ACTTION - IMMPACT XX - Assessment of Pain Outcomes Clinical Trials of Chronic Pelvic Pain and IBS

July 13, 2017

A Matter of Record
(301) 890-4188

ACTTION

INITIATIVE ON METHODS, MEASUREMENT, AND PAIN ASSESSMENT IN CLINICAL TRIALS (IMMPACT-XX)

Recommendations for the Assessment of Pain Outcomes in Clinical Trials of Chronic Pelvic Pain and Irritable Bowel Syndrome

Thursday, July 13, 2017
8:13 a.m. to 5:02 p.m.

Westin City Center
Washington, D.C.

C O N T E N T S

C O N T E N T S (continued)

P R O C E E D I N G S

(8:00 a.m.)

DR. TURK: Good morning. Hopefully, most people have had a chance to get some coffee, breakfast snacks, and be here. My name is Dennis Turk. I'm glad to be one of the people involved with this particular meeting, as well as others that you'll be meeting along the way.

I want to thank many of you for coming to this meeting, all of you for coming to this meeting, especially our colleagues from Europe, who make the long trek, as they have done for many other IMMPACT meetings.

I just want you to know this is the 20th anniversary of the IMMPACT meetings. This is the 20th IMMPACT meeting. And I'll give you a little bit of background about what IMMPACT is and what ACTTION is for those that don't know.

For those of you that have been to previous meetings, you can do your e-mails or do other things because you've heard some of this before. So before I get started with some more formal

Page 5

presentation, I really want to give you some housekeeping details so you'll be familiar with these, and this is things to keep in mind.

First of all, even before we get there, note the microphones in front of you. They are quite sensitive. They are voice activated. You don't have to push any buttons to turn them on or turn them off. When you speak, please give your name because this is being recorded, so be careful how you whisper to your next-door neighbors and what you say about people because we're going to know who you are.

Those minutes of the transcript will be made available on the ACTTION website, so anybody who's interested, who was unable to attend the meeting will have access to those. I'll also say that, of these speakers, we will ask them, with their permission if they're willing, to make their slides available to us also for us to put up on the ACTTION website.

What we've noticed in the past is some people like me stick in a cartoon or two that are

Page 6

copyrighted, so they don't want those on there, so they may delete those. But it will all depend on whether there's any proprietary information that people have. But the idea is to make this as transparent and as available to everybody since there are approximately 44 people at this meeting and there are many more people who are interested in this topic.

So let's just go to the housekeeping. Make sure you sign in on the registration. And you'll have to do that on both days, and sign in and sign out so that we know that you're here. Cell phones, put them on silent, please. There's nothing more distracting. Those of you who are speakers and many of you who are going to be speakers know that, when a cell phone goes off, put it on vibrate or turn it off totally.

I mentioned about the microphones. They're voice activated. Say your name first. The way the microphones are set up is that if a lot of people want to ask questions, once five lights are on, it will not let anybody else get in. Then, when

Page 7

someone is done speaking, then they'll be able to get back on.

So it's a maximum of five at a time, but as the lights go off, then you can come back on. But if you notice there are 5 and you're trying to speak and nothing's happening, that's because they're set up. They're activated that way.

The restrooms are out this door to the left, to the right. Go left half a corridor or quarter of a corridor and then to the right on down there. Other things, lunch is going to be served in the Vista Terrace room, which means you have to use the other elevators versus these to go to it.

For those that were here last night for dinner, it's the same room, I believe, that we had dinner in last night. It's on the mezzanine level in case you don't know, so when you go to the other elevators, go to the mezzanine level, the Vista room. There will be a sign there for you to see it.

So you know that check-out on Friday is going to be at noon, so please make sure that one

Page 8

of the coffee breaks and one of the opportunities that you have, that you check out on time so that you can get out of here. However, let me caution you that, although check-out time is at 12:00, we won't leave this meeting. We're going to lock the doors. You can't leave until we have developed a consensus about some recommendations that we can make to improve the field.

So although you're going to be checked out of your room, you can't leave the hotel, you can't leave this room until we have a consensus of what those recommendations are going to be, at least in a draft format. So therefore, don't try and escape and don't think you can get out early if we finish things, because we haven't finished them until we say we've finished.

Taxis can be ordered for the airport and the people at the front desk, Valorie, who is sitting in the back there with the blonde hair, and Andrea, who's outside right now, they can help you if there's any problems with that, help you with shared taxis. Any assistance, they are the people

Page 9

who can really help you when it comes to logistics.
If you haven't met Valorie, she's sitting there right in the back. She has the blonde hair. They both have blonde hair, so Andrea and Valorie both have blonde hair.
These meetings, as far as the organization and the logistics, couldn't happen without Valorie and her group putting things together. So we thank Valorie and Andrea for all the work that they do. They are extremely helpful. Any questions you have for those that have not already been involved, they can help you out on many things you want.
So that's it as far as the logistics. Make sure you say your name even though you think people may know who you are, because the person's who's transcribing, sitting in the back doesn't know who you are, and she can't always see your name tags.
If you don't have a name plate in front of you, by the way, you should make sure you have one so that people can see who you are. That's it for the housekeeping.
Why are you here? What is this? Well, I

Page 10

mentioned to you that this was the IMMPACT meeting. It is the 20th. And what you're here for, in case you're not sure why you're here, this is what the meeting is about, Recommendations for the Assessment of Pain Outcome and Clinical Trials of Chronic Pelvic Pain and Irritable Bowel Syndrome.
If you are not here for this meeting, now is the time to leave. Okay? So make sure you're in the right place and everybody knows that's what the meeting is about.
Now, you've heard some acronyms. So what is IMMPACT? Now, for some of those people who have been here before, they've seen some of this. But what IMMPACT is not; this is what they're not. They're not the International Micronutrient Nutrition Prevention and Control Program. If you're here for that meeting, it's down the hall.
They're not the Interactive Mass Model Proximity and Collision Testing Organization. You're not here for that.
We're not the Immigrant Public Action Coalition of Trenton, New Jersey, although Governor

Page 11

Christie has pushed that we should be part of that.
We're not the International Maine Maritime Potato Action Team, although Shannon does come from Maine, so she's also promoting that, the organization.
We're not the Infrastructure Management Mapping planning Coordination Tool. All these are IMMPACTs. But that's not what we are.
We're not the double impact taekwondo, although it sometimes feel that way, and you're going to notice that we do tend to focus and do things like that to get people to work on these things.
So that's what we're not, although that picture of Bob Dworkin at the bottom there, in case you don't know him, he's sitting over there, he does have a way of getting people to cooperate and work together. So whenever we see any debates, discussions, disagreements, we bring up Bob, and he will take care of you, take care of all issues. He'll resolve those.
To note, there's a coveted award, IMMPACT.

Page 12

It's a nice acronym. The Dworkin Award, he actually got the award, the initial award, for the most tortured IMMPACT acronym. And this award went to the 1916-1917 recipient for the In-Hospital Mortality for Pulmonary Embolism Using Claims Data, IMPECD [ph].
If you have someone you'd like to nominate for next year's award, we'd be more than happy to put them on the list, or if you'd like to come up with your own acronym, feel free to do that.
So what is IMMPACT? We know what it's not. It's the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials, I-M-M-P-A-C-T. The logo is on the left, and that's important because if you want to find out more about IMMPACT ever, and if you go to Google, make sure you put I-M-M, because if you put I-M, you're going to see impact, the normal spelling. You're going to find out all kinds of unusual interesting things that will entertain you, but not who we are. So that's what the organization is.
It's an international consortium of

Page 13

participants from academic research, governmental agencies, the U.S. FDA, U.S. NIH, U.S. VA. We have representatives from the EMA periodically, industry, consulting and research organizations, and consumer advocates.

So that's who we are. It's an invited meeting. We try to bring people to represent all the different disciplines that are relevant to specific topics. We try to make sure that we involve the appropriate regulatory people from the right divisions or the right organizations when possible. We thank them all, those that are here, and many more tend to drift in late because of traffic in the Washington, D.C. area, but we welcome them and appreciate their support.

The mission of IMMPACT is to suggest -- and that's suggest. We have no ability to dictate, to require, to make mandatory. We can only suggest methods for improving the design, execution, interpretation of clinical trials for pain.

So that's what we're trying to do. The idea is not to promote any products or any biases of our

Page 14

own type, but just to improve the quality of how studies are done to in fact get to the point where we in fact can try to improve the quality research, expedite the research, improve the speed with which new treatments come along. So that's who IMMPACT is.

IMMPACT is part of ACTTION. Remember, I told you about acronyms; so here's another one for you. It's the Analgesic, Anesthetic, and Addiction Clinical Trials, Translations, Innovations, Opportunities, and Networks, ACTTION, A-C-T-T-I-O-N. You'll notice there's a theme here, double letters, to try to make sure that you can find us because if you go to Google and type in A-C-T-I-O-N, you will have a hard time finding who we are.

What does ACTTION do? And it's ACTTION.org. And for those that are not familiar with it, go to our website to be able to find out as much as you could want to know about us. But ACTTION is a public-private partnership with the United States Food and Drug Administration. We appreciate

Page 15

working, cooperating collaboratively, and the support from the FDA.

The mission of ACTTION, little bit broader than just what we saw for IMMPACT, is to identify, prioritize, sponsor, coordinate, and promote innovative activities with a special interest in optimizing clinical trials -- you'll see where IMMPACT fit within this -- with a special interest in optimizing clinical trials that will expedite the discovery, development, and improved analgesic, anesthetic, addiction, and peripheral neuropathy treatments for the benefit of the public health. That's what ACTTION is. IMMPACT is one program, one initiative within the broader ACTTION initiative.

