, in a biopsychosocial framework, a matrix framework, multidimensional framework of some kind, that you are targeting this idea as if people can somehow isolate their thinking from their environment, their, yeah, I guess their environment, their experiences before, expected experiences after, so forth.

It seems to me within that multidimensional

framework, what we are trying to do is try to improve some, I guess the word consistency, reliability, it seems much more likely that one could try in this environment to almost do something like I guess Nat is doing, which is try to control aspects of the environment like time of day, the size of the room, the lighting in the room, the temperature in the room, things that we, the experimenters, can control, and reduce variance in that way, than trying to train people as if they, that there is some idea that they can capture something reliable about their experience of pain, as if it wasn't vulnerable to their mood, the lighting in the room, the temperature and so forth, all these other variables.

So I just don't think that it's, I'm not confident that this is worthwhile.

MEASE: David.

DAVID HADDOX: Thank you, Dave Haddox, Purdue Pharma, and prior to that largely academic pain medicine practitioner. I'm very interested in what you guys are doing, Shannon, but the question I have is we hear a lot about the sorts of variables you're looking at in terms of minimum, maximum and average.

In my clinical practice, I spent a lot of time trying to find out the modal pain score, because to me, at

least what I was hearing from my patients, that was more important than some fleeting flare that was there 15 minutes and was gone, is how do you feel most of the day. And I wonder if you have any comments about that?

SMITH: So one of the things that we initially put in and then have since been debating about, is really encouraging people to recognize how long each pain intensity lasts. So, yes, your least is this, yes, your worst is this, yes, your average -- I mean when you're calculating your average, don't just bias it based on what your worst was, think about did your worst last for a minute, 5 hours, you really need to think about the intensity and how long it lasted.

HADDOX: I had problems with some people in my practice with the concept of an average. I don't think they understood it to be what we understand it to be. And that's why I would ask them so just thinking about yesterday, most of the day what was your most common pain score, and how do you feel about that? And that's what I manage my treatment on largely.

SMITH: That's an interesting way. I mean we were trying to have them really think carefully, like what was your pain when you woke up this morning, how long did it last at that intensity, what did it increase or decrease

to, what was it in the middle of the day, really walking them step by step during the training, hopefully that would then lead to them being more able to do that when they're in any clinical trial. So I think sort of similar approaches.

MEASE: Vibeke, then John, then Jas.

STRAND: So I just think that we've been having a good conversation but my concern still is about the theoretical. So instead of taking them from their experiential thing, which may change at any time, but to have them now try to put their pain in the context of something they could imagine that could even be worse, that's where I wonder about the value. Because now you're getting to a place that's beyond what they necessarily know, and I was just wondering how many of the patients actually told you that they already knew what the worst pain could possibly be?

SMITH: So I have only looked at the first 20 cognitive interviews that we did and I would say probably two of those people said that they didn't need any further help.

MALE VOICE: Vibeke, Dennis and I here are completely confused here by your point, what is the difference between the BPI item of the worst possible pain

I can imagine and asking the patient what the worst possible pain they can imagine is? We're not following the distinction you're making. It's not the worst pain you've ever had or the worst pain you can ever, the worst OA pain you can imagine your OA pain ever becoming, it's the 10, in our training, and I think in Charlie Cleland's measure, you guys know Charlie Cleland's measure, is the worst possible pain you can imagine, which for me is snorkeling in the Caribbean and being ripped apart by a shark. I've never had that happen to me --.

So anyway, so we're not understanding, Vibeke, the distinction you're making between being ripped apart by a shark and what you are saying you would like to ask --

MEASE: So for Lee it's the dentist's drill and for you it's a shark. Lee.

SIMON: I just have a comment, a concern. So the comment is, I feel like I deal with the photographic negative of this experiment every day in clinical practice where 16 percent of my practice are patients on worker's compensation, all of those patients have a disability attorney, and they are trained by their disability attorney to have a pain that is no less than 7, but somewhere between 7 and 9. And no matter how functionally better they get, they have to have a 7 or a 9 or I'm going to redo

their percentage of disability.

So those to me are the patient rating training experts in our society because they train people that, you know, again, and those patients will come, they exhibit all the functional benefits that we expect and we kind of are in this winking game where they'll tell me they're doing more and they're doing better, but they won't change their pain score. And we all know what's going on. And, the insurer, workers comp is the insurer of last resort in economically depressed regions in the United States, and that is how you get your osteoarthritis managed, and that's the way it goes.

So I think, that's kind of the photographic negative of this where your function gets better but you don't change your pain score. That is sort of the clinical sort of commentary. My only concern about this is that I think there is something slightly anti-therapeutic about having people really focus on their pain score all the time. And I think that there is a downside to that, and whether it's in the context of a clinical trial or in clinical practice, the more you have a comment and a set of discussions anchored on that number, I think the less well people are going to get at the end of the day because their whole life is going to be thinking about, well, where am I

now, is it an 8, is it a 7, do I not want to do this because it's a 6, so that's my thought.

MEASE: I am reminded of the fact that physiatrists really teach us to not lead with a question about pain but a question about function. Let's see, Ernest, you've been really patient, and then Jas after that.

ERNEST CHOY: I'm going to revert to my OMERACT mode. So one of the things about OMERACT is that it has to be data driven. So as far as I'm concerned, this training is about increasing the measurement accuracy of a measurement. So what I would like to see is before and after training that the scoring is improving consistency. So what you could do is give the patient a standardized series of nociceptive stimulus, a series of pain, mild, moderate and severe, and if you do it at 3 months afterwards, after your training, (indiscernible) score and the covariance decrease. If you can persuade that the covariance decrease, you have showed that the measurement is showing improvement, we should do it.

MEASE: Great, so that's your next --

SMITH: Yeah, that would be a great next study.

MEASE: Jas.

JASVINDER SINGH: So I want to echo what Phil

said and what David alluded to, and this might be again maybe a subgroup or a new group you want to look at. But in downtown Birmingham my patient's educational level is on average about 10th grade in my clinic, which is higher than some of my colleagues in San Francisco and other places. And when we do the numeracy test in our patients and this goes back to the average concept and also the worst concept, that more than half of them can't add, can't average, and they fail the numeracy test, the simplest of those. And these patients typically among that group of patients are those in the lower socioeconomic strata, but there is not a perfect correlation between the education level and the health numeracy level. And all our measurements at a 0-10 other than what is it that now requires some numeracy competency.

So it might be interesting for you to look at the subgroup of patients, either by SES or a numeracy test threshold to see how many of them actually understand concepts other than what is your number between 0 and 10 now. And even for that, I have to sometimes break it down into a dime and a nickel and a quarter for my patients and so what they're doing is a 5, 10, 25, 50, not a 10, 20, 30, 40, 50, 60, 70, 80, 90, 100. And we do this as a matter of habit even in our clinic with the ACR responses, but people

are not scoring it like that. So it might be interesting to look at that.

MEASE: Penney.

PENNEY COWAN: People with pain, there's a couple of things. First of all, the more they think about their pain, the more they're going to suffer, so I'm with Dr. Markman that we do not want people to think about their pain all the time. But secondly, for many people with pain, not in the clinical trial but in the real world, they want somebody to believe their pain and to tell them, you know, this is your worst pain and then to continue to rate it lower, I mean they need that validation that this is the worst pain ever. And even if it may have improved, if they're still looking for treatment they're going to explain how terrible this pain is. And for many people the pain that they are experiencing at the time is the worst pain, because that's what they're experiencing right now.

MEASE: Roy, and then Monique, and then Lee.

ROY FREEMAN: This has been pretty fascinating. You know, I think most of us have delivered these questions to patients for the past 20 years dozens of times, and I can say now personally I've never really thought about this with this degree of detail. And I think it comes down to the concept of worst pain imaginable. And I think the

question actually is not a particularly good question because when we ask worst pain imaginable, are we asking worst pain imaginable in your limited experience of pain, which is one question, your worst pain imaginable, in terms of the greater world out there of pain, Lee being attacked by a shark while having his teeth drilled (laughter). Which is an entirely different question.

And, you know, clearly what you are aiming to do is have accuracy, reproducibility, and to bring it down to its most concrete during the course of a clinical trial so you can show the effect of a drug, and quite clearly what you have to do is to make sure that your anchor is the same. You can't go from the shark attack to your limited experience of pain. So it must be the same anchor at the beginning and end, but which is superior I have no idea, and I think this really needs to be data driven as was suggested.

MEASE: Monique.

MONIQUE GIGNAC: I'm just trying to justify why I'm sitting here, so I think that there is an issue, I agree that hyper vigilance regarding pain is a problem when people are constantly focusing on it. And because there are all these real things, response shifts and other things that go on and difficulty understanding pain, I am

wondering then is their value in linking pain more directly to function. Regardless of where you are moving on your pain, are you able to take care of your children, are you able to participate in the activities that are meaningful and matter to you? Can you work? Are you a burden to your spouse and driving him or her crazy? Those kinds of things perhaps have some value and would ultimately help in understanding what's a real meaningful pain score.

MEASE: Lee.

SIMON: So just to close the loop on what has been discussed and what Bob referred to, so we're presently doing two trials with Nat's technique, one of which was the drug that doesn't work called hyaluronic acid supplementation, which we all know for OA is usually not very helpful, and yet if we demonstrate an increased effect size in the measurement system, that's one thing. And then another one, which is an oral therapy for osteoarthritis, which might be again informative in that particular regard, but I really want to emphasize that this is not a small subset of people that actually get into the trial passing this bar. It's actually the 75 percent of the patients whereas it's the 25 percent who don't get in.

And in analyzing at least 30 to 40 trials, and to think about this issue and then go in front of the FDA and

get them to accept it as an approach to select patients, it became evident that it's consistently a small group of people that tend to have a problem in reproducibly answering the question rather than a large group of people. And all the system is doing is selecting them out using a technique that says, okay, every time we apply this, do we get a similar number in the response and, you know, it's cycled and you get four or five or six measurements and if they're consistent, fine, but about 25 percent of the patients, it's a 1, it's a 9, it's a 6, it's a 4, and you get rid of those people as soon as possible.

Now why people do that is a different question, which I think you're getting to and I think that we will learn a lot from what's being done here, I just don't know if it's the right way to get people to participate in clinical trials, but that may not be the purpose.

MEASE: Veeraindar.

VEERAINDAR GOLI: Thank you, Veeraindar Goli, Pfizer. I just wanted to make a comment on this anchoring that you are talking about, and Roy mentioned about anchoring patients with worst possible pain and least possible. And what I'm trying to struggle with is you're rating patients with chronic pain, but the anchors we give them are for acute pain. You know, being ripped apart by a

shark or being drilled, all these are acute pain analogies and we try to anchor our patients to that, what we always tell them, chronic pain is different from acute pain.

So I'm trying to understand how we can take that leap from acute to chronic pain, whereas in chronic pain you have an affective component and you have a nociceptive part, and trying to tell them that being ripped apart by a shark is a 10 whereas the guy is sitting there and he's saying his pain is a 10, you know, so how do you reconcile to that?

MEASE: Great question.

GOLI: And the second part to that is we say that the pain scale is linear, is it really linear when a patient can reduce from a 9 to an 8 much easier than from a 5 to a 4? So how do we measure that?

MEASE: John Farrar.

FARRAR: So the 0 to 10 or the 0 to 100 scale is not linear, never been linear, shouldn't be considered linear and should be analyzed with non-parametric statistics, period. Now that isn't what's done, that's the way it should be.

Okay, I think we need to be very clear about what it is we're trying to do with this discussion. What we're trying to do is to improve the instrument we use to measure

the thing of interest. If we had a blood pressure cuff that sometimes gave us an 80, sometimes gave us 100, sometimes gave us 120, we would throw the thing in the trash. And it's not to say that the patient is not having high or low blood pressure, we just don't know the answer.

In patients we have this same issue with the 0 to 10 scale, there is a subgroup of them who for whatever reason are not able to use it as an effective communication tool of the rating of their pain and will add random variation to all of our studies which makes it harder to achieve a P value, albeit it won't change the direction, right? It's not going to change whether it's a negative direction or positive direction.

And so one of the questions that's come up here is how best to do that, which is testable. And that's what Nat is doing with the things that he's working on, that's hopefully some of the things that Shannon will be working on, and it makes perfect sense to me. I don't think we here, today, have the information necessary to move forward with that ascertainment.

With regards to what David was saying in terms of the process of taking care of patients, we all reorient our patients. You know, they come in and they say they're a 10 and they're sitting comfortably in the chair, and you say,

okay, I need some help here. So I would argue that we actually, in my clinical practice at least, do training with my patients so that I can understand what goes on.

The last thing I would say is that we use the 0 to 10 scale, and it's an outstanding scale for clinical trials because it shows change very nicely and patients report it well. The thing it does not do is to ascribe meaning to that pain. In our hospitals we've been using 0 to 10 and if they're above a 5 they should get some attention paid to them. I actually have backtracked hugely on that process. I think actually what I really want to know is, is this tolerable pain, is it mild, moderate, severe, and is it tolerable, is this okay.

That's not a good way to do it in a clinical trial. It's a different way of thinking about things. But in clinical practice it's I think a reasonable way to do it and we need to be careful not to mix those two up. I do think that maybe we ought to think about getting back to the connection (laughter), so I'll leave it to you to do that.

MEASE: Right, we've been at this now for about 40 minutes and we're just about to start out panel. But I think it's been an incredibly interesting discussion, I hope that it's okay to have allowed this to go on. If we could -- David, do you have --

HADDOX: I have a function question.

MEASE: Okay, hit it.

HADDOX: Is that okay, is that a good segue?

So again, going back to my clinical practice, after I left academics for a brief stint before I joined Pharma, I was practicing part time in Birmingham, so I know your patient population very well. And that leads me to this question, we were using the BPI in our clinical practice in Birmingham and in Atlanta, and it dawned on me after sometime that I was having a substantial portion of my population who had real problems in understanding what interference meant on those life domains in the BPI, the sleep interference, the activity interference. And I noticed that terms shows up on, I think it was the WPS-RA I think Monique you talked about. And I'm just wondering, we wound up actually modifying the BPI for our clinical use.

So, for instance, when we said no interference to completely interfere with sleep, my patients didn't know what that meant. And so I changed it to no effect, completely prevents sleep. Then they said, oh, I know what that means, that's a 6. And I'm just wondering has anyone else run across that? This was a huge issue for us, because we had this disconnect between what they did in the waiting room with our assistant and then when I was doing

my clinical interview, I was saying but you said this completely interferes with your sleep but you said you're sleeping 4 hours a night. And he says, yeah, but I don't sleep like I used to, and I thought, oh, okay. I'm just curious about that word interference, do people get that concept?

MEASE: So does anybody on the panel want to

--

GIGNAC: I'll start. I'll start. I think there's a lot of complex language that we use that does have different meanings to people. I think we are asking people very often, as we've been talking about here today, to do a lot of mental gymnastics, and that can be very hard for them, and making global sweeping statements about roles and activities can be hard. All of which really pushes for more complex, often longer measures, you know, we heard yesterday about the social role, participation, they're often quite long and ask you about different aspects of the same dimension in your life. And by kind of targeting in that way, you do get hopefully a better picture that if you fed it back in some cognitive debriefing to your respondents, they would say that's right, that is me.

On the other hand, that's exactly not what I'm hearing people who are doing interventions and trials want

because they're worried about burden on their patients. So I think we do have a fair bit of work that we need to do to try to find simple ways, simple language, to ask people about their activities.

In the work research with the global items, that's why we haven't been very quick on throwing something out because a different, I mean now if you used all four global questions, for example, that's only four questions, but different ways of asking is something difficult, is it hard compared to your coworkers, is it hard compared to your best day, it can kind of get at it and they often actually are related. Although people have trouble, they're more related to one another than we might think.

MEASE: Any other comments? Yes, Veeraindar.

GOLI: This is a function question. So practically, in terms of clinical trials, what I am hearing is that function may not be generalizable to all patient populations across the board, because it's difficult to see whether function is always inversely proportional to pain or not.

So I guess the first task is to try to identify what subgroups of conditions of patients, what are the phenotypes of patients who would respond to a function question where it would be appropriate to ask a function

question, and what groups of patients where it is not appropriate to measure function?

MEASE: Laurie, yes.

BURKE: Laurie Burke, ex-FDA, and I'm one of a group of many ex-FDA people and we're all connecting with each other after leaving FDA, it's really quite exciting (laughter).

MEASE: Is this like AA or something (laughter)?

MALE VOICE: We have group step meetings, it's exactly right.

BURKE: So I think that by way of introduction and summary comments, in maybe launching today's discussion on what we're going to do about function I have a few thoughts, can I share those now?

MEASE: Please.

BURKE: Okay. So I mean we really need to decide what we're talking about and what the end goal is of what this discussion is, and that's been a little frustrating for me sitting here because we're not going to solve the problems of the world in terms of function in two days. So in my opinion, the end goal is to have a framework to talk about function and chronic pain for the context of labeling, advertising, but also health technology

assessment, where we're talking about the evaluative context of use. I know that was mentioned, you know, we're not talking about clinical practice, and we're not talking about other ways of using these outcome measures.

Now that would be my opinion, I'd be interested to know whether the organizers would agree. But I think that it's important to include a health technology assessment piece because that's where the world is going, it's not just about regulatory, it's about who is going to pay for this and what's the structure for making those prescribing decisions on the part of the clinicians. And that is still an evaluative context in my mind.

I know that -- from what I've observed, no one is really doing that well yet. I mean taking the concerns of regulatory on a global scale and combining them with the concerns of the health technology assessment bodies on a global, with a global consideration. And also in the US, the advertising concerns. But in my mind they're all the same criteria for evaluation in terms of the general standards. In the discussions that we've had at ISPOR, the Dublin meeting last Fall was focused on health technology assessment and it was very clear if you look at the presentation from Mira Pavlovich there, who is leading the EUNEHTA activity, the EU Network of HTA bodies, integrating

with the EMA approval process and doing a pilot that they are, the standards that we are talking about in US regulatory environment are similar to what they're talking about. I don't think that we have to meet enumerable number of standards in our discussion about what we're trying to do with function.

But I also think that it's critical to pay attention to what Ashley and Elektra presented yesterday with the way that we consider direct versus indirect evidence of treatment benefit. So in each context of use, meaning disease, entry criteria, study design, there is that idea of what direct evidence treatment benefit needs to be established in the thinking of those who are designing their clinical trial and developing their measures.

And then there is the issue of proximal versus distal, the core signs and symptoms, the immediate impacts and the more distal impacts, in order to sort out what it is we're measuring. But the other thing that is new that you haven't had so much discussion about in the past is the meaningful health aspect. Meaning if you have a performance item, what meaningful health aspect is that performance measure actually trying to get at.

So if we could get, and of course I think it's

really important to get to question one on this, and I know you're going to do this throughout the day, but I thought a little bit about what the structure might be for doing that. And I propose that we have different levels of concepts. There's general types of concepts that can't be measured easily with a single score, participation, fitness, capability, things like this.

Then there are the meaningful health aspects within a particular population that might be measured, that might be approximated with measures is what I'm trying to say. So that would be physical activity in daily life, functional status perhaps, functional limitations might be another one. And then so what measures do you actually use to approximately measure those meaningful health aspects, so those are the actual tests. The physical performance tests like 6 minute walk test, the PRO diaries that measure ADLs, IADLs or other general activities, the accelerometer that measures movement, and then with those, what are you actually measuring in terms of what's meaningful for treatment benefit.

So with the, to measure mobility, for example, do you do that with a PRO diary, or do you do that with an accelerometer, or do you do it with both? Physical performance, do you do that with -- what are you getting at

with the physical performance? You could be getting at interference, you could be getting at participation perhaps, but those connections, or ambulation, that's another general term that would be a concept of measurement. But the connection between the measures and these concepts I think are things people, you know, we say well what are we talking about and someone will say, well, we're talking about physical activity, but we're not just talking about physical activity, we're talking about physical activity in daily life, and then what does that mean with respect to these other more general participation, fitness, capability, ambulation concerns.

So I think that's a big discussion and if this group could get started on that today it would be really helpful.

MEASE: Really helpful overview comments and I know that Dennis and Bob and I had a heated discussion about the point that when we saw the huge difference between self perception of performance and actually performance as measured on the activity, we realized that we're measuring completely different things there, and probably both are important aspects to get at. But before, but these will all be helpful to come back to. I'm curious if anybody wants to tackle Veeraindar's question about the

phenotype of the patients that are being measured in terms of physical function and are all phenotypes appropriate or are there some that are inappropriate? Does anyone want to jump in on that, either on the panel in the audience?

FEMALE VOICE: I would only make one comment, I mean, yes, I think we need to understand this, but maybe we need to think about this as a bit of an iceberg. And some of the folks that you see, and by you here I'm talking about the group sees, are those the tip of the iceberg, the difficult, the folks that are having real difficulties there is something extra going on. But below the surface are many, many people who, for whom we can understand function, we have been looking at function for years and years, we've been measuring this very well, it's complex, just like pain is, but I think marrying the pain to the function is what your patients want.

I hear from people that "I go to my physician's office and he or she gives me this whole bunch of scales, they get very excited if I move 2 or 3 points on their pain VAS scale, but they don't ask me if I'm back at work, they don't ask me if I can pick up my kids and hug them, or take care of my grandkids, and that's what I want." Is your treatment successful if it doesn't do that?

And I know it's complex and it's not that easy

and other factors are involved and we, I didn't think we were talking only here today about clinical trials, sort of FDA kind of labeling kinds of issues, I know that is a big part of it, but it's not the only part.

MEASE: Right. Ernest.

CHOY: So I think that there are phenotypes, and certainly for conditions like fibromyalgia, if the patient has been inactive for 5 or 6 years because they haven't been diagnosed then it is much more difficult to rehabilitate this patient to any degree of physical activity because they go into this chronic illness state you find it very difficult to rescue them out of.

But coming back to the comment about physical function, in terms of health technology assessment, I come from Europe and clearly this is a very important piece, all expensive treatments have to undergo health technology assessment by a body called NICE and physical function is the single most important driver of that assessment.

Basically, they look at various measures of how to compare different expensive treatments and whether they will be paid. So you want something that is comparable across different diseases so they tend to look at very generic measures called quality of life measures, SF36. The problem about this is, in fact, in certain diseases the

ceiling and falling effect are very severe and actually we presented our own data just to justify it. While we can't use them in a simple way for that health technology assessment, but consistently physical function tends to link to health care utilization very well. So it becomes a very important piece certainly in the UK but I also know in different countries, in assessing whether something will be funded.

The one thing that is different in different countries, how they, what they look at in that health technology assessment. So the UK typically we will not include societal costs into impact of the illness. So we wouldn't look at user/caretaker burden, and we wouldn't be looking at work productivity as the benefit of the intervention, but there are countries that do. So there is still quite a lot of variation in how that assessment is being done.

FEMALE VOICE: May I?

MEASE: Yes, please.