Who is IMMPACT? Who is involved with these things? I've sort of alluded to this already, but so far, over the 20 different meetings that we've had, we've had 200 participants, some of whom have been to multiple meetings, so over times.

We've had people from academic and related participants from 12 different countries, four

Page 16

different countries here at this particular meeting, countries as far away as Australia, Belgium, Canada, Denmark, Finland, France. You could read those for yourself.

So we have people from lots of representation, although, predominantly, the people in the room, because of the ease of getting here and cost, are from North America. But we do have people -- and thanks to Ralf Baron and to Katy -- I can't see where -- she's hiding in the back -- that came over from the U.K. Thank you both.

There are representatives from 90 different academic institutions, whether they're university based or hospital based, they're all academic and they all have scholarly interests, so we have tried to arrange for lots of them to be here.

As we've mentioned, we have participants from different governmental agencies, including the Department of Defense, the Drug Enforcement Agency, EMA, FDA, National Institute of Health, SAMHSA, and the VA. And we thank the VA for being here, for your help. Thank you for supporting us.

Page 17

These are people who are all involved in understanding these different things that we're doing.

We've had support for different meetings, different numbers, different organizations, different companies for from over 46 different pharmaceutical companies since the beginning, and in the future, we may also be having support from device manufacturers.

We have consumer advocacy representatives. We've had five different organizations here. Chris Veasley, thank you for being here, trying to keep us all not forgetting who the end user is going to be, which is the person that we're developing these treatments for.

We could easily get lost up in the methodology, data, and analytic approaches, and the outcome measures, but we have to remember that these are all geared toward an end user. So we try to have people represent those individuals here as well. We have several private consulting organizations who have representatives who attend

Page 18

here, who are interested in analgesic trials.

From the FDA -- and I'm not going to read these off to you, but we've had some different divisions from CDER and also from the Center for Device and Radiological Health, the Office of the Commissioner, all of whom have been at different meetings. And obviously, depending upon the topic, different people would be at different meetings, and not everybody's going to be at the exact same meeting.

From NIH, we see lots of different institutes. I'm not going to read them for you, but we've had representations from a lot of the different NIH institutes. We've also had from the Rocky Mountain Poison Control. We've had representatives from the EMA, as I mentioned. And we're happy to have someone from the Critical Path Institutes, Stephen Coons, who's in the back, who's at this particular meeting.

The idea is we want to involve -- the message I'm trying to give you is we try to involve the relevant people as much as possible to help us

Page 19

get to our objective, which is to improve the kinds of studies that we do to get to help people who have various chronic pain problems. For ACTTION, it's more than just pain. It's also anesthetics, and peripheral neuropathy, and addiction.

What do we do? Well, since 2001, I told you, we've had 20 different meetings. You're not going to read these now, but just to give you a flavor, those that want to know this, you can go to the IMMPACT or the ACTTION websites, IMMPACT.org or ACTTION.org, to see what's going on at these different meetings.

We attempt to put up all slides when they're available from the different meetings, the presenters. You'll notice that we have presenters, if you look at the program, and then we have lots of discussions. The intent is more important things in my opinion goes on during the discussions, not only the formal discussions and the panels, but also over coffee, and over dinner, and over lunch. So we intentionally build in lots of those things.

Page 20

But these are the first 14 different meetings of some of the topics. If you want to know more about them, you can see them. The slides when possible are always put up on the website.

Every one of these meetings, almost up to the last one so far, we have produced manuscripts that get published describing the consensus and recommendations.

Remember, we can't require. We can't mandate. All we can say is, based on these groups of individuals who all contributed to the discussion, these were the best recommendations we could come up with as you're thinking about either designing your clinical trial, the outcomes you may be using, the kinds of data analytic strategies you may be using, and how you go about involving patients.

We've got a lot of different topics that are coming up. The meeting that you're at today, obviously, is this specific one.

Notice by the way that one of our meetings was co-organized with OMERACT, for those of you

Page 21

that are familiar with OMERACT. And that was an assessment of physical function because the people in the rheumatology world were equally interested and dedicated to working in this area, so they co-collaborated with us on this.

We've had one meeting on pediatric pain, so there was one of those consensus meetings. Predominantly, we've been with adults, but we did have that one meeting. And you'll also note when you go and see the manuscripts, we have multiple authors.

Everybody who's at the meeting is invited to be an author. They can decide, yes or no, they want to sign on. To date, other than some people from regulatory agencies who have not been able to for different types and purposes, all the academics, all the other people who have attended have all been authors on these people.

So what you'll notice is the manuscripts, when you see them, often have 40 authors on those things. So congratulations. Depending upon the alphabetical order, you could be high up. I'm

Page 22

changing my name to Aaron Aardvark, and I expect I will be the first author on all these papers. And poor Sanz and Bardow [ph] is not going to be able to do this so well. He'll be the last, so it'll be okay, but those in the middle understand that.

The way that manuscripts get developed is based on our discussions, based on the information we get, based on our ability to look over these slides. We develop a draft manuscript. The draft is circulated to all of you to look at, to comment on. It's a draft.

We then will revise and deal with it. Depending upon the nature of it, we come around another time. You'll see it again. At that point, you can say, "Hey, I'm not interested in staying on this as an author. I disagree." We hope that won't happen, but it could happen.

So you're not committing yourself to be an author until you say, "Yes. In fact, I'm willing to be an author on that particular manuscript." And when people are unable to or choose not to want to be an author, we do acknowledge that they were

Page 23

present at the meeting, just in fairness for people to know who they were.

All the agenda, all the speakers, everything from this meeting will be on the website. So if anybody who wants to know was my friend Joe at that meeting I couldn't attend, they can find out about that.

So that's what's going to be going on. What do we do, as I mentioned, we publish lots of manuscripts, consensus statements, methodology reviews, commission papers on certain topics. We've conducted scientific studies or we've sponsored conducting those studies.

We've developed diagnostic classifications. We're in the process right now of -- we developed an initial classification template taxonomy, and now we have working groups on different diagnostic areas in which they're developing those.

Ursula is where, right in the front. Ursula is the co-chair of the working group that's developing the diagnostic classification for many of the conditions that we're going to be talking

Page 24

about at this particular meeting.

The idea is we have a template that describes what needs to be considered, and then when you get to the specific diagnostic, or medical areas, or problem areas, then the experts in those working groups -- and I believe we have nine of those different working groups. They're going to be encouraged to develop guidelines for the taxonomy for two to three prevalent conditions that we know that we're not going to cover every one of these.

The idea is we're hoping that other groups will say, well, you didn't have our condition, or we think that will be useful. They can ask to be involved. And to the extent that they'll follow the template that we set up, any group that has a specific diagnostic area that's not one recovered can try to develop as long as they follow the template. Those are all published, so you can see these.

Everything is transparent. I am trying to make that point to you, that everything we talk

Page 25

about is going to be available through the transcript, which is available. Slides for as much as possible will be on the website.

The manuscripts that will come out of these meetings will all be, to the extent possible, as clear as possible, what we talked about. The debates, the discussions, the difficulties, that will all be there. So that's sort of what we do.

We also are developing educational initiatives. The North American Pain School, some of you may be familiar with that, which is jointly done with groups in Canada. They invite young investigators. You have to apply to come to this. I believe they let in 30 people each year with the idea being that they get a fast-track course, a full week of being housed in a resort in Canada, in which they spend all the time doing nothing but giving lectures and discussions about pain from basic physiology, anatomy to clinical decision-making, to epidemiology, to policy.

The idea is you bring a representative. They had their first meeting last year. It was

Page 26

quite successful. I was able to attend that. ACTTION has decided to be a co-sponsor, support them with this particular initiative. We're hoping it will continue. ACTTION has committed to continuing to support this for as long as we can.

If any of you are interested, they have a fascinating website with all the details about the meeting, the presentations, who was there. They had a whole range of things, and it was really a fun meeting.

So if any of you have students who are interested in any way, shape, or form in the area of pain, they should consider looking into this and possibly applying for it.

It's for North America. There's an equivalent one in Europe. And I don't remember what the official name of that one is called, but this was modeled after that one just to make it easier for people in North America to get to. So that's the other thing that we are doing.

IMMPACT and ACTTION, we've published, as I mentioned, over 100 plus papers. It's been cited

Page 27

over 6,000 times, published in 100 different journals with diversity from addiction medicine to women's health. The women's health, we've been in those journals.

My favorite veterinary medicine, they actually have been citing some of our guideline recommendations for doing trials. I don't think, for the patient-reported outcome measures, they've been so interested. I don't know. John Farrar's not here. Actually, he was involved with one of these, and he will be here later.

So that's sort of what we are. The website, just so you see what it looks like, you'll notice along the bottom there, if you see, who is on the steering committee, meetings, what was going on, publications. We developed a measure called the SFMPQ2 with Ronald Melzack, which is an expansion and development of the McGill Pain Questionnaire, a short form that many of you may be familiar with, who has been sponsoring the meetings.

So everything is transparent, everything you want to see for IMMPACT. There will be an ACTTION

Page 28

one. In addition to all this, there will be all the other things about what goes on with ACTTION.

So that's who we are. What are the objections for this meeting? You hopefully have picked it up quite well. It's to discuss, debate, haggle important considerations, to provide suggestions regarding outcomes for clinical trials to improve the quality of chronic pain and IBS.