FEMALE VOICE: I don't think that this group can possibly address the huge variation about what to measure and, of course, if you've seen one health technology assessment review, you've seen one health technology assessment review. So it's not that we're going to solve

that problem, but I think in the context of physical function alone, or whatever this group decides to take on as its chunk for discussion, I think that there is no standard terminology that OMERACT is putting together, now we're talking about this at IMMPACT, but yet there is no global ordering of these terms that we throw around in this, that we've been throwing around now for a day and an hour, and it would be good, at least for those of us that are trying to sort this out, if we had some standard nomenclature.

CHOY: That I completely agree. So certainly in terms of physical activity, in some countries is used to describe something very different now. Because obesity being a very big problem, funders are prepared to pay for things that will increase physical activity across all illnesses. So it will be perceived as a very different thing if we use the term physical activity.

MEASE: Ajay.

AJAY WASAN: This might have been asked but I didn't quite hear an answer. So just as there are inconsistent raters of pain, do we know if there are inconsistent raters of function?

MALE VOICE: Probably.

WASAN: Do we know that? That seemed to be

important to put in the paper whether we know or not know, you know.

FEMALE VOICE: And I think there will be, but I think it is also important to, you know, I'll just give a work example because that's my understanding a little bit more. And I think it was very interesting when you were talking about, we were talking about response shifts. If I think back to my life at grad school where I thought I worked night and day, and I would say I was very productive, in fact, it's nothing compared to my life now. So different jobs and -- so what productivity means to someone, it's not just a biased look at it, it is actually things are quite different, and my ability to handle activities are different.

And so I think we have to recognize that there is a perception, a subjective element, but that it's often a meaningful one. And when people talk about productivity, are you meeting the demands that you have for your job, and it's where I think the subjective and the objective are often married. People can do it, not everybody can do it as well, but it is meaningful, it's predictive, and I think we can do it. It's hard, but we can do it.

MALE VOICE: But I guess the other piece of it is, is that work published, that the reality is

inconsistent. So the quantitative sensory testing literature does have this stuff that Lee is talking about, you know, the inconsistencies between patients in terms of how they rated heat, pain, et cetera. So is that published for functional assessments, measurements, self report, et cetera.

MALE VOICE: The actigraphy graph that was shown yesterday where presumably actigraphy is the objective gold standard of physical activity and individuals ratings... there was an enormous inconsistency. There were no error bars, as Dennis mentioned, but that gives some clue that there actually --

MALE VOICE: That would be something to include, too, in the paper.

FEMALE VOICE: That's physical activity, that's not the same as --

MALE VOICE: No, no, I just wanted to introduce the notion that there is a potential anchor and that, in fact, this very much parallels the discussion as to the notion of anchoring these measures.

FEMALE VOICE: You can look at absentee rates versus, you know, other aspects at work that are hard outcomes related to work compared to some of the perceptions, as well. But Dorcas might be better able to

answer it.

BEATON: I totally agree with what you're saying. The other, I'd raise the point that when you were in grad school, if we tried to convince you that you were not working at 150 percent productivity, it's not an issue that you just have to, in that situation, if that's your true meaning, that's your belief, that's what you think really is the top of the scale.

In terms of some of the things on physical function and documentation of it, there is a lot of work in this response shift literature that if you look at it would suggest that when you go to the other extreme of quality of life assessment, which, indeed, in many cases includes a physical functioning scale, that people can shift the way they think about a very good level of functioning or a moderate level of functioning. So I think there is some documentation of that and we could work on trying to improve it and try to make sure that we've got good measures to avoid it, if possible.

MEASE: Yes, Jas.

JASVINDER SINGH: So one other concept about physical ability or physical function that was brought up by multiple people yesterday but I would like to probably get some feedback and some discussion around this, is this

concept of being able to do an activity with a lot of pain and anguish, and suffering, and then being able to do the same activity, not more often, not more frequently, not better, but without pain.

So an excellent example of this is somebody comes into my clinic with severe shoulder pain, they can only move their shoulder up to here, we do an injection with corticosteroid and Lidocaine and 2 minutes later they can still do that, but they say no pain. The pain is gone, they cannot do this as yet, a lot of them can do that, but a lot of them can still just do that. And we saw that in a randomized trial of one of the agents we studied, that patients came back and some of them said I feel wonderful and they were still, they were moving a little bit better, but some came back and said, doc, I don't know. And then we looked back at their notes and we had 70 degrees with pain of 9, we actually had them rate the pain at their best range of motion in our study, I don't know why we did that but we had it, one of the measures, and went back, a lot of the people in the intervention arm actually did the same elevation now with a pain of 2. So that could translate into people coming back, being able to do the same activity with less pain and suffering, versus a lot of pain and suffering prior to an intervention.

And joint replacement is an excellent example of this. You know, when we ask people before and after, people say well I cannot do that, you know, I'd love to be more active but this knee is bothering me and if my orthopod can replace this knee I'll do that. What happens after joint replacement, we've shown, is that people don't become more active, people don't lose weight, in fact actually they gain weight. Probably they're happier, but they're able to -- they're happier with the level of activity they have and that might be because they are pain free activities. They don't become joggers, they don't start losing weight, they gain weight, their activity level does not change a whole lot, but they're happier with what they can do in life without pain.

So I'd be interested in how we can capture that domain of function in studies and outcomes studies, that I think is very meaningful to the patients.

MALE VOICE: There are pain interference scales, which ask how pain interferes your function.

MEASE: So anyone from the panel before Bob want to address that?

FEMALE VOICE: I don't want to keep talking but I think very quickly there are three issues that you raise that did come up in I think the discussions on social

participation and other talks and physical activity, and those are issues of the value of an activity to a person, the importance of an activity. So if it's meaningful, if it's something you must do, sometimes work comes in that category or something you love like physical activity, people will push through the pain, work through the pain to continue functioning.

So I think to understand some of this we need to start asking about value and importance, we need to start talking about control, but then there is this whole thing around behavior change, and theories around behavior change. And they go way beyond pain and function. We all know we should get a lot of sleep at night, we should exercise, we shouldn't smoke or drink too much, but we don't do it. And that has nothing to do with -- well it has less to do sometimes with measuring pain or function and more to do with understanding how people change their behavior over time. And that's for another conference.

MEASE: All right. Bob, and then we're going to need to wrap up our panel to give Dan his time.

ROBERT KERNS: so there is so much to react to. I think conceptual clarity is important here. There are lots of different ways to measure lots of different things, we just need to be clear when we're measuring something what

we're measuring, the construct we have in mind to do. At the risk of being self serving, I keep waiting for my colleague, Dennis Turk, to mention this, IMMPACT did I guess recommend the West Haven Yale Multidimensional Pain Inventory Interference Scale, a 9 item scale, I guess 12, 13 or more years ago.

It's 9 items, it has a combination of items of how much does pain interfere with your ability to do X, Y or Z, or pleasure activities, social, recreational activities, and it had items, in the 9 items, about degree of enjoyment, essentially value, right? It turned out the development of that 9 items, they were designed to be separate constructs but empirically they merged together into a single scale with an internal consistency well above 90. And that 9 item measure has performed now well for 2-1/2 decades in clinical trials of various kinds in terms of being responsive to change, and I think is an oldie but a goodie that should be acknowledge, I guess. IMMPACT did recommend it and it seems to capture some of important concepts that we've been talking about today.

MEASE: Thank you. So why don't we bring this session to a close, this has been really quite interesting, illuminating, and next up is Dan Clauw. Dan is one of those that you always use the phrase he needs no

introduction. Dan comes from University of Michigan and is going to be talking about interpreting the clinical importance of improvements in patient reported and objective assessments of physical function.

Interpreting the Clinical Importance of Improvements
in Patient-reported and "Objective" Assessments
of Physical Function

DANIEL CLAUW: That was fun (laughter). And I really can't believe that Lee and Bob let on what their worst imaginable pain is. Your adversaries now know this (laughter). When you go into battle, you don't really want people to know where your weakness is, and I can't believe that you guys just bared your soul, it's going to be published on the website and everyone is going to know. Dental drills are very portable.

So the University of Michigan asked us to give disclosures, but one of the points I'm going to try to make in this is that I think measuring physical activity via objective measures like actigraphy should be moving towards being considered to even be a primary outcome measure for pain trials. I'll give you the reasons why I feel that, but I want to make it clear that I am not working on behalf of any sponsors who are trying to do this. These are my opinions that are not in any way driven by a sponsor that I'm working with.

So I'm going to really try to do four things here, first talk a little bit more about how strong the relationship is between self report and objective measures

of physical function. Then talk about a couple of studies, there are actually surprisingly few studies, even though this is often collected side by side in studies there's been surprisingly few studies that have actually directly compared the objective measures with the subjective measures and tried to figure out what it is that the subjective measure are really picking up when you can then compare that to an objective measure.

I'll look into other fields to say should we really expect a strong relationship between self-report measures and objective measures of physical function, if we look into other symptom domains or other domains do we typically see this? And then finally, something that I'll probably leave more for the panel, what is the right measure, given that I will tell you already that there's a terrible relationship between the self-report and objective measures, this will be, again, left for the panel, is what is the correct measure to use.

So first, how strong is the relationship between self-report and objective measures, not very. Again, I will make the case, Dr. Patel did a very nice job yesterday of talking about actigraphy, I will make the case that actigraphy I think is moving towards being the gold standard for an objective measure of physical activity,

which is different than function, and I will talk about that.

But if you look across all of the studies, there is always a poor relationship between activity levels and measures of functional status or physical activity. R values range from 0 to .4, so at best there's 20 percent shared variance and that's at best. Most of R values are more in the .2 range than the .4 range.

I think that if you look at the aggregate data, and it would be interesting if anyone else has looked at this literature, if they would agree with me, I think you see a fairly strong trend towards these relationships being much stronger. They're not strong, but they are stronger, if the self-report measure is a measure of physical activity rather than a measure of physical status.

If you actually look at some of these measures of physical activity that have been used in some of the studies where they're comparing it to actigraphy, these measures are incredibly prescriptive, they really are almost like a diary where someone is asked to say exactly what they've done, how long they did it for, they record those, they score those, they get a summary score, and those measures of physical activity correlate more strongly with actigraphy than measures of functional status which,

again, if you read the way most functional status measures are asked of the patients, it's more aspirational or what people would like to be able to do rather than what they actually did. The physical activity measures really ask people to recount what they actually did, they are more like a diary that is asking people what they did, whereas most of the functional status measures are more what can you do or what you like to be able to do. And so I guess it's probably not surprising that actigraphy would be more strongly related to the measures of physical activity than the measures of so-called functional status.

In the studies that directly compare self-report and objective measures, what are the self-report measures really measuring. So I'll present two studies, one that we've done in our group, one that Nat did that I alluded to yesterday. This is a study that we did that we published a couple of years ago in fibromyalgia where we looked at the relationship between symptoms and self-reported and objective measures of activity using actigraphy as the measure of physical activity in this particular case.

This is an actigraph that we've used in our group, Susan Murphy now uses this very actively, it's a cool actigraph because you can simultaneously get ecologically momentary assessments of symptoms in people

and measure their physical activity. So you have side by side in a given individual over fairly long periods of time ratings of whatever three subjective items you want to put into the watch. We're always collecting pain, sometimes we're collecting fatigue, stress, other levels, you can collect three different symptoms in this wristwatch. And then while people are wearing this you are constantly getting a three-dimensional actigraphy so you're getting assessment of what they're doing at the same point in time.

This is what an actogram looks like, and it's actually not rocket science, it's not nearly as hard as looking at a functional imaging thing or something, is that the bigger the little spike, the more people are doing. And these little spikes are little bursts of activity that people do during the course of the day. One of the things that I often say is that it seems like what we're really measuring is the ability that people can raise to the day to day demands of life. The things that they either want to do or that they need to do, and what you are really doing is capturing with a device like this the level at which people do do these different types of things.

So this particular study that we did, that Leo Kop was the lead author in, we looked at fibromyalgia patients, and we purposefully compared them to sedentary

controls because we were really trying to get people that might have comparable levels of real physical activity, but we knew that fibromyalgia patients always rate their physical function or physical activity as being very low. If you look at the SF36 PCS score, fibromyalgia patients have PCS scores akin to metastatic cancer or severe end stage HIV. They usually are two, two and a half standard deviations from below the mean on their physical function SF36 scores. So we didn't want to compare fibromyalgia patients to very active controls because we wouldn't be able to look at this sort of what is this really measuring at the end of the day.

And we had these people where their actigraphs for five consecutive days, I want to emphasize that you do need to sample weekends and weekdays, peoples' activity levels are often much different on weekends and on weekdays, so this five day period included the two weekend days and three weekdays in individuals.

So let me focus first on the left side of the screen -- on the right side of the screen, I didn't have any way to do this. So in this, again, what we found in this study was the physical component summary score, the SF36, was two standard deviations lower in fibromyalgia patients than in the controls. Daytime average physical

activity and nighttime average physical activity were almost identical as measured by the actigraph. You couldn't even get it to try to be more similar, they literally were almost identical with respect to their average levels of activity, but what was markedly different was their peak activity levels. The fibromyalgia patients had much less peak activity than did the patients without fibromyalgia, the sedentary.

And the other thing that was different is the variability, looking at either the standard error or the standard deviation, that the healthy controls had much more variability, the fibromyalgia patients were just always pretty active. They either just didn't see those peaks -- I mean so here are just the data sort of graphically represented, we're looking at the peak levels between the fibromyalgia patients and controls, fairly marked differences between the fibromyalgia patients and controls in peak activity levels, whereas there was no difference at all in their average activity levels. And again, these are people that are two standard deviations apart on the PCS score.

And these are two different actograms, the one on the left being a fibromyalgia patient, the one on the right being a control. And again, if you just look you see way

more spikes in the person on the right indicating more of these bouts of peak activity where they're actually doing things.

So based on that study we concluded that what really the self report measures of physical activity seem to be picking up on is peoples', again, abilities to meet the day to day demands of life, do things they want to do, do things they have to do, but it's these bursts that people seem to in some way be recording when they fill out a self report measure, it is clearly not their average or their mean activity level that they're recording.

So this is a study that Nat's group published recently, Trudeau was the lead author, Nat's the senior author, using very similar methods to see if this could be used in some way to increase assay sensitivity in the course of a trial in osteoarthritis. There were 47 people in this crossover trial.

Obviously Nat wasn't just looking at objective measures, he was looking at the classic measures, and what he found was that the WOMAC pain subscale was the most responsive of all five pain measures. Some of the other WOMAC measures, the BPI, were included in this particular scale.

And then they did some analyses looking at if you

combined their pain ratings and their activity ratings, it showed that a composite measure that required that someone have a 20 percent improvement in pain or a 10 percent improvement in activity had a better separation between the active treatment group and the placebo treatment group than did any of the measures singularly, that if you just looked at pain alone or activity alone.

He again found that peak activity level was again the biggest differences between the group, and one of the interesting things about peak activity level that he commented on in the article was there was no placebo effect. The placebo group had a reduction in their peak activity level during the context of the trial, the active treatment group had an increase in their peak activity level, so it was a nice measure in that it didn't, and again, I don't think we should even use the term placebo effect, I'm just parroting what they wrote in the article. I think this is probably more regression to the mean when we talk about a functional status measure, but nonetheless, there was no regression to the mean in the objective measure, whereas there was in the subjective measure of physical activity.

And he noted also that actigraphy was actually quite a bit more responsive than the WOMAC function scale,

again, largely because of the lack of the regression to the mean or the placebo effect in the actigraphy measure compared to the self-report measure.

So the next question is, should we expect a strong relationship between self-report and objective measures. Well so there's two other symptom domains that I know fairly well because we studied these in fibromyalgia, sleep and memory and cognition, and this poor relationship between objective measures and subjective measures is more the rule than the exception.

If you look at relationship between polysomnography and subjective measures of sleep it's even worse than the difference between actigraphy and subjective measures of physical function. If you look at sleep apnea patients, the R values range between, a whole bunch of PSG measures, there's a ton of PSG measures that are derived from PSGs studied, but the best correlation between any PSG measure and any self-report measure was .24. And the average correlation was actually .09, and again, so you see this both, and then in insomnia, it was a little bit better but still pretty abysmal, the R values ranged from .05 and .36.

Memory cognition is the same thing if you look at the difference between a subjective test of memory when you

ask people what they can and can't do vis-à-vis memory versus a performance based measure. There is a similarly abysmal correlation between the subjective measure and the objective measure when you look at memory versus these performance based tests.

The other thing I want to note is that these are not at all exclusive to groups of patients with diseases, these are, these same poor relationships are found in healthy controls. So it's not like once people get pain or once they get sleep problems they become poor reporters and the problem is in patient groups these poor relationships between objective and subjective measures are also seen in control groups that don't have any of these different diseases.

So given these differences between self report and objective measures, which is the right measure to use. Well, again, I do think if you follow this literature that actigraphy has become an extremely reliable and accurate measure of physical activity. So if we agreed for a minute that we're interested in physical activity as an outcome measure, which is different I think than functional status, I want to make that clear, this is a very reliable and accurate measure of physical activity. And I think there is overwhelming data, people have gone to great lengths to

video tape people and look at what they have really done versus what the actigraph measures, and again, the newer generation actigraphs are quite good at measuring, and any problem they have with respect to accuracy or reliability, the self-report measures have log scales more problems.

So if we're ok with the self-report measures of physical function, which we've been measuring for decades in pain, we have to be really, really okay with the reliability and the accuracy of these measures, these more objective measures.

Susan Murphy who works with our group has done some really interesting studies with actigraphy and I think that one of the, in addition to the discussion that we're having about the role of objective measures of physical activity in drug development, these measures are incredibly helpful in allowing us to look at groups of patients with the same disease and see markedly different endophenotypes within a given disease with respect to some of these contingencies between what's driving someone's dysfunction.

In Susan's studies of osteoarthritis you can see individuals where very clearly activity makes their OA worse, you can see another group of people where activity makes their OA pain a lot better. These are probably more people with centralized pain superimposed on nociceptive

pain. And you could see people whom pain doesn't really seem to be influencing activity, but fatigue is the major driving force with respect to what's really limiting that person's activity. So really using instruments like this in a more academic setting I think really allows us to get a very granular look at what is really happening in people's day to day lives and take a disease that we, again, I think we know all pain conditions now are very heterogeneous, but it really does allow you to start to look at different subgroups.

Certainly, when we are applying a rehab approach to these patients, the rehab approach to these three different subgroups, the pain gets better, pain gets worse, fatigue is the driving force rather than pain would be markedly different. In fact, that's what Susan is trying to do now is use some of this information, feed it back to patients and to integrate it into cognitive, behavioral, or physical therapy, and to treat these subgroups of people differently vis-à-vis what symptom you target, whether you target pain or fatigue, and whether you try to increase their activity because you see on their actigraphs that that makes them better or where you somehow temper their activity and get them to pace and stop overdoing it. Because that's usually what you see in the people that are

having activity lead to increases in pain, it's people that have the sort of habit to overdo it which we see in a lot of chronic pain cohorts.

So what I would say is that actigraphy is a very accurate and reliable measure of physical activity. I personally think that this is, it would be laudable for us in the pain field to move towards thinking of this as a primary outcome because quite frankly, as many people have noted already, in clinical practice this is what we do, this as we try to get our patients more active and more functional, and we're not targeting pain, we're really not, in clinical practice we're not having people focus on the pain we worry about sort of inordinately having people focus on the pain. And I think, in particular, as we move to the future where we start to do trials where we combine, I have always thought that the best thing to do is combine something that will relieve pain with a light version of a cognitive behavioral program so that people understand that when their pain gets better that their responsibility as a pain patient is to start functioning more, start going back and doing some of the things that they have stopped doing as their pain got worse and worse.

I think as we move into the future state of the pain field, we may be doing trials where we're combining

drug and nondrug therapies and really as people's pain is getting better from the drug, we're motivating them, incenting them to going back to some of the daily activities that they used to be able to do, then I think it would be particularly important to be focusing on a measure of physical activity as the primary outcome rather than as a measure of pain. Which again, I think most of us clinically don't think that's how we should be directing our patients.

So anyway, that's my view of the world and now we can talk about it as a group. Thank you.

Q&A and Group Discussion: Patient-reported vs. "Objective"

Assessments of Physical Function

MEASE: Why don't we segue right on into the panel and questions can be directed. Dave Williams, you're up, as well as Lee.

Bob.

DWORKIN: So, Dan, I'm not sure I heard you say what you think the patient-reported subjective measures are actually measuring since they're not measuring activity, what is it that the patient is reporting?

CLAUW: There are better correlations between the patient reported measures and these peak activity levels than there are the average activity. They're not at all capturing average activity, they're capturing peak --

DWORKIN: I thought you were going to say it's capturing catastrophizing, depression, something other than activity, but not --

CLAUW: I wouldn't say that, but -- yeah.

DWORKIN: Have you looked at that, like what are the correlations between self-reported physical function and kind of measures of other psychosocial characteristics?

CLAUW: No.

MEASE: Ann.

TAYLOR: There is some work around patients being

believed and, oh, sorry, people living with chronic pain being believed. And kind of reporting activity status or functional status that is trying to persuade clinicians that they actually have got pain, so they over report physical function problems in order to try and get clinicians to believe them. But that has a negative effect in the clinicians and think that they are over-egging their custard to actually -- and they might not be believed, so it has a bit of a negative effect then.

CLAUW: Nobody is, well, to Bob Dworkin's question, I think we can extrapolate from probably 30 years of self-report data that self-report measures of either functioning or pain related interference are correlated probably equally strongly I would say overall, but I think it's, you know, a meta-analysis or some overarching analysis would make sense with measures of both pain severity and mood measures like depression. So it's not either, I think there is structural discrimination, in people's minds, the phenomenology, they can think about these ideas and make discriminations but they are overlapping. So that, in fact, I have always thought about this as providing some evidence kind of for our concept of a multidimensional experience of chronic pain, right? But that it's back to my point that I've probably made way too

many times, I think it's important to think about these constructs as actually distinct and meaningful and really anchored in all of our personal phenomenology and the phenomenology of our patients. But trying to understand how these specific measures behave in relation to other measures, other relevant constructs, including objective measures like actigraphy, is quite important moving forward.

MALE VOICE: Yeah, just one comment, that's been looked at probably the best in looking at cognition, self-report cognition versus objective performance based cognition, and there is always a reasonable correlation, the self-report correlates better with measures of depression and anxiety. Because people, what you're really, people are worried about their memory problems and concerned about their memory problems. I would always say to my fibromyalgia patients that would come in and say, you know, I go to a grocery aisle and I forgot what I went there for, and I said well so do I, but I just don't worry about it. You know, I don't think that it means that I have some problem, but you do see that, in particular in the cognition area, the self-report measure is more strongly correlated actually with anxiety and depression than it is the performance based measure of cognition.

CLAUW: By the way, I think the other point in this context is how different self report measures of the same construct like pain severity behave in relation to one another. And so after 30-some years of experience with different measures of pain severity that we all accept in our world, the correlations are reliably higher than these correlations among different constructs, but they're not anywhere near one, right? They're reasonably higher among, you know, within the construct that we're trying to measure, but, for example, the McGill Pain Questionnaire, pain rating index versus a 0 to 10 numeric rating scale, isn't any higher than .5 or .6 I think over many studies.

DWORKIN: But it would be very troubling if a depression or catastrophizing measure correlated highly with the PCS in your fibromyalgia patients, than either the peak or the average actigraphy measure, that would really call into question that PCS is a measure of anything having to do with activity or function.

CLAUW: It doesn't because the PCS is sort of an orthogonal scale that is sort of specifically designed to not -- the MCS will capture that, not the PCS.

FEMALE VOICE: The PCS is really affected by the vitality domain in fibro, and actually we can even distinguish it in the lupus patients who have associated

fibromyalgia, they have a very different --

DWORKIN: My comment wasn't about the PCS, it was about the self report physical function. So if a depression or catastrophizing measure was more highly correlated with a patient reported measure of physical function/activity, then peak and average actigraphy, that would be to me profoundly troubling about the validity of the patient-reported measure of physical function. And it sounds like we have those data.

CLAUW: I'll just say I disagree. In the anxiety, for example, measures of psychophysical measures -- psychophysiologic measures of anxiety, behavioral observation measures of anxiety, self-report measures of anxiety was lived, the people that studied that have lived with knowing that those different measures of the same construct are reliably correlated with each other, related to each other, but the correlations aren't that high.

MALE VOICE: So I reflect on something slightly different although it's informed by all of this, which was your remarkable statement that you would use as a primary outcome in a pain trial actigraphy as the primary outcome. And I really wonder what we're thinking about when we're actually informing a stakeholder, be it a patient, a clinician, or a caregiver, or anybody, that because the

actigraphy, peak activity or something, it's not that it's not informative, but it doesn't tell you the patient's subjective perception of what they're suffering from, and that's what an analgesic is supposed to do. Yes, and it's also supposed to make them be able to do more things, but, in fact, subjectively sitting around and not related to dental procedures (laughter), but subjectively sitting around and thinking about your nature of how pain has affected you which are all of the things associated with a patient reported outcome. Every single aspect, Bob, of what you are talking about, is inherently inflected into the questions we ask in a PRO, the WOMAC, they all related to all of those things, catastrophizing, depression, anxiety, all of that issue, and therefore the sense of the ingredients associated with responsiveness is much richer and broader than just with activity and actigraphy. Not that it's not important, which is I think the crux of the discussion, I'm worried because a manuscript is going to come out of this and somebody is going to read a line that says maybe it should be considered as the primary outcome. But it's not actually what we're looking for. We're looking to know that the patient has less pain if we can think about the context of what that means and how to measure it better, not just increased physical activity,

which is not unimportant. It is another informative construct that helps us understand responsiveness.

I think it would be, maybe this is my previous regulatory hat, I just think it would be a very difficult concept for anyone to understand in labeling that a pain drug, the actigraphy got better, but what about all the other aspects of pain that are really important. So I am just a little concerned about that.

MALE VOICE: And again, part of it was being provocative, but part of it, what I really would consider it to be is a co-primary.

MEASE: I think I heard it as being part of a composite as opposed to a co-primary.

MALE VOICE: I seem to remember a two-year period when you were in a position where you were pushing for one of the three things that was needed to be a terrible, what we now know to be a terrible measure, which is a self report measure. So all I'm saying is let's elevate it up and measure it well and have it be a co-primary and see that we have comparable improvements in physical function.

MEASE: Sharon and then Ian.

FEMALE VOICE: I'm curious about the conversation with the patient who self-reports one thing and has actigraphy that says something else. So this is a little

bit separate from the clinical trial, but what does that look like and I mean one of the questions is always about, you know, people feeling like they're being taken seriously, we've discussed that quite a bit here, you know, convincing someone that they have enough pain to warrant therapy, not to have their therapy, you know, reduced or discontinued. To be believed that they can or cannot work. And here it's interesting because we think the actigraphy is a much better measure of an actual activity but in the same time the patient presumably, honestly believes they can't be more active or they're being as active as they can.

So has there been some experience in kind of discussing those discordant results?

CLAUW: Again, Susan Murphy's doing this in the context of trials now trying to actually feed this information back to patients and say this is really what you look like to us. And she says it really is interesting to see how people respond to it. When you show them their data, it is actually pretty powerful to see, you know, what, to lead to some sort of personal insights into, you know, whether pain made their function worse, whether it made it better. I don't think we're really challenging the voracity of what they're saying by self-report, we're just

literally, you know, using this as an additional data point to say, you know, we can actually measure this accurately. You know, Dave said this to me yesterday when we were talking about this, he said the whole reason, we all love PROs and we all like the patient focus, but the whole reason that we went into PROs is because none of these things could be measured more objectively. And when we finally have something that can be measured more objectively, why don't we, why aren't we willing to use it?

FEMALE VOICE: I was just kind of curious what the reaction was, I'm not saying we should or shouldn't or one is better. But, you know, I can imagine a couple of scenarios. One is, well I'll show you can't do anything, and then they don't move, and they keep wherever that actigraphy is strapped to as still as possible because they're afraid of not being believed. I can't go back to work, I can't possibly imagine going back to work, if this thing says I can go back to work, I better be careful not to move it. Or wow, I guess I am better and then a positive outcome.

In the context of a clinical trial it may not be so relevant because presumably they're not getting real time feedback on their actigraphy outcomes, so I'm not quite so worried about it in that context. But just in

terms of the general what kind of behavior will it elicit and how do we account for that, you know, in the setting of the perspective of a patient. Maybe we just don't know that yet. You know, the same way we said someone is sitting there and complaining of a 10 out of 10 pain, but they look perfectly comfortable, you know, what is the, there's a whole psychology that underlies that.

MEASE: Ian.

IAN GILRON: I just wanted to continue the response to the provocation that there should be a primary outcome measure. And I just wonder whether in some of our discussions there's a bit of blurring that as clinicians we think we want to treat the patient holistically and look at their goals and want them to be engaged and function despite their pain versus clinical trials which are very focused and presumably the treatment that we're investigating we have some, you know, pre-study hypothesis of how it will work.

So I mean if we think that this analgesic, whether it reduces pain or not, somehow is more likely to have an effect on activity, then that may be appropriate. And it just, I mean we're at a bit of a crossroads here between, I think it depends on the disease condition as well, but, you know, I think if we discover the next

penicillin for pain we want to know that we can actually find an analgesic that will reduce pain to 1 or 2 in most patients, which is why it should be a primary outcome.

So I think the fact that we're saying we shouldn't focus on pain, that's in our current setting where most treatments don't do that much for most people, and so that I think we should keep those distinctions in mind for clinical trials.

HERTZ: But also, can I comment on that, sorry, this is Sharon Hertz. But that's an incredibly important point, because what if the reason the activity has gone up while the pain may or may not be changing quite so much, is because the drug is less sedating?

MALE VOICE: Right, that's exactly right.

HERTZ: Or, you know, some other separate thing. What if it actually is activating and are we going to call it an analgesic or are we going to call it a treatment for fibro? So that's why we are -- that's why I, because this my personal opinion, right, think that multi-domain combined outcomes are very challenging because it can be very difficult to tease this out.

CLAUW: So just let me respond to these two. I think it only should be multi-domain, I don't think that it should be a primary without pain being included. But I'll

give you the other scenario which we see I think fairly often in clinical practice, and that is people where you do give them an effective analgesic and what they do is they raise their functional status level to get to the same level of pain, because they can tolerate a certain level of pain. And right now in our trials that's a failure, that person's a failure. And that is not an uncommon scenario in treating chronic pain patients, where what you literally do is don't really budge their pain level, everyone has their own like internal set point for what they're able to tolerate, but you move them up to a different plane of functional status. And right now we're calling all those people treatment failures --

HERTZ: But at some point the pain did go down, because that's how they realized they could do more. And when we ask them to average or rate pain over a period of time by recall, we're losing that, and that may be why they're failing in a pain measure. If we did real time pain scores throughout a day and looked at a profile we might be able to capture that and then consider them successes from the perspective of the analgesic outcome, as well as successes from the perspective of the functional outcome.

MALE VOICE: So you're suggesting that the concept of average pain in the context of separation from

the function, may be giving us an erroneous look into the pain response, but also, because you're doing a multidimensional response measure, you're getting different aspects of responsiveness which may be actually more informative than less informative, given what you've just said before, Sharon.

So I'm a little misunderstanding because you mentioned that multidimensional can be confusing, but on the other hand, that's just because we may not be asking the pain question correctly. So could you talk about

HERTZ: No, that last piece is not what I meant to say if that's how it sounded. What I'm saying is there are many reasons why a multidimensional outcome may improve, it could improve because there is less pain, and therefore more function, it could improve because the drug is less sedating, and it could improve because the drug is activating.

MALE VOICE: And what's wrong with a drug that decreases the pain measure, which might be less sedating, thus they get better, based on this multidimensional function, there's nothing wrong with that? Also, if it's activating, but it also decreases your pain score, what's wrong with that?

HERTZ: Well, because we don't know if it's also

decreasing the pain score if we're using the composite. So my point is all of these measures are important and how we handle them in a clinical trial is important, but we need to understand each piece of it before we bundle it. So if someone's pain is steady but they're able to do more, that's potentially a good thing, but I wouldn't call it an analgesic, because somebody would then have the expectation that it may improve pain when, in fact, what it does is help somebody improve function independent of that.

Similarly, if somebody has an analgesic that actually reduces their pain and that's the reason their function improves, we kind of want to know that too, because then if I have somebody with minimal pain but trouble with their function for other reasons, for instance, there's more than pain that I think limits a fibro patient, right, there's fatigue and other things. Well I want to know that because that may not work in that patient.

MEASE: I've got everybody who has raised their hand written down so just so as not to get the shoulder dysfunction that Jas mentioned earlier, so he asked to pull out his needle. Dorcas, a while ago --

BEATON: But I have a self-report measure of shoulder function so I'd be happy. I think we're sort of

evolving to one of the points that I wanted to raise right from when I heard you speaking is we have to remember that we might be wanting to measure different windows or different views, and we need to do each of them well, and we don't want to throw the baby out with the bathwater. So when we're saying that we're here to talk about physical function and now we've distinguished physical activity indicators from physical function, that's good, but we don't have to then sort of say, and what I heard saying that because the correlations were low we had poor reporters, that the reporting of physical function was poor, and that's why the correlations were low. Or the difference in the PCS in the fibromyalgia patients compared to the normal patients, normal inactive patients or people. But if you look at how that score was created it was done in general population norms where pain is a very high driver of that PCS weight for the domains that went into the PCS score. So it doesn't surprise me at all that on the PCS the fibromyalgia people are going to have more. And that vitality then is a distinguishing thing about what changes in people with fibromyalgia. But compared to a general population it's the pain scores that are very heavily weighted in that.

FEMALE VOICE: All the domains go into PCS.

BEATON: But if you look at the coefficients, the beta weights, it's much stronger for pain. Because in the general population that's a driver.

MEASE: So Ernest is up next.

CHOY: So I think one thing we agree is that we have to be very careful with our wording, and what we are talking about with actigraphy is a new domain called physical activity, and we don't confuse it with physical functioning going forward, it's really important we get our wording correct.

I guess one of the points about pain reduction which we talk about as a primary outcome measure versus what we tend to do in clinical practice, is that we always tell our patient that the pain won't go away completely, what we were trying to do is make the pain more manageable. And that may translate into greater activity, which in turn actually can help the patient to manage their pain. So greater activity can help the patient to cope with the pain better. So there is a full circle in this that greater activity may in the long-term help the patient to manage their pain a lot better. So it is not a straightforward relationship that we shall assume.

MEASE: Okay. Ajay.

WASAN: Yes. So actually Veeraindar and I were

talking about this exact issue last night over dinner, and I think the fundamental tension here is that whether you approach it as treating chronic pain as a symptom, which is a classic analgesic approach, okay, versus treating chronic pain as a disease, and having a disease modifying approach, which, of course, you would want to pursue multiple primary endpoints.

And so that's what I'm hearing back and forth kind of in the room, and I think that the science is strong enough now that we can approach that there is a need to pursue multiple primary endpoints such as actigraphy because we want to take the approach of treating chronic pain as a disease. And I think that really is a big part of what the discrepancy is here and we need to sort of take a stand on that as part of what we're going to put out as a work product.

MEASE: David.

DAVID HADDOX: I'd like to, it's hard to look at Sharon and talk into the mic but I want to pick up on something she said. If you use actigraphy as a measure and you had a drug that, going back to your mental status examination, induced agitation or induced akathisia, or induced a tremor, your activity ratings would go up with no analgesia. The other thing that, another example I think

that maybe mirrors, and Sharon, if I'm misquoting you or misunderstanding speak up, if you had a drug that, in fact, was not an analgesic but improved pain tolerance, you might see an increase in activity with no change in pain, so unlike the example, Dan, you were talking about which we all see clinically, where you do seem to get some transient dip in pain and then they move their activity up to where they can tolerate that, if you had a drug that somehow magically addressed pain tolerance without truly being an analgesic, you might also see that functional change and so you wouldn't want to call that functional change, that's an analgesic.

MALE VOICE: Why is that not an analgesic, what do opioids do?

HADDOX: Because it doesn't decrease pain.

MALE VOICE: But doesn't an opioid increase your tolerance of the pain that you have --

MALE VOICE: A lot of drugs, that's how they are working.

MALE VOICE: Right, they increase your ability to deal with whatever the pain is that you have, and they alter your perception of pain. So it's interesting you said that because we probably will be getting into the pharmacologic development programs of drugs that alter

perception and we have used and thrown around, we all do, terminology, which may actually not be accurate. So I just wondered what it is that altering your pain tolerance is not an analgesic?

HADDOX: They're very distinct concepts about, you know, pain threshold and pain tolerance. And pain tolerance is very elastic, you can have encouraged pain tolerance, all sorts of things affect that. And so if you had something that let's say was in essence an antidepressant, but affected the affective component of the encouraged pain tolerance --

MALE VOICE: Tricyclic antidepressants.

HADDOX: Well, maybe, I would argue that too, but we don't want to waste time on that, but I'm just trying to, theoretically, the point I'm trying to make is that I understand the interest from a regulatory perspective in having multiple measures but not combining them into a composite, that's to my point.

MEASE: Penney.

PENNEY COWAN: I find this discussion very interesting but I think the thing that no one has really talked about are these are just people, and they don't just bring measures with them, they bring who they are, what they do, maybe their function is improved because somebody

came home, one of their children came home from college and helped them lift that box instead of lifting it on their own. I mean there's so many other factors, we've talked about all the environmental factors, and I can remember when I would be sitting there and they'd say, oh, well this and this and this and I thought but you don't really know who I am, and you don't know what I want. So I think that there is a huge component of the individual in each of these measures that's going to make it really difficult to say this is the way it has to be because there's so many other factors.

And I just want you to think about it, they're people with a variety of different things happening in their lives outside of when they talk to you, and my question is do you really talk to them and say, okay, your activity has gone up or your pain, why? I mean do you ask them that or is it just you look at the measures, or do you actually have a discussion about it. Because I mean we're all about, I mean we hear patient centered care and all of that these days, do they do that?

CLAUW: Well to be clear, I don't use actigraphy in clinical practice, I don't know of anyone that's using it in clinical practice, we're using it in a research setting and we're testing whether if you incorporate that

into a research setting and you feed back this information to a real person that's living a real life, whether that helps them gain insights into how they might be able to modify their behavior, i.e. their activity level, to improve their pain.

So we're not using it in clinical practice, we're testing whether it might be able to be used therapeutically at some point in the future by just feeding this information back to people.

COWAN: And I understand that piece, but I'm talking about the whole conversation today, not just that particular thing, is, you know, we keep talking between clinical trial and clinical practice and I wonder, you know, people are going to read these papers and they're going to then apply this to their clinical practice. I mean the papers are all over and so it, to me it seems that we have to be careful that we at least put that human component in there, that it's important to think about, you know, what is their environment. And we've heard environment before and talking to people just to really have a better understanding of what those numbers mean. I have a problem with just putting numbers on people.

MEASE: Gary.

GARY WALCO: I want to go back to what Ajay said

because I think it was very insightful in that we know there's an extremely intimate relationship between function and patients or peoples' subjective pain experience. I think Fordai (phonetic) showed that 40 years ago. There was a recent paper out of Cincinnati with their functional program that focused exclusively on pain rehabilitation for function that then showed subsequent decreases in subjective pain ratings. So it is fairly clear that those are intimately related and I fear that in some of this discussion we still are holding onto the notion that eradicating subjective pain is the goal in treating chronic pain. And that is certainly the expectations that our patients have who come in.

I know that there are some drugs that have gotten approval showing a 30 percent reduction in pain in a small, given the number needed to treat, that was quite unimpressive. And so you look and you say, well if a patient drops their pain from a 7 to a 5, are we calling that a therapeutic success? And I really think the whole idea that you got at, we're not just shooting for that immediate short-term shift, but that chronic pain is something that goes on over time, and if we're going to do a clinical trial that's going to look at the immediate effect of an intervention in a small window period, we may

be completely missing the boat. And that's where I think we do need to be asking multiple questions and looking at multiple factors and how they come together for a given individual is going to be quite variable.

MEASE: John Markman, you had your hand up a while ago, and then Laurie, you follow John. But, John, if you want to yield the floor, that's fine.

JOHN MARKMAN: I'm going to take it, okay. So I think, I just want to pick up on Ajay and Gary's point, I think the importance of what Lee is saying I think about marrying or putting together in some fashion the subjective report and the objective measure and why it's so important. I mean to, John Farrar started out the meeting by saying he thinks that pain is a disease largely of the nervous system, and whether you belong to the localization congregation that puts in the posterior insular, wherever else, to the extent that chronic pain is a disease and a disease of the nervous system, it's a disease of rating and it affects the way you rate pain, and it's all connected.

So I think the bottom line is that, how you rate and how you subjectively feel the pain and what it's tolerability and how the objective measure is, are of a piece, and you've got to -- and looking at them together is useful for the reason I think that Laurie began to get to

in the panel. Which is that ultimately I think in order to get these medications to patients we have to have as robust and nuanced a sense of efficacy as we can get. Because ultimately that gets things paid for. And it helps give us an argument when we try to actually use the medications. And that matters so much more today.

So to me, to have a co-primary or a secondary, however you want to put that, where you put that objective measure of function with that self report, is just going to help get it paid for and help it get to patients. And I think at the end of the day, that's what I hope comes out of this, is a more nuanced sense of efficacy that makes the case for using the drugs more compelling.

MEASE: Laurie.

BURKE: I want to again make the case for conceptual clarity here, and I think that it's important to think carefully about what actigraphy is measuring. And I heard, and I'm going to back to both David's comment and Penney's comment, that it really is not measuring physical activity in daily life, it's measuring movement or acceleration, or something physical like that. So we've talked about lots of the reasons why you can't just assume that this measure of movement on a wrist or a hip equals physical activity in daily life. They can take it off, they

can stop moving on purpose, they can whatever it is. So that's the whole idea of what's the meaning health aspect, Penney, to a patient? Meaningful health aspect is physical activity in daily life and what that means in their daily life, and the actual measure then has to in some way be demonstrated to have a relationship to that. And we do, and this is the problem with performance measures, performance outcomes that we're starting to come up with standards for how to evaluate them is that there is this immediate jump to assume that it measures this meaningful health aspect when they don't.

And so I think that will help the conceptual clarity of these discussions here and reduce the anxiety on the part of the people with pain that you are all of a sudden going to turn this into something that's not meaningful in their lives.

CLAUW: So I'm going to push back. I showed the actograms and Dr. Patel yesterday showed the data as well, it does measure what people are, this isn't like a cognitive test or a polysomnogram that is really a surrogate measure and you're not sure what it's measuring, it is measuring physical activity, what people are doing, and I think people do find that meaningful. And again, the tenor of the conversation is bothering me a little bit

because I think any new thing that we move forward as being a potentially better primary or co-primary, you can shoot holes in it, but look what we have now, we have a subjective pain measure, it couldn't get any worse than that (laughter). That's what we have to compare it to. So is there a drug -- let me finish --

BURKE: I'm going to stop you right there because we have relatively objective ways of measuring self-report items. So it doesn't get --

CLAUW: See, I totally disagree, I think that self-report measures are wonderful and everything but at the end of the day they always have --

BURKE: Yes, but I also am not demeaning your position that this is a really helpful contribution to the measurement of health status in certain patient populations. But to say that you take actigraphy and throw it into a clinical trial you have an automatically valid measure of what you're trying to measure in every context of use is going just a bit too far, and I think that this conceptual clarity issue is important to think about. So it measures, what it actually measures is movement, so someone with tremor, that's another context of use you have to think about. I've heard some actigraphs measure movement when someone is in a car or a bus, you see, and so there is

the validity issue that has to be considered in every context and that's all I'm saying.

MEASE: Jim.

JAMES WITTER: I'd just like to pick up on what Penney had said before about the patient-provider relationship and some examples that are starting to emerge from that P word, if there's a word, PROMIS, in the sense of you know that what we're trying to do is create a standard, to establish a common language. And one of the things we've been doing is to look not only in clinical trials but also in clinical care and bring the same measures into the clinic. A lot of the folks -- I shouldn't say a lot, several of the folks that are intimately involved in PROMIS are involved with PCORI. And so we're starting to learn about examples of the impact, for example, utilizing CAT. When you see a patient and you take these measures and the CAT spits out the answer right away, then you go over this with the patient, what we're finding is that it enriches the conversation between the provider and the patient. And so it makes the overall experience more patient centered and you come up with a better outcome that hasn't quite broken the literature yet, but I think that's what, one of the things we're starting to learn about this. So I just thought I'd add that into

the mix here.

MEASE: Ernest.

CHOY: I just want to add the piece about validity of actigraphy and the argument that on the face of it what actigraphy is measuring is movement. While it is, but actually what you presented is not movement that differentiates active and placebo treatment, what you' are differentially saying is that peak activity differentiate from the two, and it's really what that piece say and what translated to the patient it may be to do things with increased vigor, it may be related to new activities that the patient is doing that they normally couldn't do and he is trying to translate what is picked up by the actigraphy the be meaningful improvement of the patient that we need to tie up, to understand the real impact of that.

But I don't think it's just movement because if you have tremor because of Parkinson's you'd have tremor all the time and while you have some variation it will be within the variation of that instrument, it wouldn't --

CLAUW: It's not going to measure peak, average might be affected by a tremor but peak isn't going to be affected b a tremor or by, or quite frankly by sitting in a bus. If we're measuring peak, it's not, these little nuances are not going to really be picked up.

BURKE: Well and there's another conceptual clarity issue, we're not just using accelerometry, we're using accelerometry in a certain way to produce a certain score and that's what we're talking about. I think that that's --

MEASE: David.

WILLIAMS: I would agree with Dr. Choy that tremor would not affect peak activity, but I bet you droperidol would, because it induces akathisia, which is a subjective sense to move. They are not comfortable, if you've ever seen a person who is undergoing akathisia they are not at all comfortable, they are intensely uncomfortable but they are moving around. And I just, I'm not criticizing actigraphy, I think it's a really good idea, just let's make sure we understand what it's actually telling us and control for those things that might mislead us in how we interpret it. That's really my point.

MEASE: This has been a great discussion. We thank the panel very much. Why don't we come back at 10:45 and we'll continue on with discussion. Thank you.