The idea is, can we improve the quality of studies, can we improve the consensus, some agreements about what the outcomes might be to foster systematic reviews, to foster meta-analysis, to foster younger people coming on and developing research, to provide information to regulatory agencies, which don't have to -- there's nothing binding. That's all we do is just give information. They can take that.

For those of you that might be from those agencies, hopefully this information will be useful to you. The same is true for the VA, and NIH, and the Office of Women's Health at NIH. And these are things that we hope will be useful and important to

Page 29

you. But this is what we're going to do.
To disseminate these considerations and observations, suggestions, and research, nothing will stop or die at this meeting. Even if we totally disagree, we'll write it up as if there was disagreement and not consistency. But we want to put the information out there.
So we will try to disseminate these considerations, what went on in our debates, the pros and cons, the advantages and disadvantages, suggestions for a research agenda; that is if we identify areas for which we can't make any recommendations because at this point, we're missing sufficient information, what might those studies be? And then we try to get this published in peer-reviewed journals, usually the relevant journals, depending upon the nature of the topic. We've had a lot of them in Pain, in the Journal of Pain, and Osteoarthritis and Cartilage. We've had them in some anesthesiology journals.
So we've tried to get them placed -- dental journals. We've tried to have them placed in the

Page 30

appropriate journals. So that will be discussed by this group, at least the steering committee and Ursula and Nick Verne. He's still involved with this I assume, who couldn't be here. So they will help us make some decisions. You'll have input where that might be. So that's what we're going to do.
Now, in order to do this, we need to do some herding, and you're going to not like being herded, but we have to try to gear -- we've got to get to that end. We've got a day and two-thirds, or longer, if you choose to make us do that, to try to bring this about.
So we have to do a little bit of herding, and you will feel like you're being herded. But what we've learned over there is there are some gentle arts. And this, by the way, for Shannon and Jen, who are going to be doing a lot of the yeoman's work on this, IMMPACT participants have to be herded. And what we've learned is -- I had black hair when I got started. So this is what we learned over years. Participants don't like to be

Page 31

herded. In fact, you can rarely herd IMMPACT participants, but that doesn't stop us from trying. So we'll keep working at it, and they'll keep working at it.
Participants prefer to herd themselves, but aren't very good at it. So you guys think you can do things, but sometimes you need a little guidance. Participants understand that they sometimes need to be herded, however, that doesn't make them any easier to herd. So even though you might see a value, you don't want to do this.
Harsh herding usually has negative consequences, so we've learned that. So we have to find the gentle art of doing this, and this is sort of how you do it. We find that there's a way to pull people together. This is for the rugby people who like to see those kinds of slides. And the idea is for us to work together to discuss, to debate, to look for commonalities, to look for agreements, to look for the areas that may be inconsistent, to find a way to try to improve the quality of the research, to improve and expand new

Page 32

and better treatments or improvements on the treatments that are available, that help the end user, which is the provider to the person who has that particular condition.
So that's what this is all about. I'm going to turn this meeting over. I don't know if you have any questions, quick questions, about either housekeeping or anything I've said about what we're going to do to you and how we're going to herd you. This is your chance to do it. If not, I'm going to introduce the people who are heavily involved.
Bob, you want to make a comment?
DR. DWORKIN: Yes. Bob Dworkin. As some of you may know, at noon today, there's going to be a webinar where the National Academy of Medicine panel that was asked by the FDA to prepare a report on how to deal with the opioid epidemic crisis.
We'll be having a webinar at noon today where they roll out the National Academy of Medicine response to the FDA's request. That's from noon to 1:00. So Valorie has been lovely enough to arrange that we'll have in this room,

Page 33

starting at noon, that webinar for any of you who want to see what the NAM is going to suggest to FDA.

We'll also be serving lunch upon the mezzanine, and maybe we'll be able to end a little bit before noon so people can run upstairs, grab some lunch, and then come down if you're interested in hearing what the National Academy of Medicine has to say.

DR. TURK: Thanks, Bob.

Valorie, anything else that we didn't cover?

MS. THOMPSON: Not a thing.

DR. TURK: Anything from the media people, from the audio/visual? Anything that we didn't cover? They've got their thumbs up.

Any questions before we get this or before we start?

(No response.)

DR. TURK: What we've done is, ACTTION has a director sitting over there, Dr. Dworkin, and has an associate director, me, and has assistant directors, and two of them are here today, Shannon

Page 34

Smith from the University of Rochester and Jennifer Gewandter also from the University of Rochester.

Most of you should know who they are. If you don't, you will get to know them because they can herd. They are good at doing this, and we're counting on them to do this. So I want to turn the meeting over to them, and they'll give a little bit more information and then get started.

So Shannon, you're going to be the starter. For any of the speakers, there is a pointer, a laser pointer, and just push the green button to go forward on the slides, and that's all you got to do. And the people in the audio/visuals will take care of everything else.

DR. SMITH: Thanks, Dennis. So I am Shannon Smith. As Dennis mentioned, on behalf of Jen Gewandter, Bob Dworkin, Dennis Turk, and myself, I want to again welcome you all and thank deeply the steering committee who has helped us plan this meeting, as well as Valorie, Andrea, and their team. They've been really instrumental in getting this meeting together. So thank you to all of

Page 35

them.

I just want to announce first, before we start the meeting, that for all the speakers, we're going to be giving you a little sign that says when you have 4 minutes left and then a sign that tells you when your time is up because we have a very tight-packed schedule, and we're already behind schedule. So we're going to do our best to try to stay on schedule as much as we can.

The first two moderators I'd like to introduce are Tony Lembo, who's at Harvard University, at the medical school as well as Beth Israel Deaconess Medical Center, and Ursula Wesselmann, who's a professor at the University of Alabama in Birmingham.

So the two of them will be moderating this first section, so I'm going to hand it over to them now. Thank you.

DR. WESSELMAN: Yes. I want to thank the organizers for putting this topic actually together for IMMPACT. And we were saying yesterday evening at the dinner how productive it is for all of us

Page 36

who work in the field of pelvic pain, IBS, and other visceral pain syndromes to actually have a forum to get together because we are all in very different subspecialties.

We will have four speakers this morning. We will have a coffee break after the first two speakers. And then, if you look in the program, you will see that we have very much time actually for the discussion. It's an hour and 40 minutes, so what I encourage you to do is to please make note of your questions. We will discuss mainly at the end of the lectures unless there is any burning question that needs to be addressed right away.

So we have a lot of time for discussion, so please think about the questions we would like to discuss as you hear the four speakers.

It's a pleasure to introduce the first speaker, Henry Lai, who is a urologist at Washington University in St. Louis. He will talk on interstitial cystitis, overview and assessment of pain outcomes, and implications for inclusion criteria.

Presentation – Henry Lai

DR. LAI: Good morning, all. I'm going to spend the next 20 minutes talking about interstitial cystitis and bladder pain syndrome. A lot of what I'm talking about is controversial about a case definition. What is IC/BPS, and what are the current outcome assessments that we have, and what are new ways to move forward?

So what really is interstitial cystitis? I always like to start the PowerPoint with this description because it kind of captures what the patients are really facing. So I will read it out to you

We have all met at one time or another patients who suffer chronically from their bladder, and we meet someone who has suffered chronically, periodically, and constantly have to urinate or often go at all moments of the day and night, and it hurts every time they urinate. It's very miserable. It's very distressful. It affects their physical health and their mental health.

It is a miserable condition to have. I

mean, I can only imagine what life is like if you have to go to the bathroom every 30 minutes throughout the day. You can't sleep. You can't rest. It disrupts your work schedule. It hurts all the time. It affects your sexual life, physical health, mental health, and it's not going away. It's like you have a perpetual urinary tract infection, but you don't. It is a miserable condition to have.

But what is the contemporary case definition of interstitial cystitis and IC/BPS? Now, this is the definition that has been endorsed by the AUA, which has more than 22,000 members in the United States and across the world, as well as SUFU, which stands for the Society of URODYNAMICS, Female Pelvic Medicine and Urogenital Reconstruction, which is a subspecialty, a specialized organization to look at pelvic conditions.

So this is the definition endorsed and being used to define IC/BPS. The patient needs to have pain, pressure, discomfort perceived to be related to the bladder. They have to have low urinary

tract symptoms such as frequency, urgency, nocturia, in the absence of infections and other identifiable causes such as cancer, et cetera.

Now, this case definition is actually very similar to what is being adopted by the Europeans, by the EAU and ESSIC, and this is very similar. BPS is a condition where patients would have chronic pelvic pain, pressure, discomfort related to the bladder and at least one urinary symptom. And you have to rule out other confusable diseases.

So if you look at the commonalities between the contemporary case definitions across both sides of the Atlantic, you realize it is a chronic condition. It's characterized by pain, pressure, and discomfort in the bladder or the pelvic area. It has to be associated with urinary tract symptoms such as frequency, urgency, and nocturia.

It is based on the report of pain and urinary symptoms by the patient. As you know, there's no pathognomonic pathology, imaging finding, and perhaps with the exception for a smoker or a patient with Hunner's lesion, there is

no pathognomonic cystoscopic finding.

There are no biomarkers that we use commonly in clinical practice. It is essentially, like a lot of the other pain syndromes, a clinical syndrome. It is consistent of a heterogeneous population. So I wanted to lay out the background right here.