(break taken)

**Group Discussion I: Considerations for the Assessment of
Physical Function in Analgesic Clinical Trials Using
Patient-reported Outcome Measures**

DWORKIN: Okay, we're going to start. Well, so before starting to talk about physical function, it seems to me and I thought we'd just spend a couple of minutes seeing if this is true, that there is a sort of tacit consensus in the room that the topic for the next IMMPACT meeting and we don't know when that will be, should be to revisit how do we assess pain in pain clinical trials. Given the discussion yesterday and today and the apparent lack of any real agreement on how do we assess pain, that seems like a worthy and obviously critically important topic for an IMMPACT meeting. Is there anyone who thinks that that's not a reasonable idea for the next IMMPACT meeting?

I mean I've got a list, Dennis and I have a list of other ideas that have been suggested, but boy, when we looked at it last night kind of how do we assess pain in a pain clinical trial really seemed to trump everything else on the list, which included things like effectiveness trials, trials of mechanism based treatment, visceral pain trials, all of those things are important but at least it seemed to the two of us that how do we assess pain in the

pain clinical trials is at the top of that list.

Does anyone want to disagree with that consensus that I hear, that we hear for the next IMMPACT meeting?
Dan Clauw.

CLAUW: No, I'm not at all challenging it, I'm just saying that I think it would be a wonderful topic, I would just suggest that you might expand the participants to include functional brain imagers, and people like that, not that it would ever be used in a trial, not at all, but what we've learned using functional imaging and things like that about pain perception to inform the discussion about, that's all I'm saying is I would have a broader group of people that might be able to be forward looking and saying this is what we now know about the neurobiology of pain.

DWORKIN: Okay, so Dan agrees with the consensus but wants us to invite Irene Tracey and I'm good with that. Any other comments on the apparent consensus? Okay, we've made a decision already and it's not even lunchtime on the second day, which is that the next IMMPACT meeting is going to revisit something like, you know, how do we optimize the validity assay sensitivity of our pain assessments, which is typically, though maybe not always, the primary endpoint in a pain clinical trial. So thank you all very much.

TURK: And if our OMERACT colleagues or friends

are interested in that, maybe there is a future opportunity to consider also --

But only if they are willing to have dinner with the IMMPACT people (laughter).

MALE VOICE: The only question I would have with that is given the discussion this morning, do you want to refine it?

DWORKIN: It sounds like chronic because I think there would probably be more of a consensus about our current approaches to assessing acute pain in a three-day bunionectomy trial. What I think we heard over the last day and a half is there is much less agreement about chronic --

MALE VOICE: And that also is going to be of more pertinence for the OMERACT group.

DWORKIN: Yes. Dorcas.

BEATON: Would it be also looking at drug interventions or would it be looking at potentially rehab interventions?

DWORKIN: Okay, so this is important and Dennis was going to respond to this question when it came up. So Dennis, I call on you to answer Dorcas' question.

TURK: Rather than make it as a question I'd make it as a general comment because it's relevant for this

meeting as well as for any subsequent meeting, is to keep in mind that although we periodically slide into talking about drug trials and regulatory issues, that the purpose of these meetings, the purpose of all the impact meetings, has always been to try to address issues that are relevant regardless of what the nature of the trial is, as long as it was related to clinical pain of some type, or it could be pain, it doesn't have to be clinical, because we could even consider laboratory studies and have.

But the idea is that as we think about the kinds of things, the kinds of issues that are relevant, they should have broad relevance, there may be some things that are unique to the regulatory area and that's fine, there may be some things that are unique to drug studies and that's fine, but these should not be viewed in any way as if acupuncture studies, physical therapy, rehabilitation studies, all of those would be equally relevant. In fact, on the list that Bob mentioned about the topics that we've had for sometime in the future, we've looked into the possibility of having rehabilitation studies as a particular focus to make sure we get an appropriate emphasis.

So I think, and now let's go into psychological treatments, we also talked about having meetings about

psychological interventions. Dan mentioned the combination of treatments which is also on the list, is to start looking at how can you improve the design, the conduct of studies that are combination treatments.

So the whole purpose of these meetings is to be aware, this meeting as well as a subsequent meeting or any future meetings, is that they're improving the quality of clinical studies with the goal to get better answers to questions, develop better treatments, to provide better care at some point, to people who have pain problems, pain patients.

BEATON: Will that subsequent meeting be focused on chronic pain patient subsets and identifying critical subsets, or not?

DWORKIN: I would assume we'd start off with chronic pain, in general, both musculoskeletal and neuropathic, and visceral, but we should certainly have discussion about whether the different contexts of use of different types of chronic pain might require different approaches to assessing pain.

TURK: Let me add to this, and we can talk about this now and that's perfect to do, but all of the IMMPACT meetings have always had a steering committee who takes back the comments from the group, and I assume when we

planned this with OMERACT for this meeting we tried to work together to come up with a program that covered the range of topics, tried to identify speakers who were knowledgeable about those topics, tried to identify background articles as you all have for this meeting to sort of bring people up to speed. So that regardless of what we initially decide here, that will all be discussed by steering committees which will then try to take into consideration all the issues and try to flesh out what we think we can do in a day and three-quarter meeting.

DWORKIN: So unless anyone has anything else to say about pain intensity, let's move on to the topic of this meeting. And so what we want to do, what we thought would be very valuable to do before the lunch break is to go back to the point that Laurie made today, Ashley and Elektra made yesterday, which is really the place to start, and I think we all agree with this, is kind of what is the concept of interest, to use their term, that we're interested in here. And in thinking about that this morning it occurred to a couple of us that the article, the manuscript that Ann and Kristine will be drafting, the very first question they're going to have to consider and we should take a vote on this, are we talking about physical function or physical functioning, and I realized I didn't

know. And maybe it doesn't make a difference, but let's just get another, let's have two consensuses occurring very quickly.

Does anyone want to make an argument for why one of those is the better term than the other, function versus functioning? Dan?

CLAUW: So I would make an argument that it would be good for them to outline the difference between the two. I've been using these measures for 20 years and until I did this talk, I didn't really look carefully enough at the measures to know that there's differences between physical function measures and functional status measure.

DWORKIN: So I said function and functioning, are those synonymous, because I think that's different than functional status.

MALE VOICE: I think this is an example, and I wrote down a whole bunch of terms that we've been using at this particular meeting, which would be valuable to include in some definition --

DWORKIN: I'm arguing that function and functioning may be synonyms and we can just have a show of hands for preference, unless there's someone in the room who knows of an actual content difference between physical function and physical functioning as terms. Everyone is

looking like they're about to fall asleep, so I'm assuming that people think this is a boring decision, so I will suggest --

TURK: Dan Carr has his hand up.

MEASE: Dan.

DANIEL CARR: Just listening to the words involved, to me, functioning sounds more like what you actually have done, and function sounds like how can I characterize you. And example might be if someone comes in and we're interested in their pulmonary function, we could do a vital capacity where we urge them to deeply exhale and breathe in as much as they can and then exhale that we measure it. But that's not going to tell me how deep each breath was that that person drew in the prior month because they never were asked to do that.

So I would, to me, functioning, I'm not saying to discard either but I'm trying to address your question about are there distinctions and functioning to me sounds like what was actually done, and function sounds slightly more abstract to characterize what the person's physiology is.

DWORKIN: So the other argument I guess in favor of functioning would be my recollection is that's the terms we used in previous IMMPACT articles. So is there a

consensus that when Kristine and Ann draft the manuscript, that they'll be using the term physical functioning to describe what we have spent the last day and a half talking about?

John Farrar just threw his hands up like that, and so I'm going to take that as a consensus that everyone in the room agrees (laughter) that it's physical functioning. Thank you on behalf of everybody, John, thank you. So now we're going to move from --

TURK: What Bob has done is done the low hanging fruit. He took the easiest things first, and now that you're primed we can move into something more difficult.

DWORKIN: So Dave Haddox is not going to let this go by --

HADDOX: Not that easy. We had a lot of discussion about physical activity as opposed to physical function --

DWORKIN: We're going to get to that, that's next. So Laurie Burke is unhappy with me right now a little bit because she thinks I skipped a level asking you whether it's physical function or physical functioning. She said to me during the coffee break that that's the second question you should ask, the first is is this manuscript meeting, whichever way you view it, about function or is it about

functioning, or is it about physical functioning? Because I think we're, correct me if I'm wrong, where Laurie was coming from is that we have a kind of presentation on social participation, and on caregiver burden, and on work productivity performance participation, and are those domains, social, work, and caregiver burden, part of physical functioning. And Laurie was saying we should talk about this before we talk about function versus functioning and I have --.

MALE VOICE: I would just say going back, because I'm old and I remember the early IMMPACT core domains, we had emotional functioning and it was a separate --

MALE VOICE: Yes.

MALE VOICE: So consistent with that kind of framework, I think we want to stay in the realm of functioning that's not emotional and then break it down. I think we called it physical functioning.

MALE VOICE: We did.

DWORKIN: All right, so what do you and the gentleman on your right think about social --

MALE VOICE: I think it can be accommodated. If we can define physical functioning broadly, incorporating social role functioning, et cetera, then we can live within that.

MALE VOICE: What about just saying
functioning?

BURKE: And you're saying emotional functioning?

MALE VOICE: Emotional is separate.

BURKE: And social functioning is within physical
functioning, is that what you're saying?

MALE VOICE: Yes. I mean that's not maybe the
best, but again, we already did that. The first IMMPACT
defined those core domains and labeled it as such.

MALE VOICE: Dave.

MALE VOICE: I would argue that we should
probably have functioning as the top level, but then I
think physical functioning is another category, but then
social functioning incorporates both the emotional
functioning and physical functioning, you know, so it's
almost like another category, you almost have like physical
functioning, social functioning, and emotional functioning
but --

MALE VOICE: My recollection of the earlier
IMMPACT meeting is we separated out social, but said we
knew so little about the domain of social functioning we
couldn't recommend it for assessment.

So what about work, is work part of physical?
John?

MALE VOICE: Before leaving the functioning issue, I think one of the ways of thinking about this is that there are models for quality of life that include emotional function, social function, and they are all defined as functions, right? And social is kept separate from emotional, it's kept separate from physical, and then we talk about spiritual functioning, if you like or believe, and then there is health.

To migrate too far away from that here I think risks making this not relevant. And so I would argue that they're all very important, but that what we're talking about here, at least from my perspective, what I thought was the focus was primarily on physical function and I completely agree with Dave that social functioning involves both physical and emotional function and your ability to do other things involves other pieces, there is clearly overlap. So I think we would keep those separate.

MALE VOICE: We have our third consensus, pain for the next IMMPACT meeting, functioning rather than function, and that we have three, that there is functioning at the highest level and that has at least three components, emotional functioning, physical functioning and social functioning. And that this meeting has mostly, with a little bit of an exception, focused on physical

functioning.

MALE VOICE: But those are not orthogonal so I think that point is important to make.

MALE VOICE: So then my question was what about work, is work, the work things we heard about yesterday, is that part of physical or part of social?

MALE VOICE: That's where I think things get a little bit cloudy, which is that I think it's clearly part of physical functioning, you need to have enough physical function to do the job, but it is also part of being involved in the social engagement. So that is why it gets complicated to think about.

DWORKIN: So Dorcas, Laurie, then Gary.

BEATON: And this is where I think my model, similar to the ICF, not that we have to take that or one of the established models that would put work, life, raw functioning as sort of almost an integrative, because although we might see physical functioning as a key component of work, if you don't have social skills to negotiate a modification, if you don't have the cognitive ability to concentrate on your work, you are going to get your hand caught in the machine.

So there's all, it's almost like an integrative, another layer, another more integrated layer of

functioning, so I could see it at a different level, potentially the social participation and the work participation.

DWORKIN: Laurie.

BURKE: Well if you recall the diagram that Ashley presented yesterday, this is the reason we developed this diagram is because there is the core signs and symptoms, then there is the impact of those on subsequently more distal things, and productivity and health related quality of life are out here. They are directly, they provide direct evidence of treatment benefit, they're incredibly important, no one is saying they're not, but they are, what we should add to that diagram, but they are really important, no one is saying they're not, but they are, what we should add to that diagram is at each progressively more distal concept layer, there are more and more other things that have an impact on those concepts. So I would suggest that this discussion limit itself to physical functioning, determine what those core concepts are within a particular context of use, and maintain that stream of physical functioning acknowledging all of these other things that we're not going to talk about because we can't possibly be complete in that regard.

DWORKIN: Gary.

WALCO: I actually was going to say something very similar in that I think that clinically working with kids, the endpoint that we shoot for is return to school. And I think with adults, the endpoint you often shoot for is return to work. And that's not simple. I mean you've got to take the other various factors into account. So I think that we can say that and then put it aside.

The other issue I wanted to raise is the area that was covered in discussion, and I don't know where you plan to fit this in but it's also caregiver burden, because I think that's huge and it's not just so directly related to physical functioning.

DWORKIN: So those three areas, work, social and caregiver, are impacted by physical function, it seems we have agreement, are all importantly impacted by physical function, but they are impacted by emotional function and as Laurie says, they're kind of distal to what the focus here has been, which is physical functioning.

Jas.

SINGH: So Laurie, where is the empiric evidence that the so called distal are distal? So I'll give you an example, when my daughter gets sick, she's six years old, when she gets a fever from a cold, she has a fever which is proximal, but she also takes off from school which is

distal, but I don't think it's distal. I think maybe when we think about chronic conditions we think that you need to accumulate things over time to start missing work and have health care quality of life deficit, but I would argue that I don't think there is empiric evidence to say these things are distal. People start missing work quite frequently, quite early.

BURKE: Okay, I think the confusion here is that you are mixing up the definition of proximal versus distal with direct versus indirect. And that's why we define both. Okay, so direct/indirect is the biomarker over here that really doesn't have any implication on how that patient feels and functions today. My cholesterol, I'm telling you, is high, but you know what, I feel pretty good. So that cholesterol is distal to how I feel. Okay, that's just the way we define distal, if you don't like distal, we don't have to use distal. I'm sorry, direct/indirect, now I'm getting them mixed up.

Okay, so there's the difference between direct/indirect and proximal/distal. Those are just the terms we use, you can use something else. There is the point with proximal/distal is not that staying home from school is less relevant than having a runny nose and a cough, they're both very important direct measures of how

the patient is feeling and functioning. The difference is there's all these other things that have an impact on whether this kid stays home from school. As a former teacher, I can tell you there's a lot of kids that come to school with a runny nose and sore throat, okay, so that's, it's not just the fact that they have a runny nose or sore throat, that has an impact on missing school.

DWORKIN: I think Gary was next and then Bob. Did I call on you already?

WALCO: You called on me already.

DWORKIN: Okay, Bob.

MALE VOICE: So that's if you buy into a biomedical reductionist model of pain, which I don't think we do. The predominant model of the field is a multidimensional biopsychosocial model and we should be clear about that and not stray from that view. So the idea you said, it's a blur to me, the biomarker, or even the idea that being sick, actually sick is an illness term, not disease, I think, you know, historical roots, and somebody labeling themselves in a fever as something is different than somebody else that doesn't experience the fever is anything that interferes with their functioning and so forth and so on, right?

So anyway, I would just, I think we stray far if

we don't agree that we're talking within a biopsychosocial framework.

BURKE: I don't understand that, honestly I don't, and so I think that how is, what framework did you say I'm operating under?

MALE VOICE: A biomedical reductionist --

BURKE: No.

MALE VOICE: Because you started with the idea of a biomarker, I guess that was --

BURKE: Well that's not at all within this proximal distal framework, biomarker is not even on the scale. So I think that's what I wanted to make sure that we're communicating here. Biomarker doesn't show up on this proximal/distal scale at all. Biomarker is in a whole different place, that's an indirect assessment of treatment benefit.

DWORKIN: So Dorcas and then Dan Carr.

BEATON: So would you say then that if we use this child with a fever that at the same time the child with the fever is going to have an impact on their feelings, their sense of pain, uncomfortableness, their activities during the day and their ability to go to school are happening as an impact of the fever?

BURKE: Right.

BEATON: And so we should try to get measures or be ready to quantify those different things and that there's different factors that have an influence on why that may or may not be happening.

BURKE: Right, and there's different audiences that are more or less interested at different layers of this proximal/distal. The parent is very interested in whether the kid stays home or not, the clinical trial reviewer at FDA is very interested on the runny nose and the cough.

MALE VOICE: So Laurie, is this over simplified, what I'm thinking is that we seem to have agreed if we're interested in functioning, there are three primary domains, emotional, social, and physical, and that then there are these other aspects of life like work, going to school, that are a function of complex combinations of the three primary domains. Like so whether I go to work early or late on Monday morning is probably a function of my physical functioning and my emotional functioning and my social functioning. And it doesn't, you know, the time I arrive at work on Monday can't be reduced to one of the three primary domains. Is that what you're saying with the words --

BURKE: Yes, and if you're a child it depends on

your caregiver functioning.

MALE VOICE: Right, so Bob, are you okay with that, with losing the word -- because whether I go to work on Monday morning at 8 or at 12, is a biopsychosocial phenomenon, but it is clearly driven by what we're saying are our three primary domains of functioning. Let's give Bob Kerns a chance to object.

KERNS: Well, we're linking, the lynchpin here is the concept of pain, so actually -- well, I'm not sure.

DWORKIN: So Bob is going to get back to us after the lunch break. Jas, you had --

SINGH: So I think, getting back to that issue, that assumes that we understand everything about physical functioning, emotional functioning and social functioning, that they're all independent and that presence of one of those is not influencing the other. So we are saying that the work productivity or going to work on Monday is influenced by all these complex things, but we know that physical functioning is something we can see and feel very well and emotional function is something we see and feel very well, and social functioning is something we can see and feel very well. I'm not sure we can.

I mean I know that we have measures we've had for these for a few more decades than we've had for the other

construct, but I'm not sure that those things, the arrow doesn't go back and the arrows don't connect as much as the other ones.

MALE VOICE: Well, no, I don't think any of us assume that physical, social and emotional are uncorrelated. I think we would all agree that they all influence each other. So I guess the thing I do assume that I wonder if you're challenging, is that those three buckets are reasonably discriminable to keep the terminology of functioning is composed of the physical domain, the social domain, and the emotional domain.

SINGH: I'm okay with that, but I think that --

MALE VOICE: That's as far as we've gotten. The next step is your question --

MALE VOICE: Well what do you mean work by paid work for example? So I think in our culture we don't think about work as just paid work.

MEASE: But we're saying work is something out there that's another level where it's influenced by all three of the primary domains.

MALE VOICE: : As opposed to work in the yard?

STRAND: We should call it participation, because that's work within the home, that's work outside the home,

and that's family, social, leisure activities, that's what people want to be doing.

MALE VOICE: Role participation, okay. Dan.

STRAND: I had something else to say.

MEASE: Vibeke then Dan.

STRAND: Sorry. So, you know, we had this discussion about rheumatoid arthritis and proximal and distal, and I am concerned about it in the context of that discussion, too, in people with chronic pain, because chronic pain now impacts many other things and it leads to other perceptual changes and modifications in life and so on and so forth. So I have a really hard time, once something has become chronic, deciding what's really proximal and what might be distal.

And, for instance, you know, in an inflammatory cause of chronic pain like rheumatoid arthritis, there is a very significant amount of fatigue that goes along with the pain, and the physical function is affected not just by the pain but by the actual inflammation.

So I worry a little bit about trying to make that distinction too clearly. I think that's what you're trying to talk about with the biomedical model.

MALE VOICE: So Vibeke, let me interrupt, I think we should lose the term proximal and distal.

STRAND: So do I.

(Cross-talk)

MALE VOICE: So I would propose that we have the core functioning domains of physical, emotional and social, and then we have some compound domains, or I don't know what we could use, but not distal, sorry, Laurie, some compound domains that are things like participation and whether the kid goes to school.

BURKE: Well you don't have to be sorry because that's exactly what this proximal/distal model is explaining is there are things that are disease defining, core signs and symptoms, and then there are other things out here that we do not know the relationship but they're associated with this condition like fatigue in RA, like other social, emotional things, and physical function, and we cannot say that the RA is the thing that is fully responsible for the change in physical functioning, but we know that it has partial responsibility for it, but there's all these other effect modifiers.

MALE FOICE: So it's not only participation is a compound domain, but I think sleep is, because sleep is clearly a function of emotional and physical, et cetera, and fatigue seems to be a compound domain also that doesn't simply get reduced to physical, emotional or social. So

we've got three core and an unspecified number of compounds.

Lee is smiling, and those of you who know Lee Simon, if Lee Simon is smiling I'm happy (laughter).

MALE VOICE: It's because he doesn't have to go to the dentist.

DWORKIN: Bob.

KERNS: I think it may be worthwhile spending a couple of minutes thinking about the idea of, again, chronic pain as a disease maybe, I don't know if we need the word disease, but chronic pain and then starting to avoid talking about pain and functioning as if there is a direction of relationship. I think it's best to think about them as interrelated and because we have thrown out in casual ways the idea that peoples' pain reports or reports of pain improve when they're working, for example. So work certainly influences one's experience of pain.

Really I think we want to be careful even in our discussion about not throwing out terms that imply causality or direction or impact. So anyway --

FEMALE VOICE: Like a phenomenon.

KERNS: So our field is moving in the direction of what we have talked about for a long time and in a deeper way about how frankly the biopsychosocial model

seems to keep working, right?

DWORKIN: So I think it was Ernest and then Dan.

ERNEST KOPECKY: So Ernest Kopecky from Collegeium. I guess, unless I've missed this, we've kind of come up with these core domains but I want to come back to this vocational or this work domain. It seems like a lot of efforts in clinic are to get people back to something and one of those is to social activity and the other part of it, which may be even a larger part of it, would be get them back to work or get them back to school.

So I'm struggling with how we kind of made a vocational category a subcategory of social, and not actually one of the core categories. So in my mind, the way I see it is that we're looking at two perspectives. We have a patient perspective and we have a caregiver perspective. And under the patient perspective we talked about functioning and then we would actually have emotional, physical, social and vocational core domains. And then from there we've had lectures and we've talked about different assessments that we could potentially use under a clinical situation or a clinical research situation that we would apply to judge whether a treatment is making them better.

DWORKIN: Okay, so the way I understood it is

that people in the room thought that this kind of participation domain, work participation, role participation domain, was not secondary to the core three but it was some compound consequence or interaction with the core three. You're suggesting that it's so important that even though there might be some bidirectional arrows, we make kind of role/work/school participation a fourth core domain. What do people think?

KOPECKY: You could make that same argument, Bob, for the social.

DWORKIN: Right, that social is a compound --

KOPECKY: Exactly.

DWORKIN: So Ernest is modifying our figure one and suggesting that we go to four core domains rather than some core domains and compound domains. It was Dan and then Ajay.

CARR: So what I was going to say I think overlaps with some of what Bob Kerns has said and it's directed more towards the preparation of a manuscript and clarity. I think that there are, there is one section of a manuscript that I would like to see look something like this. There would be a statement that there's a consensus that views chronic pain as a disease or disease process or condition, and hence we are motivated to develop outcome measures in

the parallel way that have been applied to other disease conditions like pulmonary disease, cardiac disease, diabetes, et cetera.

With that thought in mind, and being very impressed by the actigraphy data, I would suggest perhaps a three by four table at some point where the three elements are, if you consider physical, emotional, social function, and the four columns would be first an immediate continuous objective and ecological measure. And that could be actigraphy --

DWORKIN: Could you save that for about, at the rate we're going, two o'clock this afternoon? Because I'm still at the level of the very top of figure one and you're going all the way down to the lower levels. But that's, what you just said is exactly where I hope we're going to end up a 1:30, 2:00, 2:30 this afternoon.

CARR: It's an idea to take away the arrows at least in something, to make it a table instead of an arrow diagram, and to just draw parallelisms -- I'll just finish the thinking but I'll keep the piece of paper from my notes. So one would be immediate, the next would be a patient-reported outcome, the next would be an assay such as the curvy walk type thing that is a surrogate for something but we don't ever spend our time walking around

the cones, and the fourth would be some other level, higher level still.

So I just want to throw that idea out now, we can keep it on hold for a few hours, but it's trying to get from where this discussion is going to a manuscript that we can agree on.

DWORKIN: So I was remiss because I didn't say what you just said, Dan, and the kind of idea here obviously is over the course of the remainder of the day to flesh out exactly what you're saying. I was just kind of -- we're starting at the top and then we're like let's drill down so that we have a fully -- well as fully fleshed out as we can concept of interest of functioning. And that's why Ashley and Laurie are up there, because if we get off the track of fleshing out the concept of interest their role is to pull us back on track.

So are we still on track for fleshing out the concept of interest?

BURKE: Yes, but I'm concerned about the concern about a directionality, and I think that the reason for, we can drop proximal/distal if you want, but for that idea is that if you just measure something out here with what we call distal, if you just have a score that represents work/role participation, vocation, whatever you want to

call that, we cannot necessarily link that back to an intervention without having some more specific things in between. And that's the purpose of that idea. So there has to be some sort of directionality involved.

DWORKIN: So I think we get to that, Laurie.

David.

HADDOX: I think what you are trying to say is that you've got things that an intervention can intervene on in a relatively pure fashion, as opposed to something that's distal where there's so many other factors at play that this intervention, the effect of that intervention, is either indeterminate or it's diluted.

And the example I used in my clinical practice in the old days was we made a distinction between release to work and return to work, because I could release you for work, I could say that your psychological, your physical functioning, your analgesia, you're fine to go back to work, you're good by me, and you still didn't go back to work. And I didn't want that to be on me because there were other factors out there I could not control. Your employer might say we don't have light duty, sorry, pal, it's fully in board or you're out. And I think that's a good example of a distal effect where the analgesic and then the rehab was great, but the guy never went back to

work.

DWORKIN: But I also think, and this is just what you said, what's distal, to use your terminology, depends on the intervention. So it might be that a pill, that getting a person back to work is distal, but if you offer that person a \$200,000 annual raise, that gets them back to work very quickly. So the intervention is going to have an impact on how distal or not.

So let's go back up to the level are we all happy enough, Lee's left so I can't use him as a barometer, with four functioning core domains of physical, social, emotional, role participation, something kind of nonsocial role participation, we can --

FEMALE VOICE: Just call it participation.

DWORKIN: Just call it participation. So all right, so we've got a kind of taxonomy here, if you will, of functioning, four core components of functioning, domains of functioning, so then I guess the most important question is going to -- after Dorcas' comment.

BEATON: I am just wondering, what I heard us talking about is that this disease of chronic pain is impacting on peoples' lives in four areas that are important to consider and one is sort of this functioning level and then one is this compounded or integrated

functioning, are we not --

DWORKIN: I think it was Ernest who convinced us that participation is so important it should be considered a core domain in and of itself.

BEATON: I have absolutely no concern with it being a core domain, like I am right there, I think it is really important because I think it's right there in peoples' minds and what they want to be able to do. But what I am wondering about is if we wanted underneath, I'm not sure if it's going to help us to subsume it underneath the term functioning, as opposed to saying these are things about a person functioning, what they're able to do, their capacity and what they're performing. And then when we put it in a social context of having them to do a job or care for a child or manage a home, that's when it becomes more of a participation.

So that would not diminish it at all one iota from being a core domain, at all.

DWORKIN: You're saying it wouldn't be under functioning?

BEATON: There's a lot of different models that would suggest that it's a different phenomenon, not a different importance, it's just a different phenomena than just functioning within your skin, it's almost like within

your skin, then you're applying that to a situation that has --

DWORKIN: So you're saying there would be a functioning with three core domains of social, physical and emotional and then on a par with function would be participation.

BEATON: Absolutely on a par.

DWORKIN: So she's actually promoting you, Ernest.

BEATON: Yes, I would not say it's in any way diminished.

DWORKIN: Bob and then Ajay.

KERNS: So with this I feel a little uncomfortable not hanging my hat on something like, you know, there are people that have thought about this other than this group, like the ICF, the people that came up with the ICF. So if we can some way tie this to some credible framework that is already out there that is accepted in the world of people that think about functioning. And by the way, in this, you now, I think if I'm comfortable with the physical function or physical -- yeah physical functioning domain as long as it's the fully spectrum from, you know, like the integrity of the nervous system, right, you know, do you have nervous system capacity to move your leg, all

the way to, I don't know, what's the farthest, you know, actigraphy, like gross motor functioning, right?

DWORKIN: So we will certainly look at what other approaches, ICF, et cetera, have done, but I think OMERACT and IMMPACT have credibility. And if we decide that what we just came up with in the last half hour seems to work well for pain clinical trials, I think we have the credibility and the authority to kind of propose something.

Ajay.

WASAN: So I agree that scientifically it makes a lot of sense that participation is a core domain, but we have to really think about the political implications. So good examples are in the physiatry literature when papers come out about the importance of work and importance of return to work, then the occupational health literature, the health insurance industry, the worker's compensation system seizes upon that as saying, oh, if your treatment doesn't show return to work and it's a core domain, we're denying it. And that's a huge implication that can really happen very proximal to a paper coming out because of the reputation and the esteem that IMMPACT has. And people model the approval of what is considered a good treatment on the IMMPACT results, on the IMMPACT recommendations. And that's exactly what all of your fantastic work has shown in

20 years.

DWORKIN: We've written paragraphs that have addressed your concern. Ernest, were you --

CHOY: No, no, it's a different side, different thing about how OMERACT works with ICF. So actually when we come up with different diseases, we don't put participation as a domain. What we do is we map it to the ICF participation framework, with each domain we map it to the ICF framework. We've done it for different conditions when we have different domains individually map the domains to the ICF framework and we work with ICF on that.

DWORKIN: Was there another hand? Chris.

VEASLEY: Oh, you're going to love me for bringing this up, but I think we have, I'd be remiss for my pelvic pain colleagues if we didn't talk about this, and if we're going to talk about all kind of levels of functioning, sexual functioning is something that should be added.

DWORKIN: Where does it belong, is it emotional, physical or social?

VEASLEY: I think it's different, it's usually, when we talk about it, we talk about biopsychosocial sexual --

DWORKIN: So this is a wonderful example of a

kind of compound domain where it can't be reduced to one of the three core domains.

Dorcas, were you going to say something?

BEATON: I did.

DWORKIN: Laurie.

BURKE: So I think that it's important also then to clarify how we're going to distinguish between these meaningful health aspects like physical activity in daily life versus, which is sort of a participation. So it sounds like if we are starting to add these participation domains like vocational and sexual, then we have in this, you know, physical activity in daily life, getting off the couch and doing stuff because you can as opposed to other -- and so is that really directly under physical function with neurologic function then?

DWORKIN: So I don't know if this is the same question, but it seems to me now that we sort of have a consensus of three core domains and then participation and some compound domains like sexual function, sleep, and there's sort of a consensus, the next question is what are we, what have we done in this meeting and the manuscript that Kristine and Ann are drafting, what is that going to focus on.

So we've kind of mapped out a taxonomy but we

haven't yet decided are we limiting our remaining considerations, review of the literature, et cetera, to physical, or are we also going to include social, are we going to include participation, are we going to include sexual. And I think your question comes right after that question. So we've got a little bit of a consensus map but what are we going to really focus the rest of our discussion today on and the manuscript on?

BURKE: Participation is really about whether patients, people, can do what they want to do. Now they may perceive that they want and have to work because they need the livelihood, but the social, leisure, family activities are pretty much up there and a lot of people will sacrifice other things so they can have that. So I think that we can't put participation below.

MALE VOICE: Clarification, maybe this will make it better, maybe this will make it worse. Let's assume for a moment, let's not argue about it, that we had these four core domains, any one treatment doesn't necessarily have to bring about changes in all four domains as all we have said in the past is that when you are considering your study, you have to consider assessing those domains. But if I have a treatment that's specifically focused on fatigue, then I might have a different core outcome or primary outcome that

I want to use, if I have a rehabilitation treatment I may have a different one.

So it's not as if, thinking to Ajay, this is not to suggest that any one treatment must deal with all these domains, but rather, when you're thinking of your study, you should be considering whether these relevant domains are appropriate in yours and deciding what your primary is may be different depending on the treatment focus.

WASAN: I totally agree with that, all I'm saying is that people have twisted that very proximal, very soon after these things come out, and we have to be aware of, so describing participation, making it broad, making it clear that it's not just about return to work, you know, and there's all these other social, economic factors that determine that, that kind of language becomes very, very important in a document like this. So that's why I bring that up.

DWORKIN: Dorcas.

BEATON: Just to add to that, the World Health Organization, although we have this nice framework with these boxes in it, spent probably years struggling with the difference between activity limitations and their participation. And they are actually the same list of codes. So it's not quite as easy and so I think we have to

acknowledge that this will be a thought process of how do we separate the activities that are part of my within the skin experience of getting dressed versus having to get dressed really fast because I have to get to my job. So that changes that.

DWORKIN: Chris.

VEASLEY: I guess what I'm struggling with still is, you know, hearing yesterday that physical functioning would not make it into a label claim for FDA, and that it's a distal, possibly a distal thing that you're reviewing, if you're reviewing a treatment for pain.

If we were to make -- where I'm struggling is if we were to make recommendations for all of clinical trials just because you want to understand, these are the things that you would want to understand change with a therapy, that may be very different than what we would recommend to a clinical trial, would then go onto the FDA for approval of -- because I'd like to hear from kind of the industry colleagues in the room. Because what is data going to be, if it's not a primary endpoint for approval in the FDA process, when is this information, how is this information going to be helpful over time? Is it post marketing, is it just for general understanding? Does that make sense? I mean clearly if we are going down two different paths for

making recommendations, at least in my mind the recommendations might be different.

DWORKIN: So I think Chris' question is what is the industry view of what would be secondary endpoints in a clinical trial.

VEASLEY: And FDA, where are they considering this, you know, these physical function measures, social function, participation measures, this process?

DWORKIN: So Ernest, you were next in the queue, do you want to take a shot at Chris' question before you ask your question?

KOPECKY: Sure. So from an industry perspective, for us, we really hinge off of what guidance we get from the agency and the label claims are primarily on the primary outcome measure. So this is going to be a secondary outcome measure. Right now as it stands it's not that helpful from that marketing perspective, which is different than I think the focus of what we're talking about here in looking at these domains and how to define a core set of measures for each of these domains. I think that we've got a number of different levels of things going on over here that we have to be clear that we separate out. But that's kind of my view from an industry perspective as we stand right now.

DWORKIN: Sharon, so the regulatory view of secondary endpoints that don't get into the label.

HERTZ: So we're not here to discuss that, the regulatory perspective, and if you have questions about products under development or approaches, we always are happy to entertain meetings through proper channels. So I don't think that the conversation should really be focused on -- so I know there are practical aspects here, but this discussion is about the science and what's best to understand the condition and get the information and clinical studies that's relevant for the scientific question about understanding what's involved in the study, be it drug or nondrug therapies.

So I'm not going to answer that question because it's not part of this conversation.

DWORKIN: All right, so Ernest, I wonder if an answer, not a regulatory answer to that question, is that secondary endpoints give you a more complete assessment of the impact of your treatment, right? It kind of fleshes out what the treatment, whether it's a drug or acupuncture, is doing in the patient's functioning in daily life. That may or may not be the regulatory answer, but that would be my answer, you know, does this treatment, in addition to relieving pain, also have a downstream effect on increasing

their social participation or not.

KOPECKY: That's the answer on the scientific medical side, which is different than your original question which had a regulatory component to it.

DWORKIN: Chris' question.

KOPECKY: Or Chris' question. Can I get to my other question now?

DWORKIN: Absolutely.

KOPECKY: So I just want to make sure that when we go into the manufacturing we clearly differentiate this concept of participation that's evolving, because even in the last set of discussions, we have applied participation to multiple domains. And is still hear, we talk about the social domain, in my mind the social domain is very distinct and clear from a vocational or work domain, and when we've talked about this, we keep kind of skating across the two with this now higher level concept of participation which I think may be confusing. So we'll have to kind of really narrow that down when we're talking about what exactly participation means. Because I see that across all these domains.

CHOY: So I think along that, I always think that we should keep participation as the ICF defined it, because otherwise we're going to muddy the water. If we're trying

to reinvent what participation is, we must stick with what ICF defines participation as.

DWORKIN: Okay, so in the interest of moving things along I'd like to make a proposal that I think follows from what Ernest just said, both Ernests, so that given what we've discussed at this meeting and the way we kind of conceptualize the meeting, that of the four buckets we've been discussing, emotional, social, physical and participation, that we focus the rest of our discussion today and the manuscript on physical functioning and we acknowledge that the social domain and the emotional domain, and the participation domain are equally important to physical functioning but the focus of our efforts is physical functioning because we really didn't prepare to do participation in depth, or social functioning in depth, and we certainly did absolutely nothing about emotional functioning. And so that the discussion we've had so far for the last 40 minutes or so is really just the introductory paragraphs of a manuscript setting the stage, what's the world that the manuscript was within and then what part of that world are we going to focus on.

I'm waiting to see how much disagreement there is.

KERNS: Bob Kerns violently disagrees.

DWORKIN: All right, we'll start with you.

KERNS: I'm kind of surprised about that because I didn't think that we've spent a day highlighting that relative to anything else. And within the domain of physical functioning we kind of rah-rah actigraphy, and I agree, but that's one relatively narrow dimension of what I think we're talking about as the domain of physical functioning. We didn't talk at all about fine motor functioning, for example, with people with arm related problems, or other aspects --

DWORKIN: So you're arguing to keep in social and participation --

KERNS: I'm just saying if we zoom in, if our goal is actually to pronounce, bless actigraphy, then yes, but I don't, I don't think that's our goal, right? So I don't know what we're going to talk about other than that and gross motor functioning, physical activity --

DWORKIN: Ann and Kristine talked about a whole lot of stuff that wasn't actigraphy and that also wasn't social and participation.

BEATON: And if we tease out, some of them were more participation oriented, if we tease those out of that list there were still quite a few, and then the performance based measures which might be the Mulberry Pickup Test or

the Jepson Hand Function Assessment --

KERNS: But we didn't talk about those, we didn't talk about a lot that's in that domain.

BEATON: So it's like pulling apart physical functioning now and what's in that spectrum.

MALE VOICE: Bob, so what I don't understand about what you're saying, if we are going to cover all of these domains except emotional because you don't think we really focused really on physical, then we have three times as much homework to do,

KERNS: Well I think we could focus on physical but then even there we're going to have to acknowledge huge limitations I think.

DWORKIN: So why aren't we suggesting narrow it to physical and leave social and participation for another meeting, another day? Does anyone think we should kind of exclude emotional and include in one effort from this meeting social functioning, physical functioning and participation? Vibeke.

STRAND: I just don't see how you take, how you look at physical functioning without looking at it in the context of participation in activities that patients want, people want to perform. So I don't see how you can say social, leisure, family activities aren't included, I don't

see how you can say work within the home isn't included, and I don't see how you can say that work isn't included.

So to me, I don't see how you measure physical function without asking those very questions.

DWORKIN: But couldn't I respond and say you then need to include emotional functioning because I'm less likely to go to work if I'm depressed?

STRAND: Well then we can go back to HRQOL and talk about that kind of stuff, too. But I mean the fact of the matter is we need something that puts the physical activities in the context of life, not just activities of daily living but life.

DWORKIN: Dorcas, help us out here.

BEATON: Well one way to think about it might be the example of return to work that David gave earlier. You might be able to improve somebody's ability to do their job, and I might be measuring that using --

STRAND: I mean the questionnaire that I really like to share is the one that asks about all those things pretty simply in a single page. And it's only been validated in arthritis, chronic arthritis. But I think it is actually very applicable to lots of different settings. And it looks at more than just, you know, work outside the home. That's all I'm trying to get at. Something fairly

simple that shows the impact of the relief of pain on the ability to engage in those activities that are most important to you.

BEATON: Participation level.

STRAND: It's the WPS, I mean I've got some slides, but --

MALE VOICE: Couldn't we acknowledge that it's physical function in context is important to consider and when we think about physical function in different contexts we need to still find ways to measure that and then we do on to talk about different ways to measure, acknowledging that these all occur in some type of context. They don't exist in the ozone, but we acknowledge it and then start showing what are some of the kinds of measures and how they apply to things like in the actigraphy example, which, by the way, we did not intend to be endorsing. All we did was pull it as an example.

STRAND: I think that would be great, I think we put actigraphy there, I think we could put this instrument there and all of that.

MALE VOICE: And on the self-report measures as PRO measures, many of them are referring to the contexts in which people are doing things. So it's not that we're forgetting the context, it's just if, in fact, you endorse

the fact that when you consider physical functioning it's essential to consider physical functioning in context.

These measures are available and as you are looking at them you have to think about the person and the situation that are relevant.

DWORKIN: So I think, and I don't know if this is what you are going to say, Laurie, I think we're still at the level of physical functioning, actigraphy is all the way down here, and we have to spend three hours talking about, before we all get on our planes, what's in between physical function as a concept of interest and the kind of very specific measures like actigraphy. And we haven't even talked about what do we mean by physical functioning, we're still stuck at the level of are we focusing on physical functioning or social functioning and participation.

KERNS: Why not write a paper that's -- I think there is grist for the mill here, major contribution, even outlining this four domain kind of view and all the ideas about how they're overlapping, interrelated, bidirectional, and so forth. I don't think we dug -- and then subsequent IMMPACT meetings are to go into each of those, including physical functioning. I think there's a lot of advance that's been made here by staying at the --

DWORKIN: So for those of you who've been at

previous IMMPACT meetings, what happens next won't surprise you, Dr. Bob Kerns who is associated with Yale University and the West Haven VA in Connecticut, has just volunteered to spearhead an article that presents the IMMPACT/OMERACT roadmap taxonomy framework of functioning which will give at this kind of 35,000 foot level, the rationale of these four core domains, physical, social, emotional and participation. And Bob will be the senior author on that paper. The manuscript he will draft will be circulated to everyone in this room and you are all welcome to be co-authors. Sound good?

MALE VOICE: And it will be done in the next four months.

DWORKIN: So what Bob's suggested, so he's going to do that manuscript, so that's going to be the first manuscript that comes out of this meeting, and then what Kristine and Ann are going to work on, which obviously follows on Bob's overview, could be a topical review, you know, I don't know, might be able to do it in 2,000 words, I don't know, something that would be interesting to consider. What Kristine and Ann would do is to then focus on the physical functioning bucket in Bob's kind of framework in the manuscript you would be drafting. Does anyone disagree with that?

MALE VOICE: Bob, you can't put your hand up.

DWORKIN: Laurie, please don't disagree, he just agreed to write an article.

BURKE: I agree with that (laughter). What I disagree with is calling participation another core domain. Because as we begin to think about where participation goes, it is not separate from physical, emotional, social, it is within. And I agree with Vibeke completely about your last comment, too.

So I think, I don't know, Dorcas and I have been drawing all kinds of different diagrams over here, but I think that participation, and I won't use distal, but it is distal to physical, emotional, social, and it's part of all of those domains. So I don't think it can be a forth domain and that's why I --

DWORKIN: So I think in terms of moving forward, we don't have to resolve this now, though what I would like to propose is that since we're ten minutes away from lunch that Bob, and Laurie, and Dorcas, and Ashley have lunch together and talk about participation while they enjoy whatever, and Ernest, and Vibeke can have lunch with them. So we have one of the lunch tables is now reserved and closed for further attendees, but Philip wants to throw a monkey wrench into this.

MEASE: Not at all, just to be clear, that a lot of this work has been done already. And we have a lot of work, as multiple people have said, done by the ICF and also OMERACT 2.0, which in many ways mimics what the ICF structure is. And so those would be part of the introduction of any such article is to say this is what has come before and where does our work fit into this.

DWORKIN: And so that's what Vibeke, and Bob, and Laurie, and Ashley, and Dorcas are going to work on. And if there is room at the table Philip is going to join them. And so what we would like when we reconvene after lunch, if Bob, since he is spearheading this effort, can just summarize for us, he's not listening, if after lunch Bob summarizes for us in 10 minutes what you guys decided about participation and these other three domains.

So that's going to happen at lunch, we've just added a manuscript to the queue, so I guess before lunch are we all in agreement that the second manuscript which obviously needs a lot of work in our discussion after lunch, really focuses on the physical functioning bucket. And what -- so as a concept -- is it concept, yeah, concept of interest, and what exactly do we mean by physical functioning based on all the presentations we've heard and all our prior experience, how do we flesh out that concept

of interest of physical functioning, what are the components.

Now one thing I think we had a lot of talking about is that there is physical functioning and there are physical activities and that's not exactly the same thing. So after lunch someone will be at the white board I think and what we should all try and do is come up with some framework, taxonomy, roadmap of what exactly we mean by physical functioning, physical activities, I don't know what, and then eventually and the faster we get there the sooner we all go home, we get down to the level of actigraphy, a discreet measure, or the multidimensional pain inventory. And at some point in between physical functioning and actigraphy at the bottom of the flipchart, we're going to have to deal with do we mean kind of physical functioning in general or is it pain interference with physical functioning. And it sounded like those are two different things. So it might be two separate flipcharts. Does that sound like a plan for what we should all be thinking about at lunch and what we're doing after lunch?

Philip.

CONAGHAN: What I was I think alluding to yesterday and because I can't see what Dorcas' and Laurie's

diagrams look like right now, is if we took figure three from that Tomey paper and just head it up as a straw man conceptual basis for just the physical functioning. I agree with the overall plan but if we have that up on the screen as a starting point, it might save us a bit of time. It's ICF based is what they've done, it falls within OMERACT's sort of principles, it might just be a starting point. But Dorcas, you probably know the literature better.

DWORKIN: Do we have that on a slide? So let's have that slide up after lunch, a framework for our after lunch discussion.

Any other, we've just got a couple of minutes, I think we should break for lunch at 12 for those of you who haven't checked out yet, any other comments about this plan that's sort of evolved over the last hour? Bob's article and his writing team and what we're going to be doing after lunch, what we are all going to be thinking about at lunch?

TURK: Don't be afraid to speak up, we promise we won't ask you to be the head of a paper (laughter).