Our contemporary definition is actually quite a departure from the NIDDK criteria for IC/BPS research. That was developed more than 30 years ago. The reason I wanted to bring this up is because this IC/BPS research definition from the NIDDK is still commonly used by regulatory agents and in clinical trials. So I want to give you a perspective of what it was.

In 1987, the NIDDK established a committee to streamline the research for IC/BPS. And in 1988, the year after, they emphasized cystoscopic finding. The context of this definition is, at the time, there was really no research definition of interstitial cystitis. So it was meant to be a starting point where the NIH can enroll patients

Page 41

and study the conditions. So they have to lay out some inclusion/exclusion criteria and kind of define what the condition is.

It was thought at the time that IC/BPS is a bladder disease rather than a pain syndrome. It is based on expert opinion and expert consensus, but the intention at the time is to have a research definition -- I want to emphasize a research definition, not a clinical definition -- to enroll relatively uniform populations so you can study them and to have some kind of objective criteria to enroll them into the clinical study. That's the context.

To have interstitial cystitis according to this definition, the patient needs to have pain. They need to have urinary urgency. But much more importantly, they have to have some kind of objective cystoscopic finding in the bladder.

When you look inside a bladder, you have to either see Hunner's lesion or you have to see glomerulation, which is a submucosal hemorrhage inside the bladder. Now, there are other criteria

Page 42

such as your dynamics criteria, how long you have to have the pain, and you have to rule out other things like infection. But the gist of this is that you need to have pain, urgency, and a certain cystoscopic finding.

Now, this is for a good intention, but unfortunately, this research definition becomes a clinical case definition where people use it to diagnose IC/BPS. So clinicians start to use this definition to define or diagnose IC/BPS. Sometimes, unfortunately, they will do a cystoscopy on the patients and find that while there's no Hunner's lesion in the bladder and there's no glomerulations in your bladder, therefore, you don't have IC/BPS, and therefore, I don't really know how to treat you.

I still see that these days, and I think this is very unfortunate that the research definition becomes the clinical criteria for clinical care.

It's also somewhat unfortunate that this becomes the de facto definition of what IC/BPS is

Page 43

from a regulatory perspective, enrollment in clinical trial, clinical trial design, and drug approval.

The real problem with this research definition is that Hunner's lesion is very uncommon. It's only seen in about 10 percent of the patients. Glomerulation, that you are required to have, is really nonspecific in this condition. And the majority of the patients that fit the contemporary definition of IC/BPS are actually not covered by this definition.

So I have to admit that patients with Hunner's lesion in the bladder is a different group of patients. They are what we call the classic interstitial cystitis patients. So they have focal visible distinct area of information in the bladder that you can see on cystoscopy. So it's almost a sunburst pattern that you can see in the focal area inside the bladder with radiating vessels to the side, and sometimes they bleed when you hydro-distend them.

You can see this in office cystoscopy. You

Page 44

can see this with hydrodistension. Sometimes, it will resemble carcinoma in situ of the bladder that tempts you to do a biopsy, but you will see chronic information on the biopsy specimen.

So Hunner's lesion, I think, is a distinct group, in a different group among the population. Patients with Hunner's lesion are typically treated differently from the rest of the IC/BPS patients. In general, you have good response.

The AUA guideline will say that you need to try to identify Hunner's lesion. If you see them, you could treat them with fulguration or injection of cannula, which is triamcinolone, into the bladder.

The data are limited because they are small series, single-center series, but a lot of them suggested you get a pretty reasonable response if you do either a cannula injection or fulguration into the bladder. It last probably about a year or two, so we need repeated treatment.

The problem is that most of the IC/BPS patients that we see these days actually don't have

Page 45

Hunner's lesion, perhaps up to 10 percent. The other 90 percent of the people have normal cystoscopy, and they are the difficult ones to treat. So what do you do with them?

Glomerulation is another definition that you need to have, another criteria you need to have in the NIDDK definition of interstitial cystitis. There has been a long overdue systematic review, but this paper basically said that glomerulation is not specific, is almost irrelevant to IC/BPS. The conclusion from the systematic review is that there is no convincing evidence in the review literature from the last few decades that glomerulation should be included in the diagnosis of phenotyping of IC/BPS.

I think they sum it really well. We should not be looking for glomerulations to define the conditions. The real problem is that the NIDDK criteria that are still in use actually miss the majority of patients that will fit the contemporary case definition of the disease.

This is a paper from the Journal of Urology

Page 46

that summarizes it really well. So it says at the bottom, "Strict application of NIDDK criteria will have misdiagnosed more than 60 percent of the patients with the conditions" because it is just too restrictive to be used by conditions for diagnosis of the condition.

So it is good as a research definition to enroll a homogeneous group of patients for research, but it will miss a lot of people that actually have the clinical condition. It may not be good for the clinical diagnosis.

So I try to summarize some of the problems with the criteria because it really doesn't address a large unmet need of patients and society. There are a lot of patients who have IC/BPS who do not fit the criteria. So if we use that criteria for regulatory reasons, for clinical trial, and drug development, it is doing almost a disservice to patients and society overall.

We are really restricting to a very narrow minority of subgroup of patients. But the reality is that it is a heterogeneous patient population,

Page 47

just like any other pain syndrome. It is a clinical syndrome, and we really desperately need novel treatment. And I will allude to that later, the fact that we don't have objective biomarkers, imaging, cystoscopic finding, et cetera. It doesn't mean that the patient needs to suffer and stay with the old way of doing things.

Now, this is a very nice paper that actually compares the people who fit the NIDDK criteria versus the ones who don't fulfill the NIDDK criteria. This study basically enrolled patients that fit the contemporary case definition of IC/BPS and do a cystoscopy on a subset of the patients, and identify ones that fit the NIDDK criteria and the ones who do not. And they look at different things in this comparison.

So if you compare the patients who fit the criteria versus the ones who don't fit the old NIDDK criteria, there's really no difference in urinary biomarkers among the ones that they have looked at, where they did the bladder biopsies of those two groups of patients. There's really no

Page 48

difference in bladder biopsy features. They look at a number of biomarkers and really didn't see much.

They look at the symptoms, the clinical presentation of those two groups of patients. It looks like, other than an increase in urinary frequency, nocturia, and decreased bladder capacity, and the NIDDK group, there really isn't a lot of difference in the other clinical presentations, either.

Now, the clinical reality is that the IC/BPS population is heterogeneous if you do cystoscopy on them. I think I alluded earlier that only about 10 percent of the patients have Hunner's lesion. Not a lot of patients have glomerulation, and it's really not specific to the conditions. In a majority of the patients, the bladder actually looks fine.

So in this study, they define the patient population into three different groups, mild symptoms, moderate symptoms, and severe symptoms, and they look at cystoscopic finding. I mean, it's

Page 49

really all over the place. It's only a minority of the patients, in this case about 10 percent, that have Hunner's lesion inside the bladder, but you've got varying cystoscopic finding in the clinical populations.

Now, there has also been histologic studies done. Patients are enrolled based on the clinical criteria, and they look in multiple ways in the bladder, biopsies, and see what other differences. They have actually done some innovative clustering algorithms to divide the patients into three groups.

The majority of the patients actually have a normal bladder. The bladder histology is completely normal. There is only a very small number of people that have a loss of urothelium with edema, and information, and glomerulation, and mast cells in the bladder. There's a middle group of about 8 percent that lost to urothelium, but without a lot of growth evidence of information in the bladder.

So the clinical reality is that the

Page 50

population is heterogeneous when you biopsy them. That leads to the recognition that it is a syndrome with a heterogeneous population. And the ESSIC, the European guideline, even suggests that we should further classify and phenotype the patients based on what we see on cystoscopy and based on what we see on the bladder biopsy because they are actually different groups of patients if you try to classify them a little further.

So the criticism on the NIDDK criteria is not only from this side of the Atlantic Ocean. Even the EAU guideline set the diagnostic criteria described by the NIDDK almost 30 years ago and was formulated for research purposes only and is inappropriate for clinical care, clinical trial, et cetera.

So IC/BPS is a very difficult syndrome, just like other visceral chronic pain syndromes. It's very difficult to treat because it is a heterogeneous population. There isn't a lot of objective biomarkers that we could put our hands, on and we have really very poor understanding of

Page 51

the etiology, what's causing the conditions.

Some would argue that there are deficient urothelium in the bladder that allows the urine to be exposed to the bladder and cause other issues in the bladder. Others will argue it's a neurologic condition involving central sensitization, peripheral sensitization, et cetera.

The current treatment is the linear algorithm. You do first-line treatment, second-line treatment, third-line treatment, fourth-line treatment. But what we really need to do is to move towards individualized treatment or IC/BPS. We need to define a phenotype and pathophysiology of the conditions, and then map the phenotype and pathophysiologies to specific treatment so that we can improve the clinical outcome. For example, if there is peripheral dysfunction in our patient, you may want to consider myofascial physical therapy.

So to move forward in IC/BPS, we need to define a clinical population. We need to recognize it's a heterogeneous population. We don't want to be very restrictive, but in fact, we need to

Page 52

recruit patients and phenotype them in the clinical trials. And Dr. Clemens tomorrow will talk about what we find on the MAPP in terms of understanding the pathophysiology of the symptoms.

We also need to have better tools to assess clinical outcome. This is traditional, ubiquitous, and it's almost standard clinical tool to assess clinical outcome. It's called the IC Symptom Index and IC Problem Index.

Essentially, it's a composite score that combines bladder pain, urinary frequency, urgency, and nocturia symptoms of patients. But there have been some recent psychometric studies that show that you should not be combining the outcome with a composite score that combines both pain and urinary symptoms.

So the MAPP developed a new score that separates out the pain symptoms from the urinary symptoms, and they should be measured differently, so there's a psychometric study.

As part of the MAPP, we follow patients over the course of a year, and then we were able to

Page 53

identify different subgroups of patients. There are some patients whose pain gets better over time, some that are stable, and some are worsened. And there are people that improve the urinary symptoms over time, some are stable, some are worsened, and the two groups are not exactly the same.

We have done some studies as part of the MAPP that Quentin will probably talk about more tomorrow that the longitudinal outcome over a year was somewhat different between the people who have improvement in their pain symptoms and improvement in urinary symptoms.

There are certain predictors of patients whose urinary symptoms get worse over the course of a year, and there are certain predictors of patients whose pain gets worse over the course of a year. They do overlap somewhat, but they are not identical.

I think it's better to measure pain and urinary symptoms separately because, as we know from both the MAPP study and clinical care of patients, some of the urinary symptoms can improve

Page 54

with certain treatments such as neuromodulation in a stem [ph]. And the pain component may or may not improve, so they track differently.

There are other potential outcome measures we could look at. For example, we've been looking at flares, which is very common the patients. The urinary pain frequency, urgency, gets worse with flares. Longer flares are associated with worse pain and urinary frequency.

We have been doing focus groups to capture aspects of flare that are important to patients. Perhaps this is something we might want to consider as a potential outcome in future clinical trials because it does impact patients' overall health.

There are also potential biomarkers that we could look at. We have to submit we currently don't have validated biomarkers for IC/BPS. It would be ideal if we could have some diagnostic markers to identify the patient population and also measure the outcome. But the fact that we don't have it doesn't mean the patient needs to suffer for the next 30 years.

Page 55

So I would say just move on, have better clinical case definitions of IC/BPS, and consider looking at some novel outcome measure for the condition. Thank you for your time.

(Applause.)

DR. LEMBO: Thank you, Henry. That was great.

We're going to hold questions, as we said, until after all the speakers. Our next speaker is Michel Pontari. He's professor and vice-chair of the Department of Urology at Temple University and the Lewis Katz School of Medicine. And his talk is going to be on prostatitis. Thank you, Mike.

Presentation – Michel Pontari

DR. PONTARI: I want to thank you for inviting me to this very interesting meeting. This is the NIDDK classification of prostatitis. This was adopted after a consensus conference in 1995 and published in 1999. Type 1 is acute bacterial prostatitis. These are people who actually have an infection, a tender prostate, come into the hospital with a fever, and dysuria, and get IV

Page 56

antibiotics.

Type 2 is chronic bacterial prostatitis, actually relatively uncommon. People who actually have a bacterial infection in the prostate, they get treated with antibiotics in between episodes. They are asymptomatic, and don't have pain.

Type 3 is the most common, about 90 to 95 percent of patients, which is what we're going to deal with today. Chronic pelvic pain syndrome has been called the "headache in the pelvis." It was arbitrarily divided into 3A and 3B with inflammation in either seminal plasma, express prosthetic secretions, or post-prosthetic massage urine, and 3B is no inflammation. So far, there have not been many clinically or any clinically significant differences between these.

Type 4 is asymptomatic inflammatory prostatitis. These are people who have no symptoms, but on biopsy or for some reason have a post-EPS, and will have inflammation but without pain.

So type 3 combines a prior diagnosis of

Page 57

chronic non-bacterial prostatitis and prosthetic prostatodynia. There was a classification from '78 from George Drach and raised this question of significance of inflammation. There is no correlation between inflammation and symptoms, and the term "chronic pelvic pain syndrome" is used instead of prostatitis because it recognizes that pain may not be from the prostate.

The term "prostatitis" is a horrible term because it implies that there is inflammation or something with the prostate. These guys have pelvic pain that may not be coming from the prostate, which is important.

What we use is the NIH definition, and the key symptom in prostatitis is pain. What separates what we call prostatitis from BPH is pain. Guys come in with frequency, urgency, "Doc, I get up at night." We call that BPH. They come in, "Doc, I get up at night and I got this pain." We call that prostatitis.

The NIH definition is genital, urinary, or pelvic pain for at least three months with or

Page 58

without voiding symptoms in the absence of uropathogenic bacteria, in the absence of other causes of pain such as malignancy.

So in terms of epidemiology, the UDA study found that there was about 1800 visits per 100,000 population. We did a study with the International Consultation of Male LUTS. Over 24 studies of prevalence was 7.1 with a median of 6.7. It's higher in Africa than North America, and the incidence is about 3.3 per 1,000 men or about 267,000 cases a year in the U.S.

So to help study this, the NIH formed the CPCRN, the Chronic Prostatitis Collaborative Research Network, in 1997, and there were six sites across America. We enrolled 488 men with chronic prostatitis over a four-year-period. The mean age is 42, so these are young guys who have this condition. And the range was 4 percent with less than 25 and 13 percent were greater than 55. I have patients with this condition between the age of 16 and 88 all over the map. It isn't just one age that gets this.

Page 59

In this, we study symptoms, bacterial studies, symptom scores, and we did three clinical trials. So for a symptom assessment, we developed the NIH CPSI, the chronic prostatitis symptom index, which is a validated, self-administered index. We compared symptoms to patients with BPH and asymptomatic controls and came up with three sections.

What happened in the development was there was a review of prior literature with an inventory of symptoms. There were several studies prior to this, de la Rosette in '93, George Brabalis back in '90, and a large study by Rich Alexander using an internet survey that catalog some of the symptoms that these patients had. And this was the basis for going to focus groups of 6 to 8 patients from 4 sites and talking about their pain symptoms, urinary symptoms, quality of life, physical functioning, a lot of the impact domains.

There was an initial draft of 55 questions covering pain, urinary symptoms, sexual symptoms, quality of life, and economic impact. There was

Page 60

cognitive testing, and then there was a revised draft with 5 centers. We sat down with 2 patients, said, "Do you understand this?" and trimmed down the 21 items.

Then with this, we also gave them the AUA Symptom Score. This is the standard assessment of symptoms in urology for lower urinary tract symptoms. Nocturia, frequency, urgency, essentially 7 different categories that we use, and 4 demographic questions.

The control groups were men with BPH and asymptomatic controls, so there was not surprisingly a difference in pain. The men with BPH had less than 10 percent pelvic pain. I think that in and of itself is actually interesting because men who come in just BPH, if we don't ask them, we're not going to get that they have pain. So 10 percent of guys we thought, oh, you just have BPH actually had pain and much lower than the other controls.

The top four pain locations became items 1A and D in the index, so the penis, testicles,

Page 61

bladder, and the perineum. The frequency of pain became item 3. The intensity was a 10-point scale, which is item 4. And we added ejaculatory pain. If there's one symptom that seems to be almost pathognomonic of men who we think have prostatitis, it's post-ejaculatory pain.

For urinary symptoms, these men had a lot more dysuria than everybody else, so that became question 2. The AUA Symptom Index was high in both the prostatitis and the BPH patients and almost equivalent. So what was done was that 2 symptoms seemed to recreate the entire AUA Symptom Score. That was obstructive, voiding symptoms, and frequency, and these together became 5 and 6. So these weren't selected for any reason other than they recreated the rest of the symptom score.

For quality of life, there were 8 questions over 2 domains, psychological distress and physical limitations. And these all seemed to perform equally well, so we picked two of them, put them on the score, and then the overall quality-of-life item became number 9.

Page 62

So this is the score. It's 9 questions, 43 points. There is the pain subscale, which is 1 to 4. There is urinary, 5 and 6, and the quality of life is at the end. It's asymmetric. How this is scored, you can have a maximum of one on some, you have 10 on the others.

There was a study that Quentin did on rescoring this on a 0 to 100 scale, and it didn't seem to make any difference. So the max score is 43 and anything above 15, we would consider significant symptoms.

So how responsive is this? Kaye Propert looked at patients enrolled in our first clinical study, and this was responsiveness to change over time in 174 men in our first CPCRN study, looking at a total and 3 subscores versus the global response assessment. This GRA had seven items, 3 on the other side, and of no change.

So patients who improved in total pain and quality of life were highly responsive, and then as you went from slight improvement to marked improvement, it became more responsive. Urine

Page 63

symptoms were only responsive in those who had marked improvement, and there was a small response in any scale for those who became worse.

So if you get worse, the NIH CPSI is not going to reflect that very well. If you get better, the pain and quality of life is going to be a lot more responsive than the urinary symptoms.

What they found here, too, is that in between sections of the GRA, in between categories was 4 points. So 4 points seems to be the smallest perceptible change, but the ROC curves indicated that 6 points was a better choice for who has a clinically significant response.

A slightly different study was done by Turner out in Washington, looking at primary and secondary care, not tertiary care sites. They compared the NIH CPSI to a grade A chronic pain scale. Pain and quality of life were markedly associated with this scale and urinary symptoms had a low correlation.

So again, the pain and quality of life were responsive to change; urinary scale is not, similar

Page 64

to what we had seen in Propert's study. And the recommendation was you can use the NIH CPSI, but add another validated pain measure to it.

So how did this respond in our trials? We did three trials. The first one was 6 weeks of either ciprofloxacin, 500 twice a day or Tamsulosin. And this was a 2-by-2 block, so you either got placebo and placebo, cipro and placebo, Tamsulosin plus placebo, or both drugs.