DWORKIN: Dennis promises you, I make no promise to honor Dennis' commitments. I don't make promises like that. Any other comments from the panel? All right, well let's break now for lunch, those of you who have to checkout at 12, you should do so. Let's reconvene around

1:15 and see if we can come up with a consensus quickly and
leave early. Thank you all very much.

(lunch break)

**Group Discussion II: Considerations for the Assessment of
Physical Function in Analgesic Clinical Trials Using
Clinician, Observer, and Laboratory Outcome Measures**

just thought we would show again the two framework slides that were discussed, one is the ICF framework. Can people see this that are over -- well Jim Witter knows it by heart, so he doesn't need to see it. And then we'll show the next slide is the OMERACT one which incorporates many of the same concepts but then also there's some slight differences. So we're going to go back and Bob is going to, part of the discussion is how this also melds with the biopsychosocial continuum. So, Bob --

DWORKIN: So I'm just getting over my anxiety, anger, et cetera. I did what all smart people do, I enlisted the help of another colleague, so Phil has agreed to collaborate, provide leadership for this group. I think it turns out we're on the same page, and having talked over lunch with our developing group, evolving group, I think we really are kind of converging in an important way.

So I think what I'm imagining now and I think that Phil is in agreement, the team is in agreement, is really a paper that's staying at a 30,000 foot view, kind of providing a conceptual framework for understanding

functioning in the context of pain or maybe pain, we really haven't talked about that, with an eye to laying out a framework for essentially kind of a research agenda, but also importantly consistent with OMERACT and IMMPACT, thinking about interventions in particular, kind of the endpoint.

I think we haven't talked about any of this in great detail, so this is all for further discussion, but I think we do agree to start with a biopsychosocial framework as the framework that has been I think largely universally adopted within the pain field as an important kind of theoretical context in which to have these kinds of discussions and scholarly descriptions, I guess, ultimately in the service of being able to kind of link back to that theoretical framework and see to what extent any work that emerges from this actually reinforces or supports the broader context and therefore can have some, you know, generalizable appeal, I guess that's always the value of theoretical framework.

It turns out that, at least as far as we know now, both the ICF in its foundational work is linked to the biopsychosocial model and OMERACT has had that in mind as well, and there are important papers that Phil in particular is aware of that we will largely kind of

attribute out ideas as we start to move in the direction of something more specific about pain or chronic pain.

So there are two models that I think we should spend a few minutes reflecting on and then in addition, Laurie and some others have been working kind of in parallel, you saw them up here in the earlier panel fiddling with different kinds of schema for understanding these same concepts. So I think we, Laurie is going to maybe put some of that on the board maybe as we talk or she can do it now, for a few minutes anyway, not long because we really want to move onto the physical functioning.

BURKE: Do you want to do that first?

DWORKIN: Yeah, I think we should zoom in on these two models, reflect on them a little bit. This is really, you know, very consistent with my ideas about linking to what we already have out there, right, the ICF and OMERACT's work, and then knowing that or imagining that we're likely to further refine them, build our own framework that's informed by these models, as we think more specifically about pain and pain management.

MEASE: And this year, just to mention, at lunch we discovered in PubMed that Stucki had created a new paper about mapping to the ICF structure and so one of the activities could be doing a mapping exercise to the ICF

structure. One of the elements of the conversation beforehand was that, as you can see, there's a bidirectional arrow here between activities and participation and activities, body function, et cetera, et cetera. So the question is where does participation lie in the hierarchy we were discussing before lunch.

And then maybe just to contemplate it a little bit more, just take a moment to look at the, because this was shown very quickly yesterday, but just to remind us about OMERACT 2.0 and where life impact, resource use, economic impact and pathophysiologic manifestations and the different elements that go into that. And all of this is contextualized. So depending upon the population that you're working with or studying will make your context for you.

DWORKIN: So Phil mentioned Gerald Stucki, I remember meeting him, right, and so it may make sense, it seems intuitive to reach out to him, in particular, and there may be others, there may be others in this group who would like to join this. I think organizationally I'll just put this out there now that we'll probably try to convene via teleconference relatively soon. I think in terms of a timeframe for seriously working on this, for me it's more likely to be fall, I'll just declare that now, but so I

don't think this is Bob's idea, our group I think is a dynamic one, we certainly accommodate other people who have this interest, as well.

So, Phil, do you want to say anything more about either of these?

MEASE: No, I'm willing to let Laurie --

DWORKIN: Yeah, I think that's a good idea.

BURKE: This is a consolidation of ideas from the folks on this end of the table. Okay, so this circle right here, and what we're drawing out are outcomes now, it's different than just a nomenclature system such as described here, these are really outcome concepts.

DWORKIN: Staying away or not really getting too close to measurement and methods, right?

BURKE: Right, because we have to figure out what we're going to measure before we decide how to measure it, okay --

MALE VOICE: Can I just ask, even though we don't have a measurement yet, are we still framing this within the context of clinical trials?

FEMALE VOICE: Or this might be the array of concepts that could be impacted and then you might select from this for clinical endpoint, being aware that there's this whole array of impacts.

BURKE: So I'm going to define four dimensions, okay, so wait till I get through four dimensions so you can kind of see where we're going. Okay, right here, this is the person. And also within this person are all the attributes of this person that you would define as your entry criteria, your inclusion criteria, your exclusion criteria of any sort of a study, an intervention study where you're going to measure outcomes. So it's there, demographics as well as the severity of the disease, the subgroups, important subgroup potentially, whatever you think makes sense to define as the persons of interest, okay.

So then these three domains are the physical, the emotional, and the social. And I'm not saying proximal versus distal, but on the other hand, the concepts become more specific to more general. I think that's a better way to say it, more general concepts out here.

Well right here in this wedge, participation, because participation outcomes are influenced by the physical, the emotional and the social domains, rather than making it a separate domain out here somewhere. So then at any point in time you can think I want to measure work participation which would be this concept right here perhaps, or whatever concept you want to think of, you want

to measure as an outcome.

Then the next domain, I mean the next dimension are the impacts, and this is all at a single point in time, this is not -- time is a fourth dimension that we're not going to be able to draw on here. This whole thing is for what -- so we're going to measure this outcome at six months, say, so that's what that dot represents. If you want to measure it at a different time you have to move the whole thing somewhere else. And here are the other impacts --

SIMON: Laurie, you used the term general --

BURKE: You like that, fine.

SIMON: Well I'm asking a question, which side is more general, the left side or the right side?

BURKE: I think over here, but that's to be debated.

SIMON: So you're saying that the left side is more specific and the right side is more general?

BURKE: That's how I'm thinking of it, you don't like it that way?

SIMON: I'm finding it interesting because you are now picking out very specific things like participation on the right side, whereas the left side was about general things of, you know, the three different topics were pretty

general. So I'm just not sure how you're defining it.

BURKE: Something right here would be walking.

SIMON: Why would that be there and not to the right?

BURKE: Okay, you're getting way too specific for where we're at in this discussion, okay? That is really something that this group needs to, first of all, see if there is any -- does this resound with anyone?

(cross-talk)

BURKE: I haven't finished my four domains yet. Okay, so this, okay, in terms of dimensions, this, coming in this way, are the interventions or the impacts on all of these outcomes. So this would be coping mechanisms, there would be environmental impacts, there would be motivation on the part of the person to function in some way. There would be personality, all kinds of things in this direction that you have to decide whether you're going to measure them or whether you're going to randomize them, assume they're the same, or what are you going to do with these. And then the dimension coming out of the page is the measure.

Okay, so if we're going to measure this thing, there's lots of different ways to measure it, which one approximates that thing the best, and that's the measure

development piece that we're not going to get to probably today in this conversation. Okay and then the fourth one is time when you actually have to move the whole box, whole three-dimensional piece to a different place because you're at a different time.

All right, now you can tear it apart.

CHOY: Maybe I can help with Lee's question about general and specific. So I guess on the left hand side when you look at any specific physical function, and use the example of walking as the instrument, because we start with the condition you will have a set of instruments that is very specific to that patient population. But when you take out what your result and then take out participation, you expect the impact of participation you are going to interpret it across many different conditions. Or you can compare that an intervention in one disease on the level of participation at a societal level. Does that help?

SIMON: No. Okay.

CHOY: That's how I was thinking about the specificity is related to how you need to capture that information with your tools and then when you take out the participation, actually what you expect, that you were able to look at the result on participation across a lot of diseases. That's what health technology assessment is

about, you have tools that you can generalize and understand different technology and how it impacts on society.

SIMON: But this is a single clinical trial.

CHOY: No, I think this is just a conceptual framework, it has nothing to do with any specific clinical trials.

SIMON: But what's the big oval circle to the left of that, that is referring to inclusion/exclusion criteria?

BURKE: Yes.

SIMON: You're going to have that for a population?

BURKE: It could be a, yeah, it's the group of people that you want to study.

SIMON: That you want to study.

BURKE: That you are planning to study.

SIMON: Right, that's not a population. That is a population, but it's a specific isolated population, this is not a health technology assessment, this is something else.

BURKE: It could be, you could have health technology assessments at all levels. Persons of interest.

MALE VOICE: I think we're worrying a little bit

about some semantics here that we perhaps could get away from, which is that I think the idea is that there's physical, emotional and social function and those are connected, overlap, interact with effect, with whatever word you want, with a whole host of other things. We need to come up with as simple a framework as possible for this. If you try to draw the Venn diagram for this of work and all these things they would all overlap in multiple ways, and it seems to me that there has been a lot of work done over 50 years that's decided that the simplest way to separate these is into these three, or at least a way of separating them is into these three functioning quantities, and that that is the, allows you to then look at all of those pieces as you move down into work and other pieces.

So however it gets culled, I think we're getting worried about how this is going to be implemented. I think what is being presented here is the concept that we start with these basic functions, physical, emotional, social, understanding that they overlap, and saying that we're going to start there and we're going to see how those interact and overlap with the others. I don't see a problem with that, and if we, as Bob would like to say, I mean it seems to me that there could be consensus about this, the issue is it's based on literally 50 years of work by a lot

of people and there is a real advantage to using something that is generally accepted as a basis to move forward.

MEASE: Both generally accepted and also can be mapped to, so to speak.

MALE VOICE: Exactly.

MALE VOICE: Laurie, I have a question, so the four arrows, you called those impacts. Did you mean by that term that those are factors that impact those interventions?

BURKE: Yeah, the concept that you're planning to measure. So they're things you have to think about in the context of your study.

MALE VOICE: So a semantics suggestion because an impact often is, in addition to this meeting, it is often, you know, the end result, maybe those are modulators?

BURKE: Define the situation of measurement almost?

MALE VOICE: Sorry?

BURKE: They're defining characteristics of the situation, which you are measuring at that level. So maybe, yeah, moderator would work for that.

MALE VOICE: I was saying modulator.

MALE VOICE: Without thinking too much about it I could imagine it being either a moderator, it could be

moderating factors or mediating. I don't know.

MALE VOICE: Being a little bit of a devil's advocate here, I think it's, just from my perspective, a little bit complex having that other dimension of time or length of distal/proximal. In the initial diagram that was demonstrated, three circles were given for P-E-S and then there's overlap in some elements of it. It seems to me just a little bit more simple. And here you have this dimension where you'll have to decide where a specific element here is either closer or farther away so it gives just another level of complexity which might have benefit, I don't know, but might be too complex.

BURKE: Sure. Well the reason I think it's beneficial is because if you want to measure an emotional functioning as a general concept, then that is why this is the general end, Lee, because there's a lot of sub domains or sub-items, at least, that have to contribute to the total score of emotional functioning. As opposed to down here, if you are just going to measure happiness or, I don't know, what's a single item emotional function measure.

MALE VOICE: But it would be the same thing if you just measured the whole circle versus one point in the circle. If you have three circles that intersect, you have

an area of intersection of all circles, you have an area of just two --

BURKE: Well the problem with that, the reason for doing this was our discussion about participation. And is participation a fourth domain of functioning. And I propose to you that it is the result of the three domains that all aspects of functioning have an impact on a conclusion about participation. That's the reason for this, as opposed to the Venn diagram.

MALE VOICE: But you would get that if all three circles meet at one point.

BURKE: Okay, well you can draw it that way if you want, I mean this is just really to illustrate the meaning of participation as opposed to it being separate --

MEASE: So I'm seeing, Bob I know wants to keep the conversation brief, do you want to move on?

TURK: Yes, because we haven't gotten to physical --

DWORKIN: So I would just say probably by way of process that our group form a group and then become students of the two foundational frameworks where people have already done a lot of work, ICF and OMERACT, and then we'll move into thinking about the applicability of those

models and frameworks for helping us think about a model that's I guess more specific to our work and pain or chronic pain more specifically.

So before we move forward let me see if I'm correct in an assumption, Laurie, Philip and Bob. I'm assuming that the framework your team will be developing, is going to have within it a big space for physical functioning, right? Okay, and so then what our see as our goal for the next two hours and it is probably not enough time but we'll do the best we can, is to figure out in this domain of physical functioning where you've got physical functioning at the top is the way I view it, and actigraphy at the bottom, how do we flesh out this concept of interest to physical functioning, what are the components of physical functioning that are important constructs, concepts, that we ultimately want to get down to specific examples of measures of those components of physical functioning?

So is that where we are, is that kind of reasonable summary, that we're going to now flesh out the domain of physical functioning? And kind of everything between physical functioning and actigraphy. Now it seems to me that we've got to start with a definition of at least a working definition of physical functioning before we can

talk about what are its components. And so I don't know, based on yesterday's discussions, whether physical activity is something that's within physical functioning or is somewhat separate.

So, Dan, did you have your hand up before?

CARR: I did, Bob, I wanted to interject two comments, the first being that we already saw yesterday some pretty good definitions that could be adapted already where much of the wording is already as people pointed out thought through and acceptable. Then at the outset, I think that before we go too far down the road of definition, we keep in mind how these different concepts will actually be measured to avoid having some disjoint where we have a beautiful theoretic construct and no way to actually measure that, when there are a lot of ways to measure it.

So let me jus bring up again and I'll keep quiet about this, a straw man, which would be that somewhere in this scheme or manuscript there might be a table made, and I'm going to call it a three row, four column table, where the three rows are the functions that are to be measured and a hierarchy of measurement as a straw man could be proposed where, to begin with, one has this continuous immediate objective and what was previously termed ecological monitoring.

So for activity, that would be actigraphy. One day when we're all wearing FMRI's that could be our anterior singular cortex activity, that could be our "moodometer," like the glucometer that is going to give us instant readings. So one is objective, immediate, continuous, and in situ. And then further down, there would be a patient reported outcome. Further down, there would be a surrogate test like the equivalent of pulmonary function testing or walking in a zigzag way between cones, things that we don't normally do --

DWORKIN: So Dan, maybe this is a cognitive deficit on my part, I'm having trouble thinking through your example because I'm still stuck at I don't know how we're going to define physical functioning. So it's hard for me to think about -- so I want to come back to that, but I think we need to have some working definition of what we mean by --

CARR: Let me just do this, I only have two or three more sentences, having gotten this out on the recording this will be factored in ultimately because I know you will be able to figure out what we mean by physical function, I'm confident.

So anyway, the last remaining things to measure then might be an observer or clinician impression of these.

So I just wanted to throw this out as a non-hierarchical, no arrow thing to keep in mind for the future and now you can return to the definition that you're struggling with.

DWORKIN: So, okay, would someone like to propose what we mean by physical functioning so we can then get to frameworks like what Dan just presented? Do we have a definition of physical functioning, does it include activity as a component? Ian.

GILRON: Very simplistically, and I'm just trying to think about this as an inpatient clinical trialist who wants to cut to the chase and go which measures can I use. And so I'm just wondering if we could sort of start with some examples and maybe work our way backwards, so we haven't talked about sexual function at all in all this, and that could be important in doing a pelvic pain trial, let's say, and I'm also keeping that in context.

So I'm just wondering whether we could define it as outcomes which rely largely on physical function, so, well, I mean it wouldn't rule out an occupational outcome or something that requires participation, but something that physical function is very important for. But we're not defining the set of outcomes that we want to mention or measure, we don't want to define the measures that we're going to mention.

DWORKIN: So I have to defer to Ashley and Laurie because maybe I'm off base here, if we're talking about fleshing out a concept of interest do we need to start with a definition of that concept of interest or not?

BURKE: Yes, we do.

DWORKIN: So Jas and then Ernest.

BURKE: It would be helpful to get suggestions about what falls within physical functioning. Like for example, do symptoms fall within this red area of physical functioning, does neurologic function fall between, I don't think it does --

DWORKIN: I assumed it was things like activities of daily living and walking upstairs and -- Jas.

SINGH: So I think Dorcas showed us and WHO has done this more than a decade ago, why do we need to redefine physical function?

DWORKIN: So what is that definition?

SINGH: She projected it in her presentation, we should --

DWORKIN: We have that definition.

CHOY: I'll read that out for you now, it said "ability to carry out various activities that require physical capability ranging from self care to more vigorous activities that require increase in decrease of mobility,

strength or endurance.

DWORKIN: Do we endorse that definition as a definition of physical functioning on which we're basing our considerations for the assessment of physical functioning in pain clinical trials?

SINGH: Yes.

DWORKIN: All right, we have a definition of physical function. Philip.

MEASE: Yes.

DWORKIN: So this group endorses the ICF definition of physical function. So maybe we'll all get to our flights earlier than we thought, has the ICF gone beyond that definition to specify what the important components of physical functioning according --

MALE VOICE: Yes.

DWORKIN: Thousands. Thousands. That's not helpful, we want fewer than that.

BEATON: They divide maybe the narrow part of the thing would be acts, movement, lifting, acts, tasks, doing up buttons, managing toileting tasks and then social acts, tasks and social integration.

DWORKIN: So is voluntary physical activity different from physiological activity?

BEATON: And I think the one that might be the

most challenging might be the acts, when we're talking not about a specific ADL activity but an act.

FEMALE VOICE: A volitional effort.

MALE VOICE: Different from pulmonary function which we're not talking about. We're not measuring pulmonary function, I think Dan used the example of somebody of a straight physiological measure, that would not be part of what we're talking about. It may be required but it's not the domain we're measuring.

MALE VOICE: You're referring to like heart rate?

MALE VOICE: Yes.

MALE VOICE: Interesting.

MALE VOICE: I'm trying to separate out a voluntary activity that someone performs versus an involuntary physiological response.

DWORKIN: So do we have two categories of acts and tasks?

FEMALE VOICE: And social integration.

DWORKIN: So social integration, is that we have a separate bucket of social, does social integration go in that other bucket?

BEATON: Social functioning, is that one wedge, social integration is when you are moving around into societal role, which is our participation.

FEMALE VOICE: So an example of social integration functioning --

BEATON: Well when you get to the level of social participation there is probably the blurring so it might be the work, the home role, the homemaking role, the parenting role, so it's taking all of your different parts of function --

DWORKIN: But Dorcas, isn't that participation?

BEATON: Yes.

DWORKIN: So it would be nonparticipation, social physical functioning.

BEATON: There might be parts of social participation that you might measure at the level of just physical functioning but that would be harder. I think that would move more towards asks and tasks being the main parts of physical functioning.

BURKE: So do you think all this social integration goes in this participation wedge?

BEATON: That's my personal understanding is when you get into social role in participation you are out in participation.

BURKE: So it's in the participation wedge of physical functioning?

BEATON: Yeah.

BURKE: And what is an example of task?

BEATON: Getting dressed, managing toileting

--

BURKE: Oh, ADLs.

FEMALE VOICE: But it's more than ADLs. Much more than ADL. ADL is, that's not where we're at anymore, we want to do a lot more than that.

BEATON: May be recreation, could be tasks required for my job that I can be working for a longer period of time, I have the endurance --

FEMALE VOICE: But doesn't that go in here?

BEATON: That one is actually applied in the social setting.

FEMALE VOICE: I can do my home work better --

BEATON: Yeah, I can physically write and then participation would be I can write fast enough to keep up with my job and my student --

FEMALE VOICE: Oh.

BURKE: Work related activities.

FEMALE VOICE: But they are work within the home, too.

BEATON: Absolutely, but it's more at the -- those acts and tasks are within a person, they aren't

integrated into the social role yet.

FEMALE VOICE: Well good, then this works that the participation piece is within physical function but it is also part of participation.

DWORKIN: So to go back sort of to Ian's question, do all the measures that Kristine and Ann talked about yesterday, to the extent they were focusing on physical functioning, fit into these now too more specific buckets of acts and tasks?

TAYLOR: Yes, but there are some gaps.

DWORKIN: For example.

TAYLOR: A lot of them are, I think a lot of the physical function measures are clinician driven, so what do we want to see in the people who are living with chronic pain do as opposed to being participation driven, how are we going to measure what the person living in pain wants to do. And that might be something that is pulled out of this manuscript, saying this is where the gaps are.

FEMALE VOICE: So at the measuring level there might be a need to generate things for the patient that are specific to that individual.

CHOY: It's historical because of how those instruments were developed. There are many instruments that are quite old and they were way before the days that we

would be starting to move to a patient focus group and all those work.

(Cross-talk)

DWORKIN: So Jas and than Ajay.

SINGH: So if we go back to the previous slide, it tells us from the ICF model, participation is not a sub box of activity limitation. So there's an arrow that goes in both directions and therefore it is not a sub domain under activity or physical functioning. So what we are maybe defining with this model is a departure from that model, and do we have enough empiric evidence from patients to have the departure?

DWORKIN: So it seems to me that is relevant to what Bob and Philip are doing, in the kind of overarching framework where does participation with respect to physical functioning. What I want us to focus on, because we've only got an hour and a half to two hours left, is just the physical function part. If I am understanding Dorcas correctly, if we start with the physical functioning domain bucket, so let's say physical functioning domain, there are two large groups of aspects of physical functioning with that which is acts and tasks. And it sounds like we're not coming up with any other sub domains of physical functioning beyond acts and tasks.

BEATON: That is kind of how the ICF uses it.

CHOY: Aside from that, along what Jas has said, when we think about, for example, physical activity and actigraphy, you may not want to put it in participation, you may want to put it under activity, right?

DWORKIN: That's great, Ernest. So to Dorcas, we've heard a lot about actigraphy, not because there was any notion of promoting actigraphy but because it was more coincidental. So where would the actigraphy outcome that Dan was talking about this morning fit, is it an act or a task? At least a measure, a measure of acts and tasks.

BEATON: I would agree, it sounds like it is more like an act, a very valuable act.

DWORKIN: It's a measure of the person's daily acts.

FEMALE VOICE: I would suggest you have to first decide what your acts are that you're interested in in your population. So you go through an exercise of define a person, group A, and what are those acts, is it daily, you know, how far along that continuum of physical functioning are you measuring, do you want the whole thing or do you want the lower limb functioning, or do you want the upper limb functioning, or do you want jawbone movements or what are those acts and tasks in that population?

DWORKIN: Okay, so before we get to Ajay, the question is does ICF have a list of acts and tasks?

MALE VOICE: Yes, and extensive list.

BEATON: So there have been a lot of people who have studied the list and tried to map stuff or patient's perceptions of important items and they don't always fit in, so we have to allow for some flexibility to say we might need to add in more items to the ICF.