Essentially, it was a negative study, so the NIH CPSI, there's no difference between cipro/no cipro, tamsulosin/no tamsulosin for the total or any of the subscores.

We did a second study, a 12-week study looking at Alfuzosin, which is an alpha blocker, in men who were alpha-blocker naïve and symptoms less than 2 years. We thought, wow, this is the group it's going to work in. And what's interesting is that -- and it got published in the New England Journal I think because we got the exact same response for both groups, which is pretty hard to do, but 49.3 response.

Page 65

(Laughter.)
DR. PONTARI: Incredible.
So the endpoint was a 4-point reduction in the NIH CPSI, which is the smallest perceptible change. Absolutely no change, absolutely no difference to assess, so no difference in total, no difference in the subscores.
We came close in the third one. We did a trial of pregabalin. It was a dose escalation from 150 to 600. We had found from people in this room that 450 was good for fibromyalgia, so we went up to 600 over 2 weeks at each dose.
The primary outcome was a 6-point drop. We were pretty confident we were going to have a significant improvement. For the primary outcome, we got 0.7, very close, but not quite. But the secondary outcomes all significantly improved.
Now, one caveat is the secondary outcomes are only in people who completed the trial. For the GRA, it was people who -- if you are a non-responder, if you dropped out, you were a non-responder. So we have to take it with a grain

Page 66

of salt. These were only in people who completed the trial. But it was significant for total score, all the subscores, so that any other outcomes in McGill was significant. So if you improve, it seems to go with that.
Quentin did an adaptation, where they added -- this is for the male; there's a female also -- 2C and 2D, which is pain or discomfort with bladder filling or pain or discomfort relieved by voiding. And with this scale, called the GUPI, which is what we use now, and the MAPP, the RSC curves had 7 points for this versus 6 defines a responder and 4 point again is the minimum perceptible change. So this is what we use in the MAPP.
So Henry's study was interesting from the standpoint of the men. This is in the MAPP. They asked patients, "Do you have pain with bladder filling and/or relieved by bladder emptying, and do you have urgency, or do you have none of these?"
What we talk about in the MAPP are people who have bladder pain, painful urgency, or none,

Page 67

and what's interesting is, as you can see the, 75 percent of the men had bladder symptoms. So these are men who would be, by prior criteria, characterized as interstitial cystitis, and this was a really interesting phenomenon for us.
Going back, I looked at all the old studies. I'm not sure if we didn't ask. It didn't come up. We sat with men in focus groups for the NIH CPSI, and this didn't come up. This was not one of the symptoms that we found. But it turns out that if you ask these questions specifically, 75 percent of the men have this.
Going from both to either to neither -- sorry, from neither to either to both, you see an increase in severe pain, frequency, urgency, symptom burden, depression, worsened quality of life, including IBS. So if you go more towards men with bladder symptoms, they have an increased incidence of IBS, up to about 30 percent.
This was the study that Henry had mentioned, Jamie Griffith's study. They did look at factor analysis. I think the first line I'd like to put

Page 68

in here, "Questionnaires differ in their assumptions about how symptoms cluster together." We always thought pain in urinary symptoms have to be from the same thing. Well, maybe they're not.
So looking at this, two factors came out, pain and urinary symptoms. Pain also correlated with depression, whereas urinary does not. So their conclusion was the total score is they combine pain and urinary symptoms into one score, limited for clinical and research purposes.
In terms of the impact domain of emotion, just as in other things such as the vulvodynia paper that we got, catastrophizing the men with prostatitis is important. It's associated with greater disability, depression, urinary symptoms, and greater pain. And in this study by Dean Tripp, helplessness was the strongest predictor of pain, even after controlling for depression and urinary symptoms.
Now, in terms of entrance criteria, there has been nothing like the uproar over IC for prostatitis. So there was a consensus conference,

Page 69

a very small one, here in Washington in 1998. It was international because we had people from overseas. And at that point, there was a consensus that was developed for adoption following the criteria for clinical studies in prostatitis. It used the NIH definition, the 1995 NIH classification scheme, the eligibility criteria that we came up with for the CPCRN, and the NIH CPSI.

So the first three I think would still stand. I think the CPSI we have to look at as an outcome measure because I don't think it would work as a total. This is what we came up with for the CPCRN. Used the NIH criteria, do you have pain for greater than 3 months, and the inclusion criteria would be sort of common-sense things.

It's a diagnosis of exclusion. If you have pain from prostate cancer, we don't want you here. If you had BCG, we don't want you. We even excluded unilateral orchialgia. So patients who only had pain in the testes were not included in these trials; structured neurologic disease, and

Page 70

prior prostate surgery.

So as far as the MAPP entrance criteria, there was a CPCRN and now we're in to the MAPP era, pretty similar. We had a little bit of difference in terms of the length of disease in terms of IC and the prostatitis, so we kept the old prostatitis one, 18 years old, again, non-zero score, so they're really pretty similar.

As far as inclusion criteria, the only thing I could see with the MAPP that's different than CPCRN was a history of non-dermatologic cancer. There were some deferral criteria that I think are useful in terms of if you're going to do a trial. These are the deferral criteria we used for the CPCRN.

So if you had an infection at that point, we didn't exclude you. We said, "Come back in 3 months and see if it's gone." So you couldn't have an active infection within 3 months, a recent STD. If you'd undergone a prostate biopsy in 3 months, you can come back and see if your symptoms have persisted, acute or chronic

Page 71

epididymitis from the last 3 months, and/or genital herpes in the last 12 months.

So with prostatitis, the main symptom we're talking about is pain. It's what distinguishes it from BPH. And 75 percent of men with prostatitis or chronic pelvic pain syndrome also have bladder pain.

So the implications are first for treatments, because you can use bladder medications for these, but also, these men may have to be included in whatever interstitial cystitis is. So what the bladder symptoms are, having these guys in one silo may not be completely appropriate.

Pain and urinary symptoms may not respond together. Using a combined score is probably not a good idea. The NIH CPSI and GUPI total scores are likely not useful in clinical trials, and so far, we have had minimal controversy in the entrance criteria for chronic pelvic pain syndrome. Thank you.

(Applause.)

DR. LEMBO: Thank you, Michel.

Page 72

So we're going to take a break now. It's supposed to be a 20-minute break. Maybe we could make it a little bit shorter so we'll catch up a little bit on time. We have an hour and 40 minutes for discussion, so we should be able to catch up then. So why don't we reconvene in about 15 minutes? Thank you.

(Whereupon, at 9:28 a.m., a recess was taken.)

DR. WESSELMAN: We want to continue with two more topics, vulvodynia and IBS. And as I said before, we will have a discussion right after that. And we want to keep on time so that we can finish just before 12:00 so that we can look at the webinar that is scheduled for 12:00. And lunch apparently is right here on this level, so it will be easy to eat and watch the webinar.

It's my pleasure to introduce Andrea Rapkin, who is a professor of OB/GYN at UCLA, and the topic of her lecture is vulvodynia, overview and assessment of pain outcomes and implications for inclusion criteria.

Presentation – Andrea Rapkin

DR. RAPKIN: Thank you, and thank for the opportunity to speak this morning. How is the volume level? Okay.

The most recent nomenclature for vulvodynia is now called the 2015 classification. This does not differ substantially from previous classifications, but I will point out where it does. There hasn't been quite as much dissension about this nomenclature as there has been with prostatitis and bladder pain syndrome.

This particular series of definitions was developed through a consensus with most of the groups that are involved with the research of vulvodynia, the International Society for the Study of Vulvovaginal Disorders, the IPPS, International Pelvic Pain Society, and the International Society for the Study of Women's Sexual Health.

This is a pain-based classification system, and as with other disorders, there are two main classes of vulvar pain. We're not interested in specifically vulvar pain caused by a specific

disorder. But vulvar pain is an idiopathic pain disorder of at least 3 months duration. So the previous classification system included pain of 6 months duration. This is moved down to 3 months duration.

Again, being a pain-based system, we are defined only on the basis of location triggers, temporal pattern, and onset. The specific visual or sensual characteristics of the syndrome are not included here, so we have location, localized. The most common localized area is to the vestibule, and I'll show you a picture of that for those who are not familiar with the vestibule in a moment. The other area of localized vulvar pain is the clitoris, but this is much less common.

The pain can be generalized to the entire vulvar region. This seems to be again less common. Probably only about a 10th to a 20th of the patients may have some generalized pain as well, and this would be considered a mixed picture.

Those with only generalized pain in my experience tend to be much older and tend to be

confused with patients with either pudendal neuralgia or with a referred hyperalgesia from, say, a bladder pain syndrome or other pelvic pain disorders.

Now, the trigger is important, but the trigger may be provoked, or spontaneous, or both. And in this situation, we're primarily talking now about the vestibulodynia. The reason for using on vestibulodynia is, one, it's the most prevalent type of vulvodynia and, two, it is the most well-studied type of vulvodynia. So going forth, I'm going to be focusing on provoked vestibulodynia or PVD.

How is the pain provoked? Most typically with sexual contact, vulvovaginal penetration, but also with tampon use, and in many women with sitting or with tight clothing, there is pain in the genital area.

We also have the mixed pattern, whereby there's pain that is provoked and spontaneous. These patients are part of a spectrum and tend to have more severe pain. The spontaneous pain alone

without provoked symptoms is very rare. Temporal pattern, it's unlike other pain conditions. The pain wouldn't be constant if it's provoked, however, it's generally intermittent.