DWORKIN: It sounds like they would be generally relevant to patients with chronic pain?

BEATON: They are meant to be relevant to people with any disorder. Different ones, it's a whole classification system, so different tasks or acts might be more relevant to people with osteoarthritis in the knee, or with chronic pain or --

DWORKIN: So an act would be climbing stairs, a task might be going to the supermarket.

MALE VOICE: Or buttoning your shirt.

BEATON: Yep.

DWORKIN: Ajay, finally.

WASAN: So the other thing I was going to add is in the overarching framework from our discussions I found it really helpful to think about self reported physical functioning versus measured physical functioning. And

that's a basic branch point --

DWORKIN: We haven't gotten there yet because I'm going to call on Dan, now that we know, at least now that I know what we're talking about, to mention your framework --

WASAN: Okay, that's what I'm saying, if it's somewhere up there in the framework it would be really helpful.

DWORKIN: Yes. So we've got physical functioning and within that acts and tasks. Ian.

GILRON: To get back to your question or comment about actigraphy, I'm just wondering whether actigraphy is an intermediate measure or a surrogate of an act or a task because not every movement that registers on actigraphy is necessarily a meaningful act or a task. It's a surrogate and so some proportion of that is meaningful, but in and of itself it does not represent a meaningful act or task.

MALE VOICE: I think it depends how you use it because you can do a lot of stuff, right, you can just be an act, the act, say number of steps in a day, right, or you can actually say well how good is that exercise or a specific exercise task you can say. So that's a broad term, right, actigraphy for both acts and tasks depending on how you use your measuring devices.

MALE VOICE: It could potentially be a measure for both and I don't think it matters. It really doesn't matter. The issue is we're defining it as physical function and it's a surrogate or it's a measure of the activity level of the person wearing it, whether they're doing an act or a task. I mean the -- so I wouldn't worry about differentiating that.

DWORKIN: Well I'm going to defer to Laurie and Ashley, if we don't know whether -- if our concept of interest is physical functioning and there are two types of physical functioning, acts and tasks, and actigraphy measures some mix of box, doesn't that kind of compromise the interpretation of actigraphy as a measure?

FEMALE VOICE: I think we need to first decide what acts and tasks we're trying to measure. And then decide if actigraphy represents those specific acts and tasks.

DWORKIN: So what acts and tasks are we trying to measure?

MALE VOICE: If I remember when Dan Clauw showed one of this actograms, he showed activities and he had them tied to specific things like swimming and sitting.

MALE VOICE: So there was no measurement on the swimming part, he wasn't wearing the actigraph.

MALE VOICE: But I don't know how he came up with those characteristics but he had specific acts tied to certain levels of activity.

MALE VOICE: Because the patient reports them.

MALE VOICE: That was part of the diary and it was clear that the patient took off their actigraph during the swimming exercise.

FEMALE VOICE: They're not waterproof.

MALE VOICE: Well that's changing, there are waterproof devices.

MALE VOICE: But hypothetically you could take actigraphy and could assess specific activities, with this device what the range is.

MALE VOICE: Yes, and in fact that's what you do with a diary that goes along with it. Kind of like a Holter monitor where you jot down what you're doing at the time that you are having the arrhythmia.

BEATON: That's a good analogy.

MALE VOICE: The actigraph measures motion in three dimensions, period, back to my tremor example, you know, whatever, movement disorder. The activity is measured by the actigraph and knowing what those motions were doing, the diary or some observation or something like that. And so in my mind, and as I was thinking and sketching during

lunch, that's why I was making the point with Dan earlier, the actigraph just measures motion without any meaning. If you have a diary or you have an observer or someone saying during the time that actigraph was kicking up, the individual was doing this action.

MALE VOICE: But you could do a study in which you brought in X number of people, had them do a specific activity, take their actigraphy data and create some normative information about what is the typical range of activity for someone performing activity X.

MALE VOICE: I would just add that within the next five years there will be a compendium where you apply the signal data and you will know what activities they engaged in.

DWORKIN: So coming back to Ashley's question, what are the acts and tasks we're interested in assessing in a clinical trial of a pain treatment? But it sounds like that there are dozens to choose from and ICF has already listed them.

MALE VOICE: Thousands.

MALE VOICE: But the question is do they have them for the chronic pain persons that you are interested in?

BEATON: You would select from them.

MALE VOICE: You end up mapping from them.

MALE VOICE: But has that been done for chronic pain.

BEATON: Done for chronic arthritis.

MALE VOICE: It's been done for psoriatic arthritis.

DWORKIN: So I'm a little confused, if we have thousands of acts and tasks, how do we ultimately get to measures, we don't have --

MALE VOICE: So then what you do is you take those, let's pretend that you end up with 300 items from the ICF that map to, that are of interest, and then you put up your different instruments and then ask the question to what extent do these measure amongst these 300. And if it turns out that 80 percent of them are measured by two or three of your instruments then you have done a good job. But if only 40 percent are then your instruments are lacking.

DWORKIN: So it sounds, Dorcas, you said this has been done in OA, could you summarize the conclusion of what the OA effort was?

BEATON: They end up making a short list and a long list of relevant content areas for measures of OA or whatever disease group, so I'm trying to think of what the

core, everyone uses the term core, but it's the core set of items for pertinence to OA. Does anybody know, Vibeke or Lee, do you know?

STRAND: I don't either.

BEATON: Some of them are 17 items, 100 items, and so it can be used very nicely, almost as a content validity check of some of your instruments. It doesn't replace an instrument, it's just when you're saying when you're thinking about OA remember that these are the things that you should be considering, these are the items of relevance.

MALE VOICE: So you could select OA as an example to illustrate the point saying that this needs to be further developed for other conditions that have not at this point been developed. But at least it will be the description of how it was done for anybody who wants to look at this.

FEMALE VOICE: They've addressed an awful lot of the conditions.

MALE VOICE: But not all --

DWORKIN: For example, diabetic neuropathy may well have been addressed.

John.

MALE VOICE: Just to build on what is being said

here, the reason there are thousands is because you can cut them into tiny little pieces, like buttoning and unbuttoning, and putting on your shirt and taking off your shirt, and putting on your pants and taking off your pants. There are categories though of these that can be looked at and I don't know the ICF categorization as well, but in thinking about and in looking at, for instance, why the WOMAC was chosen to be the way it is, it was focused on leg function. And so it looked at climbing stairs, walking up stairs, walking blocks, et cetera.

The point that I think is being made is that for any given pain syndrome there may be a best group of these items, there is not going to be a single measure that works in all of them because they are going to be different. And if you want to be disease specific, you need a disease specific item.

Now there are general function measures that are out there, but they do what you would expect, head, arms, body, back, leg, feet. And they incorporate some of the items of each of those to come up with a composite score. That is unlikely to be adequate for a specific disease state.

MALE VOICE: When you do a disease specific outcome and a general outcome, then you can have the

informed --

MALE VOICE: Right.

MALE VOICE: Bob, I'm not sure I understand why a thousand acts and tasks makes you uncomfortable, I mean in terms of developing measures there should be no limit to our creativity and it's contextual to -- it's contextual to the population and the treatment. But if we're going to end up coming up with recommendations for specific measures like you did in 2005, then that's going to be based more on the studies, the measures that are available and how they have performed. So I don't know, I mean we can just comment that there have been thousands of acts and tasks defined --

MALE VOICE: But if you wanted to give a thousand questions to any one patient --

MALE VOICE: No, no, no.

MALE VOICE: I'm just saying that there are many acts and tasks that are important to people and that thus far this has led to the development of certain outcome measures and this handful has --

DWORKIN: I guess my problem, I was thinking of it from Ashley and Laurie's perspective, if there are thousands of acts and tasks and I want to figure out what measure of physical functioning to use in the next trial of

postherpetic neuralgia we do, how do I get from the thousands of acts and tasks to picking a measure of physical functioning for my clinical trial of PHN. Because I don't know that the article that Kristine and Ann are going to be writing will necessarily recommend measures, but I guess our hope was that it would be an article that someone would read and it would spell out for them the important considerations in determining how to assess physical functioning. And if I'm reading an article that says there are thousands of acts and tasks relevant to physical functioning and chronic pain, boy, I don't find that helpful in picking a measure for my PHN trial. So maybe I'm missing something.

MALE VOICE: Is this meant to be a guidance for measurement development or a guidance for implementation of available measures into trials?

DWORKIN: I think we have always hoped that the articles were going to be somewhat helpful to someone who is going home after they've read the article and designing a clinical trial rather than a long-term 10 year research agenda.

MEASE: So Bob, you don't leave that up to the reader, you leave it up to the authors to at least have checked with the ICF core list and then you then pick

measures that are extant and that have done a good job and ask the question to what extent do they cover the relevant task. And if it turns out that it covers the majority of them, then you've provided some content validity for --

FEMALE VOICE: I think, Bob, we really need a good example of marching through this thinking, of course, given that I'm thinking the way everyone else is thinking. Anyway, that you have a, we need an example, whether it's OA or whatever, ankle arthritis, ankylosing spondylitis or whatever it is, we need to have an example of this patient group that we can define specifically and then pick the acts and tasks that are relevant to them in a certain timeframe as outcomes.

Then we need to pick, this is ankle arthritis, here's the acts and here's the tasks, and we base that on explaining the whole ICF thing. Then we have to find the measures that we think are going to have a relationship in some way to that outcome. So that outcome is not the same thing as a measure and that's' one thing that is important about this diagram. These are daily life outcomes that relate to the acts and tasks. So the measures aren't necessarily measuring that thing specifically. If it's actigraphy the measure is measuring motion and how do we know that the motion is related to that ankle arthritis.

That is, under this relationship here, has to be at least theorized, if not empirically demonstrated.

MALE VOICE: But I think what you just said is the key, which is that there needs to be a demonstration that the measure is well correlated with the process of interest. That's up to researchers, that's up to Dr. Patel, that's up to the others here. Given the popularity of actigraph technology and the occurrence of the bands that certainly many of our kids wear, I can't imagine that that doesn't happen rapidly over the next several years with or without us. We can leave that, but the article needs to say, as we always say, that this has got to be demonstrated and proven to be an appropriate measure/predictor, whatever it is you want, of the outcome of interest.

DWORKIN: So Philip, you've been waiting patient.

CONAGHAN: I like the, by the way, of an example all the way through, I think that will make it much easier. I think, Bob, just to come back to your question, the ICF, as Phil said, gives us some element of construct validity for this, and if we think of that as the what we can sell from this is saying here's a modern conceptual framework that's been detailed across a range of functioning disorders, here's the construct validity for this tool, and you will also need to select all the other psychometric

elements, once you get down to tool level you have now got to understand the psychometrics of the tool and its performance across a whole lot of things. And just one other issue, there was construct validity again brought up.

So I think here we're moving from the big down to the specific tool and you're right, ultimately I think our goal is to come up with tools we can recommend but I think you are not going to be able to do it all in this process, it's going to be the start of a big of a research agenda or come out of this mapping for each of the areas so we know for OA that ICF mapping has been done, we'll say for other areas it hasn't, we can look that up. And then we get down to the level of psychometrics of the tools that you are going to use about function. And again, we know that's being done, there is a EULAR process going on at present that is mapping OA and RA tools for all their psychometrics and that will be finished this year. So there are various other resources we can point to.

So I think this is a worthwhile process for roadmapping for researchers.

DWORKIN: Ernest.

CHOY: So just to clarify, the ICF framework isn't intended for you to just pick this item and put in the clinical trial and out comes a score, that isn't the

purpose. For as a researcher, what I want to do is when I select an instrument to measure something I want to check that this is valid and so I want to know what it measures, it's covering enough areas in the ICF framework. The other issue I need to decide, for example, is sensitivity to change, how it performs in certain patient populations before I decide to use that instrument in my trial. So the ICF give us some information but the validity of the instrument you want to choose, but it isn't supposed for you to just take an individual item because there is no way of combining them. There is no formal way of combining the individual items.

DWORKIN: Lee.

SIMON: You might just consider the possibility in the manuscript in thinking about the instruments to measure the items, that one might consider applying the word parsimony because there is exhaustion associated with trying to measure everything and, furthermore, the research agenda should include that there may be aspects of this instrument that measures multiple items or gives you impact on multiple items, that's how the WOMAC was developed.

OSCAN, which is for the function of the hand in osteoarthritis, doesn't work very well, and he is now working on it to improve it and it's because it doesn't

have that same generalizability to a heterogeneous disease that we may not feel that the WOMAC is great, but so far it has remained the best parsimonious methodology to measure function, even in the context of the overlay of the ICF items.

DWORKIN: Sure. So I'm going to try and flesh out, see if I understand what Ernest was suggesting correctly. Let's imagine Roy Freeman and I are designing a trial in painful diabetic peripheral neuropathy, and some measure, we learned this morning we don't know what measure, is going to be our primary endpoint of pain intensity, but we want to measure physical functioning in patients with painful DPN. And so we were struck by Bob Kerns saying that we've got this multidimensional pain inventory that assesses pain interference with function. Is what you were suggesting we take the items of the Kerns and Turk NPI and see which acts and tasks it assesses?

CHOY: Yes, exactly what you do.

MALE VOICE: You map it back to the tasks and acts to see if, in fact, they are there within what you measure with that instrument.

DWORKIN: And Roy and I then decide if those are the important acts and tasks in patients with DPN, we just kind of make that decision on our own.

BEATON: There's a process that might be a choice on a research agenda for this group, of how you choose the items on the ICF that are relevant for a disease group, there's a mapping process.

DWORKIN: So ideally someone has already told Roy and me what the important acts and tasks are in painful DPN and then we check that the NPI measures those acts and tasks.

BEATON: Correct, but you also have to be, one more step is to be confident that the ICF was exhaustive in covering your patient population's experience. Because maybe in developing the ICF they didn't have a big enough list to capture these really important things. So I would say as well as doing the mapping exercise, you take the mapping exercise and show it to groups of patients and generate more important concepts that we're missing or maybe got taken out during the mapping process but the patients say are really important to have. Sorry, patients, people with pain I really should say, patients, I apologize.

TURK: But on Monday when you want to start your study you may --

DWORKIN: We're not there yet.

MALE VOICE: You may miss that step, Dennis,

that's correct.

DWORKIN: So the recommendation in the manuscript that Kristine and Ann are drafting is that for the major chronic pain conditions that are interesting, we need a mapping with patients and clinicians of the relevant tasks?

MALE VOICE: So that's a big process. What I think their goal is, is to perhaps use an illustration of an example of how this might work, show the relevant literature, and then show how it might happen. But this is the activity of, say, a fellow for a year to do this sort of thing.

DWORKIN: To do the mapping. Right, no, I mean there is a research that in our article from this meeting would be saying that these, the mapping of acts and tasks and whether the important ones exist for some conditions, like OA, but clearly for other conditions, it's a research agenda, not that we're going to do it. That we could --

MALE VOICE: There is the ideal and there's the parsimonious, to use Lee's term. Ideally this should have been totally gone through, but since I'm starting my study next Wednesday, I'm going to take the best I can in the process that's available.

DWORKIN: Ian.

GILRON: I just want to make sure I understand. I mean from looking at some of the systematic reviews that were part of our reading, I'm guessing that there may be a few disease conditions where we have validated measures that can be recommended for use in clinical trials tomorrow or Monday. And but that would be one part of this report and then another part would be on how to develop measures going forward which could be three to eight years from being implemented in trials.

MALE VOICE: Or how to check or sort of give reassurance about the comprehensiveness of the measures that are already in existence.

DWORKIN: So, Ian, I'm going to hold back on using my example of DPN, if no one has mapped the important acts and tasks in patients with painful DPN, then, and we're endorsing this framework, then I can't choose a measure based on this framework.

BEATON: I think that's one approach in how you could use the framework, so I think it's just one choice of doing the mapping and the listing of the items from the framework. Other people use the framework and use the chapter headings to say I'm going to talk to patients about what's important in these different areas of life functioning and these different groups of acts and tasks.

DWORKIN: So I'm going to try and pin you down, Dorcas, so Roy and I are trying to decide between the Kerns and Turk NPI and the Cleland BPI as our measure of physical function in the clinical trial we're starting on Monday. And we only want to use one measure of physical function, those are the two that IMMPACT recommended back in 2005, how does this help me and Roy make a decision between those two different measures?

TURK: Bob Kerns and I can tell you which to use (laughter).

DWORKIN: But we don't have Charlie Cleland here is the problem.

TURK: That's my point.

DWORKIN: Does this help me at all or -- (cross-talk) -- but Philip just said it's going to take a fellow a year to map it.

MALE VOICE: No, no, no, if you just take the two you mentioned, you didn't have a total mapping the way you want, but these are two potential measures, you look at the items in the measure and how close are they to the functioning that you're interested in, that's the best you can do right now. It doesn't mean you aren't better but that's what you can do now.

DWORKIN: So Roy and I decide what function we're

interested in.

MALE VOICE: And then you look at these measures and say do these measures, do the items, not the scale names, do these items measure what you want.

DWORKIN: That's the way we do it anyway.

Ian.

GILRON: I thought there would be some importance in terms empirical evidence of validity and reliability from validation studies. I would think that would, I mean content validation is important but if all other things being equal, that's how I would decide between the two.

DWORKIN: Of course, I'm just trying to see what the added value of this is in helping Roy and me the way we would have done it before, I'm sorry I'm speaking for you, Roy, the way we would have done it before is exactly what you said. We'd look at the literature on the relative reliability and validity and responsiveness of the Charlie Cleland measure versus the Turk and Kerns measure, but what I'm struggling with, if we're endorsing this and if the article that Ann and Kristine are writing is going to be focused around this, what does this add to what we would have done, what I would have done before this meeting?

Philip.

CONAGHAN: I think you're adding a conceptual

framework that is patient based. And it may be that we've been using tools for years that didn't fulfill that framework. So I agree that in your trial done tomorrow I'd do exactly the same as you, I'd go and look at the traditional psychometrics of those tools and rank them according to what I want. But I accept that this conceptual framework is the way we should be moving. So in part it is inspirational but it sets up a lovely research agenda for young researchers who care about outcome measurement for us to really base them on a proper conceptual basis.

DWORKIN: I think that's brilliant and I would like you to please take what you just said and make an email tonight, that paragraph, to Ann and Kristine, because to me that's critical, that's where this manuscript that their drafting goes beyond the 2005 article.

MALE VOICE: Absolutely right.

MALE VOICE: And in 2005 we weren't there yet, now in 2014 that's where we are. Similar to the problem associated with pain, itself, as a measure.

DWORKIN: And the additional thing that Dorcas and Philip are saying, and Laurie and Ashley are very much in agreement with, is that over the next five or ten years we need to do the mapping of acts and tasks for all these other conditions where it hasn't been done, so that at some

point Roy and I could actually map the NPI and the BPI to what is important to patients with DPN.

MALE VOICE: In your 35th manuscript based on this clinical study.

FEMALE VOICE: Now I have a question about the next step of applying this model. So we're going to map the acts and tasks from ICF to whatever person group we have here, but what if we really want to measure social integration, do we have to also measure the acts and tasks to make a conclusion that an intervention has an effect on social integration?

MALE VOICE: Isn't social integration partially related to acts and tasks?

FEMALE VOICE: That's what I think.

MALE VOICE: Of course.

FEMALE VOICE: So how can you interpret social integration in a causal attribution sort of way without knowing how that intervention affected acts and tasks?

MALE VOICE: You are talking about the development of any tool. That was I think what Bob Dworkin was saying before which is that how is this different than what we would normally do to create a tool to measure what we want to measure?

FEMALE VOICE: But what I've observed is there is

a lot of tools that are developed to measure things in this participation realm --

BEATON: Because even if you wanted to be able to understand it as a causally linked thing to the acts and tasks, you still need to be able to see it. You need to see participation, you need to see pain, you need to see task behaviors and see if you can build the structure equation model that helps you understand how they go together.

FEMALE VOICE: You just bring up another question, do we also in addition to physical function have to measure the symptoms like pain?

BEATON: Well it depends on your question.

FEMALE VOICE: Well --

MALE VOICE: Are you asking if you did a clinical trial or for pain therapy, could you only measure social participation? If you are interested, sure. You wouldn't have to measure the rest, but what would you do that for? Because one of the major aspects of the pain therapy should be the improvement in pain, but also it might also include social participation.

BEATON: But in people with pain there's studies that do, there's a study going on now called Making it Work, which is about helping people with pain to be able to

sustain their work role. And the primary outcome for that is not going to be pain it will be work participation.

MALE VOICE: But you are not studying pain.

BEATON: Exactly, it's almost like the background.

MALE VOICE: You're studying work productivity or whatever it is, that you're studying.

FEMALE VOICE: And you think you can interpret the work participation as data --

BEATON: Just as an indicator of work participation, not saying that it's an indicator of somebody's symptoms, it's how they're functioning in their job.

FEMALE VOICE: But you can understand that without knowing how their actual bending, lifting, balancing, things that are relating to their work, have changed, you don't care?

BEATON: If my intervention is at that level of trying to intervene and the next thing it's going to intervene on is work participation, then I just look at that. If I was trying to intervene on their strength, so that they could go back to work, I would want to measure their acts and strength.

DWORKIN: Ajay.

WASAN: Maybe I'm stating the obvious but it seems to me that this is what the framework helps you do, right, you go to your measure of social participation, you look at the original validation paper, you see that it has some process where you can map it to acts and tasks, and if not then it at least falls into back to what you were saying about how you would use a measure, so you can understand exactly what is the value of the measure in your population. Is it more specific just for that exact thing that's asked about versus has an effect been mapped in some sort of qualitative way at least, you know? So it seems to me that's a big point and a big improvement with the framework. That is one of the goals of the framework.

DWORKIN: The 2005 paper had no conceptual framework. So this is, as Philip pointed out, a conceptual framework within which to evaluate existing measures and that also guides the development of new measures that takes into account what acts and tasks are most important to patients.

So I guess I'm going to stop talking because I want to know whether Kristine and Ann are happy since I'm not going to be drafting the paper and you guys are going to be drafting the paper. So I'm going to sit down and you guys should start asking these guys questions so you have

everything you need to draft a paper because I'm going to go home and forget about this meeting completely

(laughter).

PHILLIPS: I think that one of the questions that I have is do we feel comfortable that all of the acts and tasks that we would be interested in are covered by the ICF framework was it currently exists?

CONAGHAN: Probably not.

MALE VOICE: No.

PHILLIPS: Okay.

(cross-talk)

PHILLIPS: Is there any enthusiasm among the group to fill in the gaps.

MALE VOICE: No, it's a big task, so between now and --

CONAGHAN: I think I'd put an exploratory bid in here, we probably need to have a look.

DWORKIN: So what will happen is we might come up with an example that's missed and then can use that as an example, as a stimulus for research agenda for more complete identification of missing items.

FEMALE VOICE: And sometimes what's missing, and I dare say this, is the contextualization of the activity, so walking will be there but people will say it's not

walking, it's walking on uneven ground, it's walking when I'm trying to rush to something, so sometimes it's that nuanced pieces about beyond walking, what part of walking.

CHOY: You already said that for the next meeting you are going to invite people from ICF to work with so this is the direct consequences of this discussion.