The onset, however, is interesting, and we wish we had more information about this aspect. So when we talk about someone with PVD1 or primary vestibulodynia, we're talking about an individual who says, as long as they remember having genital contact, they've had pain.

Well, that usually doesn't go back into childhood. We don't have any real good prospective studies. So the "as long as I can remember" often goes back to, well, the first time I tried tampons, or the first time I was trying to touch the area, or I was with a sexual partner.

Acquired, however, is an individual who has had a period of comfortable or pleasurable genital contact followed by onset of pain. And often, it's a very acute onset of pain. Women may say that it suddenly started with a particular episode of intercourse that was uncomfortable, or that they

Page 77

thought they had a bladder infection or a yeast infection, and in fact may go in and be treated multiple times for these infections, only to find that it's not actually an infection.

Now, the new classification system includes potential factors associated with vulvodynia. I think these are very important because this may in the future be part of the phenotyping or subgrouping of patients, but right now there isn't data to suggest that this would be the case. So we have individuals who commonly would have musculoskeletal factors.

Now, the sensitivity of the vestibule and the sensitivity of the pelvic-floor muscles are correlated, but not well correlated, so there are different factors that may predict each. But it has been demonstrated that many women with PVD do have muscle overactivity, do have increase in muscle tone, decreased relaxation, and alterations even on ultrasound of the bulk of the muscles.

Neurologic mechanisms are being evaluated more recently with imaging studies, but there is

Page 78

some changes in peripheral neuroproliferation, plus or minus mast cells, et cetera; and of course, psychosocial as with any chronic pain condition. But what isn't evaluated I think as well in other chronic pelvic pain conditions, and it is our area to really focus on, is sexual functioning because that's the primary area that's impacted with PVD.

So likely not one disorder, it may be a constellation of disorders, but I think this is the case with many of the conditions we're talking about. Pathophysiology is unknown. There are many different studies, many different theories.

There's information on neuroproliferation, and this comes from not only biopsies, but the fact that many women with provoked vestibulodynia are treated by vestibulectomy. So yes, in the beginning of treatment for IC, there were some individuals who had cystectomy no longer done, but vestibulectomies are still performed.

There are still many studies evaluated the histopathology. It cannot be diagnostic in a specific patient, but there have been some changes

Page 79

with free nerve endings, with staining for nerve growth factors, substance P with mast cells, plus or minus to granulation.

We're looking at CNS processing, as are many other individuals, but as of yet, you cannot get an MRI and predict what diagnosis and what treatment outcome, functional or otherwise, the muscular, myofascial problems.

Now, the issue with hormones, we're going to exclude individuals who were clearly estrogen deficient; so if an individual is post-menopausal with genital urinary syndrome of menopause or if an individual has lactational amenorrhea. But someone who's been on long-term low-dose hormonal contraceptives that are known to lower estrogen levels, the question is what's going on in these patients, do they have a different picture, should they be a subcategory?

Of course, comorbidities; about 50 percent of patients with PVD have comorbidities, and the more comorbidities, the worse the symptoms. Genetic polymorphisms have been identified, but are

Page 80

as yet not useful for treatment.

Lots of psychological factors, and in particular, as I said, sexuality has been looked at quite a bit. Unfortunately, either we don't think as an initiating factor, but certainly as a result, lower desire, arousal, satisfaction, orgasm, more negative attitudes on the part of patient and partner. So partner responses are often very important in these studies, and I'll get to that soon.

Here is the vestibule. So the vestibule is that area of the vulva outside the hymeneal ring and will include the posterior hymeneal remnants as well. And the vestibule stops where the vulvar tissue begins to look like epithelialized skin as opposed to mucus membrane.

The vestibule itself in the past has been shown to have areas of erythema, and in fact, in Friedrich's criteria, erythema was one of the characteristics, but it's now not considered to be necessary. This accounted for initially why the condition was called vulvar vestibulitis.

Page 81

As with prostatitis, which we've just heard, it is no longer considered to be an itis. And the areas of erythema are there, but they may be vasodilatation from neurogenic inflammation, for example.

Some of the prevalence studies are done with questionnaire, with phone survey, and it is unclear how well they correlate with actual findings on examination, but there is some correlation to be sure. About 8 percent of women have this provoked vestibulodynia.

So the recommendations for outcome measures for clinical trials is one of the manuscripts that was developed by Ursula and her colleagues, other individuals who are not here today, Caroline Pukall, Sophie Bergeron, Candace Brown, Gloria Bachmann.

You have a copy of this, and much of what I'm going to say going forth reflects or includes some of these aspects because I think it was really very well done, looking at recommendations for outcome measures. And obviously, the purpose is

Page 82

similar to the purpose for this entire meeting, to provide consistent measures, facilitate comparisons across studies, and improve outcome measures for multiple dimensions, in particular the very important measures of interpersonal relations, sexuality, and et cetera.

So in this particular manuscript and in the studies developing from that, the IMMPACT framework was used, but the sexuality measures were added. And again, there was a focus on PVD.

The recommendations were that the psychometric properties be foremost, that there be evidence that these particular measures that are used are valid and reliable wherever possible, that the issues relate to practical application.

So we've seen a couple of survey questionnaires that are quite short, but once you're evaluating different dimensions, including sexual functioning, partner response, affect, et cetera, you can have quite a large bulk of questionnaires that are given to patients for assessment.

Page 83

It would be important that these measures had been used in prior clinical trials of PVD so we can see how they perform. And with vulvodynia, it would be nice to have some specific measures, just like with the GUPI or GUPTI [ph] -- I don't remember what you're calling it now -- there were questions that were added for female having to do with pain at the opening of the vagina and the vulva area. There are very specific measures for vulvodynia, and I'll get to some of those in a moment.

So some of the core measures to consider again are the fact of how do we define inclusion criteria. This hasn't been quite as difficult for PVD because we do have a location; we have pain localized to the vulvar vestibule or mixed. The pain should be provoked or mixed.

In terms of onset, we are still evaluating PVD1 or PVD2, but it is not clear how important these items will be. Note that, as part of the 2015 criteria, exam is no longer part of that, but all of the individuals currently who are doing any

Page 84

research in vulvodynia outcomes are including examination findings. Also, the level of pain required to fulfill diagnostic criteria hasn't specifically been established, but generally is accepted as about 3 out of 10 or greater on a VRS scale.

So if we're looking at pain with gentle contact of the vulvar vestibule, this was initially taken from the Friedrich's criteria. So as I said, it may not be part of the 2015 nomenclature, but it is part of what is important researchers feel for entrance criteria, for inclusion criteria.

The pain level with the cotton swab hasn't really been standardized. So do you brush the cotton along the vestibule? Another approach that's been more or less standardized is to place the cotton swab at a perpendicular angle and to depress to a third of the head of the cotton swab.

It's not included in trials, but in my experience, it does matter to some degree how you place the cotton swab, and again, the threshold for inclusion.

Page 85

But what about this vulvar sensitivity testing? How useful is it really? Unfortunately, the patient sensitivity on exam and their reports of clinical improvement of pain with intercourse don't really correlate that well.

Now, in my patients, we do the sensitivity testing with every visit, and I have found that there is some correlation, but we tend to find the sensitivity with a Q-tip lagging behind or being more problematic, even when they're starting to have improvement with pain with intercourse. And we know that sexuality is very complicated, obviously.

Again, self-report and objective pain ratings may not be correlated with sexual function parameters or satisfaction, again reflecting the complexities here, and the fact, as I'll get to, of the pelvic-floor involvement.

Cotton swab tests can have some false positives, so individuals with no complaints of pain with intercourse may have discomfort when their vestibule is prodded with a dry cotton swab,

Page 86

moistened as well.

So what about this vulvar sensitivity? It is increased when you have individuals who have a younger age of the first onset of vulvar pain, provoked vulvar pain. And it is more correlated when you have the provoked pain, obviously, and also correlated with pain after intercourse.

It is not associated with comorbidities, so the more comorbidities doesn't necessarily increase vulvar sensitivity with cotton swab. There aren't really significant changes or differences with PVD1 or PVD2, meaning primary or secondary, or if an individual has spontaneous pain.

Luckily, for those who have had trouble with their algometers when they use them in a research setting, the cotton swab test does correlate well with algometer findings.

Now, just what would the exclusion criteria be? Well, infectious, inflammatory, neoplastic. Neurologic is a little confusing, again, with how you make a diagnosis of pudendal neuralgia, for example. But we would exclude obvious neurologic

Page 87

trauma, iatrogenic, and the hormonal deficiencies. And I've already mentioned some of the controversies in that area of the hormonal contraceptives.

So these are the frequently studied parameters, the pain intensity and vulvar vestibular sensitivity, 1 versus 2, comorbidities, anxiety, and depression. Clearly, further phenotyping is important.

My goodness, I'm talking. Let's skip over this.

The tampon test may reflect more than intercourse pain, the situation of pain in the vestibule. And this is because many women are avoiding sexual contact, it's so painful. So the tampon test was developed by David Foster and has been validated. And this is placement and withdrawal of a tampon ad then determining what the pain is.

It's reliable, tested 3 weeks in a row, good validity, and better adherence than asking a woman to have intercourse when they have pain. They may

Page 88

not even have a partner at this time.

Pain intensity has been looked at by both the NRS and the VRS. The problem with pain intensity scales is that we can't use pain in the last 24 hours. We really need to think about "pain the last time you attempted vulvovaginal penetration" or pain during non-sexual contact activities. You can also switch from pain in the preceding month, where they may not have even tried contact, to pain during last 4 penetration attempts.