DWORKIN: Roy, and then we'll go back to Kristine and Ann.

FREEMAN: So just two very quick thoughts have crossed my mind. I think it was Philip who articulated the concept so very nicely and spoke about this as being patient based. And I wondered whether rather than patient based it should be disease based with picking up on the notion that DPN functional issues are going to be different to radicular pain functional issues. And so I think when you write this is probably should be within the framework of disease based rather than patient based.

And then without ever having seen or thought of this huge item bank of thousands, it seems to me that there should be something written about the process whereby items are chosen. And so there needs to be something with respect to how whatever item bank is developed from this huge repository, how we come to that decision.

MALE VOICE: One of the things, I liked the word that Philip introduced which was aspirational, which is that you absolutely, you proceed with what you have now and you have all the ways of measuring the psychometric properties of them, but there is a continual improve process that goes on, and one of the continual improvement parts of that is having patients either in focus groups or patient research partners, look at your various measures and speak to their adequacy.

So I think what Philip was saying about patient based was not to try to distinguish it from being specific about a disease, but just to ask patients whether or not we really are covering the disease from their perspective. And that's part of what takes time is that a fellow who is developing some of these mapping processes would ideally include in that activity some focus group activity.

FEMALE VOICE: Absolutely.

DWORKIN: Gary and then John.

WALCO: I would just ask Roy to clarify, when you say disease based are you saying that one should, like we're going to specify these different areas and domains, things that should be measured, but then specifically what you look at within those may be defined by a disease, but you're not suggesting that the entire structure start out

with disease.

FREEMAN: The latter. Well, let me say what I'm suggesting is that when you develop your instrument, which is a refinement of a thousand, that this be disease based. So that the issues with respect to function in a patient or cohort of patients who have low back pain are going to be totally different from those in a cohort of patients who have postherpetic neuralgia and --

MALE VOICE: You're talking about the specific items in the specific instrument, the domains would be the same that we're putting out regardless?

FREEMAN: Definitely.

MALE VOICE: Specific acts and tasks.

FREEMAN: So you've got a thousand instruments, how you're working, so I'm speaking about an array of instruments.

DWORKIN: John.

MALE VOICE: So two things, quickly. One is that there are other ways of thinking about this long list and Dennis mentioned it before, others have mentioned it, which is some of them are sort required parts of living that you have to do, and some of them are optional. And I think it would make sense at least to have a brief discussion in this paper about the fact that measures of, if you haven't

played the piano before measuring piano playing isn't going to be useful to figuring out whether you're better. It's an optional task. Whereas getting up to go to the store because you have to go get food, people do it with back pain, they do it with whatever. And so just that concept needs to be integrated in the selection.

The second thing is Jim Witter has been surprisingly quiet, and I just wondered what he might contribute from the PROMIS perspective on function.

WITTER: Well I was actually going to ask a question about gaps in terms of thinking about things like scleroderma, for example, or Dupuytren's contracture or really bad rheumatoid arthritis of the hand, or lupus can be the same thing. And in terms of looking for gaps and trying to figure out I guess tasks here, those who have looked at the list that ICF has which I haven't, does it include things like the things that we now do in our modern society, you know, texting and use of keyboards and things like that, I don't know if that's in there, just a question?

DWORKIN: That's a good point, it would have to do with hand dexterity, descriptive phrases and then the question would be whether or not we think that, depending on whether you would do it this way or this way is adequate

to cover it.

FEMALE VOICE: It's a good point to say we also have to refresh our concepts as all of this, the nature of work changes --

WITTER: And this is what Jim Fries has preached for years and years. But just to answer John's question, I think in terms of what we have for, contributor for physical function, I've mentioned this I think a couple of times, we follow the WHO-ICF framework, right, and we have the banks for adults and children. The adults are from, taken from the best of the best, right, and they've been improved and we're into upper and lower extremity but it is all patient report. And I think Ann, you had given the definition from Barron in your presentation, that's the PROMISE definition of physical function functioning and that's how we sort of look at it.

DWORKIN: So, Jim, the PROMIS measures have already been mapped to the ICF --

WITTER: Yeah, that is something we've been doing for a number of years now. I can point you to people that know more about this than I do, in fact, we just met some people from, it was, I'm trying to think, the organization they came from at the federal level, I'm blanking right now. But anyway, that was part of that conversation as

well.

So we've been doing this for a long time and, yes, the answer is yes, we've been doing that.

MALE VOICE: Sorry to break in, but one option then is just to say that we might want to emulate what PROMIS has done, right? I mean it's nice we have this framework but also we may not need to reinvent the wheel.

MALE VOICE: We should have mentioned that, we talked about that as being incorporated in that.

DWORKIN: So when I'm planning a DPN trial with Roy on Monday we not only look at the Kerns and Turk measure and the Cleland measure, we obviously should also look at the PROMIS measures and then we figure out which of those measures best assesses what we think are the relevant acts and tasks in painful DPN. I think that is an advance over what we have already published, because of what Philip said, that it really is a conceptual framework that guides the process.

Penney.

COWAN: I just want to go back to a statement earlier and talking about patient centered care which is certainly the watch word today, is that, you know, we do need to have those focus groups and ask them what is important to them as far as their function goes. Because

you don't know, you have no idea what it would be like to have maybe DPN, what does that mean to you, and if you give them this list, what's important to them. And I think that's where to some extent we need to know what's important to them. You guys can only guess.

DWORKIN: It's a whole other conversation, but, you know, physical functioning is important in our action, the American Pain Society Pain Taxonomy, and there is a way in which this mapping process for the conditions where it hasn't already been done, could be incorporated into the act effect, because one of the dimensions in our multidimensional pain taxonomy is physical functioning and psychosocial functioning. So there is a possible marriage there, assuming there are enough resources.

Bob.

Kerns: Just a minor point of semantics really, I think it may be helpful for us to actually in the context of the discussions we've been having about, talk about people with DPN, DP, NP, and opposed to DPNP.

DWORKIN: There are some old dogs that have trouble with new tricks.

Ian.

GILRON: So I just thought that a lot of this is conceptual and I was wondering whether there would be some

benefit to do a limited focus systematic review of maybe in a specific area or specific measures, but of clinical trials that have used measures of physical function and how they've performed, how many times are they used as a primary outcome versus a secondary outcome. I don't know if this is like a handful or this is a lot, but I mean just to illustrate to people who this can be done. Because it sounds to me like it's relatively novel beyond using BPI and NPI, that otherwise a lot of this conceptual framework stuff might get lost on people and sort of say well this is something for the future, is this something that we can really think about doing tomorrow. And it would be nice to know, you know, like if you picked actigraphy and some occupational measure that has been actually used in clinical trials of interventions that are targeted towards pain, to sort of illustrate how this might work.

So ACTION would have the resources I think to support that kind of systematic review if someone could be identified who wanted to do it. Whether it would be all chronic pain conditions or whether we would zero in on the major neuropathic and musculoskeletal where it hasn't already been done, you know, could be determined. But yeah, I think that would be something that would be worthwhile

--

DWORKIN: Kristine or Ann, how close are you to being able to do what Ian has just asked about, for example, in the case of hand OA or OA as a pain condition?

PHILLIPS: When you say how close, do you mean to publishing or to --

DWORKIN: To tackling, showing, doing a review of, say, OA pain studies and cataloguing the function measures in those and then speaking to the properties or capabilities of the measures?

TAYLOR: I mean a couple of them have already been done, there's three that have already been done in hip and knee OA, all of those are (indiscernible) reviews. There is one on neck --

GILRON: Can I just ask, are those reviews looking at validation characteristics or how they perform in clinical trials of interventions?

TAYLOR: No, they are actually looking at the specific tools, not how they function in clinical trials.

PHILLIPS: The measurements of properties.

TAYLOR: The measurements of properties, sorry.

GILRON: Not assay sensitivity, but just to illustrate this is the current status quo on what's been done. I was thinking it would be something that would fit into this paper but I mean it doesn't have to be exhaustive

but more illustrative, that this has been done and it can be done and this is where we are and this is where we want to go.

DWORKIN: So what I'd like to propose is Philip raised his hand, and once Philip asks his question I think we should take like a ten minute break, because I want Kristine and Ann to really have a ten minute break to think about if they've got everything they need, and then we'll reconvene at three, but first Philip is going to ask a question or make a comment. We'll reconvene at three with the sole purpose of making sure that before we all go Kristine and Ann are satisfied.

Philip.

CONAGHAN: Just a quickie. I mentioned before EULAR being involved, so there's a task force under Alan Tenant that currently has catalogued the outcome measures for RA and OA trials and is doing detailed psychometrics of each of those measures. So I think there is some stuff we can lift from elsewhere without having to do it all again. And there is also a parallel toolbox project that I'm not sure in how much detail that's going to into the psychometrics of tools, but that will give the sort of an answer to, at least for just OA tools and some of what you have asked.

DWORKIN: So a 12 minute break for coffee, we come back at three, and I'm going to turn the podium over to Kristine and Ann.

(BREAK TAKEN)

DWORKIN: So I kept rudely and tactlessly interrupting Dan Carr when Dan was making a point, and I'm going to call on Dan because I think now the point he was making before is really to me applicable in terms of thinking about the conceptual framework we have of acts and tasks. And so after Dan makes his point I'm going to turn the rest of the meeting over to Ann and Kristine so that they have everything they need to begin drafting a manuscript with help from everybody in this room, more help from some than others, and then we'll be able to call it a day. So Dan.

CARR: Okay, thanks, and apologizing for reiterating some stuff but I'll go through this and just leave you with the notes that I wrote down. So this more speaks to how to organize a manuscript than how to develop grand concepts and biometrics and so on. But a narrative strain at some point in the manuscript I suggest might derive from the consensus in the pain community that clinicians and researchers want to have pain reframed as many other mainstream conditions are in biomedical science

that we care about.

So if you look at the quality of pain trials, for years before this effort, they just simply did not measure up to the quality of oncology trials or ID trials, both in terms of power, number, design, so we want to bring pain more in the compass of how we deal with other diseases and conditions.

In those other circumstances there are different models or explanatory structures that are applied when observing or explaining phenomena at different scales. I'll get to that in a minute. So with those ideas in mind, just to kind of summarize a lot of our discussion, I was proposing a three row, four column table that wouldn't have arrows and would be deliberately stated as not being hierarchical but rather just presenting several different constructs or models according to the scale or nature of the thing being observed or measured.

So trying to synthesize information from the last couple of days, one level. So the rows would be physical, social and emotional function, as a straw man, and in the columns, the first column could be immediate, real time, continuous, ecological, objective, and that was inspired by the talk by Dr. Patel on actigraphy. But there are analogies, like if you did continuous oxygen monitoring,

continuous glucose monitoring, blood pressure and so on.

Another column would be patient recall based. So in an analogy with pulmonary you could say how many flights of stairs can you climb without stopping or have you had a lot of difficulty. There's lots of analogies, I don't want to get stuck on them. Third would be a clinical test, which has an element of objectivity but is not naturalistic. Analogy with pulmonary would be a pulmonary function test where you let all the air out of your chest, you take the deepest breath in and you let it out and that's measured, that's a vital capacity. That's a test, it's done by the patient, but unless the patient were asked to do that test they would never have actually done that in the prior month.

And level four I'm a little fuzzier on, we have to think this through, but it would have to do with some type of clinical assessment or appraisal of the nature that we have been talking about intermittently to carry the pulmonary analogy further. It might be if the examining clinician takes a stethoscope and listens if there are wheezes or so called rales or something like that. So there is an interaction between the observer and the patient. So it's just a table that I would see somewhere in this, it's not a dominant thing, but it's trying to draw together some

threads of the discussion.

DWORKIN: So whether we're talking about acts or tasks, those are four different ways of thinking about the assessment of those acts and tasks, and it's comparable, loosely related to what I think it was Ashley presented yesterday in terms of performance measures, observer outcome measures, clinician outcome measures and patient reported outcome measures, so that it's not exactly the same, particularly the first column in your table, but it's a kind of taxonomy of different approaches to assessing the acts and tasks that we talked about before the break. So thank you, I think that is very helpful.

CARR: Not meant to be comprehensive and it's a straw man, but I think it does tie together some of the things Ashley --

DWORKIN: And that's now down at the level of what we have, the kind of physical functioning, we have the domains, the acts and tasks, what are the different approaches available for assessing, typically in a clinical trial, whatever acts and tasks we're interested in.

So as promised, I'm going to sit down and turn the meeting over to Ann and Kristine and tell us what you need.

PHILLIPS: I'll start off. Thank you very much.

I think that Ann and I are going to meet this evening and draw up an outline for the manuscript. We'll then flesh that out over the coming days and then circulate it to everyone here for comments. One request that I have is that when you review the manuscript, please provide specific recommendations, I've done this a lot and one of the things that's very hard to address is when people write vague comments like needs to be, you know, expand this or something like that. If there is something that's missing, please tell us exactly what you think is missing so that we can provide it.

I think in terms of what we've talked about this morning, we'll start off by sort of providing some definitions and a taxonomy and some background, perhaps including in figure one some of the ideas that have been discussed this morning. We'll then talk about a roadmap that would be useful for investigators in developing an outcome measure for a clinical trial. It would be lovely if we had the time to actually walk every instrument through this process and come up with an exhaustive list but I think we are not going to be able to do that, but we could provide the framework that then someone who has an enthusiastic fellow or a research group that wants to tackle that could then use this as a roadmap to guide them

in what they're going to do.

I think the other idea that has come up from Lee is to talk about, you know, what's the most parsimonious way to use this roadmap, you know, how can you get to what you need in the here and now versus what would be the ideal situation, alluding to the fact that some of the instruments that we used were not developed in the modern era of thinking about patient input and that sort of thing.

I think that's it. Ann, do you have any more

--

TAYLOR: Yeah, I think we would include some strengths and weaknesses of the tools that are already out there and where we can we will highlight tools that have already been through a rigorous process, a systematic review, so it's clear where methodological issues exist with some of the tools that are available, and we can give some examples of most commonly used tools that have been used in clinical trials.

PHILLIPS: I think the other thing that we should mention in the paper is the potential gaps for acts and tasks, you know, in mapping something back to the ICF framework. This is an area for future research or future exploration that needs to be addressed. And I think just

pointing that out might be helpful.

TAYLOR: The other thing I think that would be very important, especially with the feedback we've been given today, is actually looking at how we can involve people with living with pain and actually directing what physical function outcome measures would be useful to engage people living with pain in clinical trials. And we could maybe suggest some of the methodologies. And not just focus groups because I mean they're quite investigator heavy, but there's other techniques that have got evidence base behind them for use and we could actually include that in our roadmap.

PHILLIPS: Can I just, one of the questions that I still have that's lingering that I think you agree with me on is whether actigraphy as being a measure of physical activity should fall under physical activity, and whether physical activity, how that relates to physical function. Could I hear someone, some comments about that? In other words, if we're thinking about physical function measures, patient reported outcomes, measurement properties of performance measures of physical function, should actigraphy be under that or should it be in a separate category in terms of when we talk about the taxonomy? I'm looking toward Dorcas.

CONAGHAN: I think I would be thinking it's under the physical, under the -- I guess if I keep thinking back to that conceptual framework of capability, to me that is one of the capabilities. And I think we should be very careful, my doyen of measurement science delved into my head many times that you've got to be careful about measurement and explanation. Measurement is where we can, I'm going to come back to Jas' shoulder example, we can measure how high you lift your arm, but there's a number of reasons why you might be fatigued, you might have a whole lot of reasons why you can lift it, not just subacromial bursitis, but that's explanation, not measurement.

So I think what you want in that, in your first function, and it's sort of the first column of what's up there on the board, is a term capability, is that what you had in mind, Dorcas? To me the actimeter and some of that interesting stuff falls in that basket.

PHILLIPS: Good, thank you. Are there any other gray areas that anyone wants to comment on? Yes, Dorcas.

BEATON: I wanted to point out that the, we found on the ICF website that there is a core set that's been done for chronic widespread pain which might be a group, a core set that is helpful for you in your review.

PHILLIPS: Okay.

TAYLOR: Sorry, one question, when we're focusing on the manuscript, do you want us to focus on pain interference or function in general?

PHILLIPS: That's a good question. Could we include a paragraph about pain interference to acknowledge it?

MALE VOICE: Yes.

PHILLIPS: Okay. Is everyone comfortable with that, do we have any --

MALE VOICE: You have to flag the issue.

PHILLIPS: I think we need to acknowledge it.

MALE VOICE: It needs its own subheading almost.

PHILLIPS: Right, is everyone comfortable with that? Okay.

MALE VOICE: But I would put it as part of function, in general, I wouldn't highlight it as sort of the key functional domain that we look at, because as you heard, there is a lot of controversy over what exactly does pain interference mean and the correlation is actually only .4 to a lot of other functional measures so you are in a, like I said, murky zone.

BEATON: There might be an overlap between an impairment and a functional test.

MALE VOICE: Right, there is controversy over whether it is actually a measure of function, right, because of all these variable correlations to other functional measures, so.

TAYLOR: I think it would be worth putting in a table of some kind of definitions, not just physical function and physical activity but some of the other words that have been used, and maybe through the process of sending around the manuscript, that's a way we can reach consensus about how we feel for the definitions of some of the other terms that we're using in the manuscript.

PHILLIPS: Ian.

GILRON: I don't have a specific interest myself, but I just, it was mentioned and so I was wondering whether there is expertise in the room to talk about measures of sexual dysfunction and whether we need to invite somebody to collaborate.

PHILLIPS: I think along with that, how to address caregiver burden, work participation, social participation, all of those things. And we can certainly include it, we didn't talk about it at this meeting but we can include it in the discussion. Does that address it?

Yes, Jim.

WITTER: If you get into that aspect of sexual

function, the person to talk to from the PROMIS perspective, and they've done a lot of work on that, would be Kevin Weinfurt at Duke.

PHILLIPS: Okay.

DWORKIN: So additional feedback, questions, comments for Ann and Kristine?

TAYLOR: I'm resigning when I get home, by the way (laughter). I'm joking.

DWORKIN: Laurie.

BURKE: Well when you say that you're going to make a list of pros and cons of certain instruments that makes me worry a little if unless you're really going to put the pros and cons within the framework that we're talking about. So, for example, it would be really good to know what acts and tasks each, or social integration concepts each instrument is measuring. Because unless you can identify what the instrument is measuring it's going to be hard to identify the pros and cons.

TAYLOR: I'm sorry, that's my poor explanation, I just meant top level, you know, why would you use a generic measure as opposed to a disease specific measure as opposed to a site specific measure, as opposed to a pain interference measure, what are the strengths and weaknesses of the kind of top level rather than going into tools

themselves. Sorry, that's what I meant.

FEMALE VOICE: Are you going to go into the idea of starting with a definition of the context abuse, the person, group, the disease, as Roy was talking about, important disease subgroup or whatever?

PHILLIPS: Yeah, it seems like the second or third paragraph will be a good place to put that, we'll be sure and include that in the outline.

DWORKIN: And that is also, of course, a critical issue in the paper that you're working on with Bob Kerns and Philip, and the two of you are working on, yeah.

Peter.

TUGWELL: Could you just give me an update on your most current thoughts of exactly what the title is going to be for each of the two manuscripts?

PHILLIPS: For each of the two manuscripts, I thought that ours would be more of a roadmap, we have an NIH roadmap, it seems that we could have a chronic pain outcome measurement roadmap for how we get to, how do we accomplish what we want to accomplish. Does that -- okay, that's one.

TUGWELL: I thought there were going to be two, did I miss something?

DWORKIN: So I don't know what Bob -- Philip, do

you have a title for your manuscript, there you are, let me off the hook?

MEASE: No (laughter). Yes, we will shortly. So you're --

DWORKIN: How about the assessment of functioning?

MEASE: Conceptual framework.

DWORKIN: Conceptual framework for functioning in analgesia in pain clinical trials, is --

MEASE: I love it, Bob, you're quick on your feet at 3:30 in the afternoon.

DWORKIN: This would be more of a roadmap for physical functioning.

MALE VOICE: Do you have an acronym for that (laughter).

DWORKIN: Not yes. But actually, Peter, in the spirit of consensus, if you've got two titles it sounds like Kristine and Philip would be happy to entertain your titles.

TUGWELL: I was just wondering how I was going to describe this to my wife at dinner tonight, I like to have very clear, crisp, one-line messages.

DWORKIN: Any other questions for Kristine and Ann? Kristine and Ann, do you have more questions? The

group is getting a little rowdy. Ajay.

WASAN: Just one tiny thing that popped into my head is in press now, in *Journal of Pain*, there is a nice paper on core outcome measures suggested by the NIH Task Force on Chronic Low Back Pain. So you just may want to eyeball that so that we're all on the same page and there's no inconsistencies and things like that. Because that may help with, you know, along with syncing what our framework on PROMIS has been, I think that would really go a long way to bolster the manuscript.

MALE VOICE: PROMIS is heavily imbedded in that.

WASAN: Exactly.

PHILLIPS: Okay, that's great.

TAYLOR: I mean I think there will be a section actually on identifying resources for people that are helpful as part of the roadmap, you know, a good reference, this would be something that I think we could all collaborate on.

MALE VOICE: Not being facetious about what I just said actually in thinking about that paper, the purpose of the paper that you just heard about was to come up with a minimal dataset that people should consider in all studies of low back pain, not that you can't add other

things, but the idea being that we can advance the field if, in fact, people use some consistent, not the measures, themselves, but consistently gathered certain types of information which in some sense is analogous to what we're trying to do, can we get people to start using some things with consistency so that in fact we can aggregate across studies and advance the field more rapidly.

TAYLOR: We also thought there would be some kind of action points for clinical trials and some action points for actually just general clinical issues to do with outcome measures so we would have those actions points or recommendations as well at the end.

PHILLIPS: What timeframe would you like to have as a goal?

FEMALE VOICE: Dinner tonight is fine.

DWORKIN: Kristine, as soon as you can get to it without kind of breaking your back.

TAYLOR: I mean looking at our schedules, we would probably start working at the end of May on this, with a view to getting a draft after that.

MEASE: Another plea back to all of you, which is when you receive this draft, whenever it happens to be, in addition to providing the more detailed kinds of comments, I mean nothing worse than somebody saying needs work,

whatever that means. But also trying to turn it around in a reasonable timeframe so that whenever you've done a manuscript, if any of you have been involved with writing with 20, 30, 40 authors, it takes a horrendous amount of time to get things done. So try the best you can to give some kind of reasonable timeframe and try to stick to that within some reasonable limits because it really does make this a very slow process if everybody drags this out for a long time.

DWORKIN: Okay, are you both okay?

PHILLIPS: Yes.

DWORKIN: Well I think we're done, thank you all very, very much for a great IMMPACT meeting. You will be hearing from us about two manuscripts and the next IMMPACT meeting on how to assess pain in clinical trials, 2014 or 2015. And have safe flights home and they're not here, but we are in enormous debt as always to both Valorie and Andrea outside for coordinating this meeting. And they've done all the other ACTION meetings so we appreciate them, and say thank you to them on your way out. Thank you all very much and safe travels home.

TURK: Thank you to our OMERACT colleagues who worked with us.

[END]