I'm going to mention a few times, if I have any time left, the VPAQ, which is a recently validated vulvar pain assessment questionnaire that has two pain intensity domains and has been able to cover many of these domains.

Also important would be aspects related to the sensory descriptors of pain. Certainly the Short-Form McGill and then the VPAQ, which I have a copy of if anyone is interested, has a pain descriptor's subscale. So it would be recommended to include the VPAQ with the Short-Form McGill.

Page 89

Temporal pattern, as I mentioned, is less important because pain tends to be provoked. I'm going to skip over the data on PVD1 and 2. There are some inconsistencies here.

What about physical functioning? Yes, the SF-36 is helpful. It has not been used in many PVD trials because the pain is provoked, as I said, and can be avoided by avoiding genital contact. However, because of the comorbidities, it's important to look at health-related quality of life and also to look at specific-to-PVD physical functioning. So the VPAQ has a life interference subscale, and there's a quality-of-life scale from the VQOLS that could be added.

Another core outcome is looking at sexuality, Female Sexual Function Inventory. The problem with this inventory is that it asks for your different questions in the last 4 weeks. For women who are no longer sexually active, this is a problem. There is in the VPAQ now a self-penetration interference subscale for self-contact, and this can be useful. So consider as a secondary

Page 90

measure the PROMIS scale minus the ejaculation questions.

Sexual satisfaction, distress, and interference should be included as secondary outcome measure, and you can find these in VPAQ, and there are other measures that can be used.

In terms of emotional functioning, vulvodynia is associated with depression, anxiety, however, they're not clinically significant levels. They're different than controls, but obviously should be measured along with catastrophizing ways of coping as well.

So in the literature of vulvodynia, the BD-II has been used quite a bit for this anxiety because trait anxiety doesn't change at all with treatment. The state of the STAI could be used.

PVD emotional response questionnaires that are specific would be important, and again, you can look at the VPAQ or the PASS-20. The PASS-20 is interesting because it's important that fear of intercourse pain be assessed, and that can be assessed with anxiety related to sexual activity.

Page 91

Of course, for any treatment trial, you wanted to look at global improvement and treatment satisfaction. And these should be adapted to women with vulvodynia, and PGIC has been adapted to women with vulvodynia.

Let's skip over to -- supplemental measures, domains that are relevant would be the social role functioning, in addition, relationship adjustment, and documenting comorbidities. Of the impacts, core and supplemental domains, the ones that are particularly relevant, 4 vulvodynia studies include interpersonal functioning, coping, and social role functioning.

These are more important than the other supplemental domains. Of course, the core domains of pain, physical, and emotional functioning, improvement symptoms and disposition are very important.

So in sum, there is no single validated PVD questionnaire for all measures. I've tried to go into measures that should be considered core and supplementary. The vulvar pain assessment

Page 92

questionnaire captures most of these core and supplemental domains, and some measures from PROMIS should be added for secondary. And that's about it. Thank you.

(Applause.)

DR. LEMBO: Thank you, Andrea.

We're going to move on and talk about irritable bowel syndrome. We're really fortunate to have Bill Chey, who's a professor of medicine from the University of Chicago, who's going to talk to us about it. Thank you.

Presentation – William Chey

DR. CHEY: Thanks so much for the invitation to the organizers of the meeting and to Tony for inclusion in this meeting. I'm going to talk to you about IBS. I've really focused on issues around measuring pain, as I thought that's what I was supposed to do. So if I left stuff out that was supposed to be included, I apologize.

Let's start with a general overview of IBS. I think everybody in the room has some familiarity with this, but there are certainly some nuances

Page 93

that bear mention that have tremendous overlap with all the other things, all the other topics that have been discussed already this morning.

So this, like many of the other conditions already discussed, is a symptom-based diagnosis. There is no biomarker that we can utilize to identify patients with IBS at the current time. It's a very prevalent problem. It affects anywhere from between 7 to 15 percent of the general population in the United States.

It has a remarkable impact on quality of life. Particularly, severely affected patients have dramatic impairments in quality of life as well as work productivity, and it really should come as no surprise that this is a very expensive disorder, billions of dollars on an annual basis in direct and indirect costs.

We currently utilize the Rome IV criteria to diagnose IBS, particularly for clinical research studies. Now, there are some differences, and you can see the Rome IV criteria on the slide. There are some importance differences between Rome IV

Page 94

versus Rome III. Rome IV was just released in May of 2016. Rome III was the standard for many years before that.

If you look at this slide, it says recurrent abdominal pain, one day per week, associated with two or more of the following, related to defecation, onset associated with change in stool frequency, or onset associated with change in stool form.

Just so you know this about the differences between Rome III and Rome IV, remember that Rome III included abdominal pain and discomfort. Rome IV is focusing really solely on abdominal pain.

Now, let me just say that I'm going to China in two months to discuss the unhappiness in the Far East with that decision. So this has created some controversy, although there is lots of qualitative research to suggest that patients draw a clear distinction between pain and discomfort, and that in some parts of the world -- not Asia, but other parts of the world -- they actually had difficulty

Page 95

describing or identifying with the word "discomfort."

The other thing is that we raised the threshold for the diagnosis, so one day per week. It used to be 3 days per month. That will obviously have some effect on the overall prevalence of the condition.

Then the last changes to mention is related to defecation. Recall that Rome III said "relieved by defecation." And the reason we made that change is because, again, in epidemiological research, it became clear that there's a small subset of patients with IBS who have exacerbation of their pain with defecation.

So the majority get relief of their pain with defecation, but a smaller proportion get actual worsening of their pain with defecation.

Remember that IBS can be or is diverse from a clinical phenotype standpoint. There are patients with constipation, patients with diarrhea, and patients with a mixture of both constipation and diarrhea.

Page 96

Also remember that, from the standpoint of the Rome criteria, we distinguish between these different IBS subgroups on the basis of stool consistency, not stool frequency, although from a regulatory standpoint, as we'll talk about in a moment, for the constipation subgroup, we focus on stool frequency as opposed to stool consistency.

Now, there are multiple symptoms that are reported by patients with IBS, not just the ones that are included in the definition created by Rome. So if you look at this particular graphic from the UCLA group, you'll actually see that patients commonly endorse complaints around gas and bloating, for example, in addition to problems with pain and altered defecation.

You might ask -- and this question comes up a lot -- well why isn't, for example, gas and bloating part of the Rome definition for dividing patients with IBS? And the reason for that is while somewhere in the neighborhood of 80 to 85 percent of patients with IBS endorse those complaints, it turns out that those complaints are

Page 97

also extremely common in virtually every other functional GI diagnosis as well as the healthy population.

So while it's a common complaint in patients with IBS, it unfortunately offers little discriminative value from healthy volunteers as well as patients with other functional GI disorders.

In terms of what symptoms are the most bothersome, actually, I think a theme that comes out over and over again as you talk to patients with IBS is the unpredictability of their symptoms. It's the inability to be able to know what symptoms they're going to experience and when they're going to experience them. And that creates a lot of secondary situational anxiety that I think amplifies their symptoms as well as drives their illness experience.

But you can see that, outside of that, the next most bothersome symptom is abdominal pain followed by distension and urgency, actually, which is interesting.

Page 98

Now, there are a number of challenges. These challenges are really the same as has been described with many of the other disorders that have already been discussed this morning, issues around overlap.

One thing that's important to mention in patients with IBS is that somewhere in the neighborhood of 30 to 50 percent of IBS, patients have at least 1 other functional GI diagnosis. So in addition to IBS, they'll have functional dyspepsia. They'll have functional heartburn. They'll have proctalgia, so a variety of other GI symptoms.

But in addition to that, as has already been mentioned, a variety of other non-GI-related conditions, which I think gives us insight into the pathophysiology of at least a subset of these individuals. I firmly believe that patients who have multiple overlapping pain disorders have more of a top-down disease as opposed to a bottom-up driven disease. And that, at some point down the road if that's validated, may well have treatment

Page 99

implications in regards to the selection of, for example, pain modulators.

The pathophysiology of IBS is diverse. So there is an interaction between a variety of host factors, luminal factors, and environmental factors. I think that the building blocks that we've talked about for many years, important to the pathogenesis of IBS, abnormalities and motility, visceral sensation, brain-gut interactions, are still operative. But I think as time goes on, we're increasingly becoming aware that these factors are really influenced by issues like, for example, permeability, immune activation, genetics.

Biosalts, interestingly, as sort of going back to the future, have been knocking around for a long, long time. But it's increasingly clear, for example, that we've been missing patients with bioacid malabsorption who present as otherwise being diagnosed with IBSD.

Psychological, psychosocial factors as the last speaker mentioned, very important. And food, I think again is increasingly becoming recognized

Page 100

as one of the main drivers for symptoms in patients with IBS. In fact, I think that the two main environmental stimuli for IBS symptoms are psychosocial stressors as well as food.

In fact, realize that somewhere between two-thirds and three-quarters of IBS patients associate a symptom onset or worsening with eating a meal. And I'm emphasizing this because, to me, it's so interesting that from a research standpoint and also from a therapy standpoint, we have focused a lot on many of these other factors, and our therapies are really largely predicated upon pharmaceuticals, but we have very little evidence in the way of how diet therapies may benefit patients with IBS.

So we're going to focus the rest of the talk on the issue at hand, which is measurement of abdominal pain and a couple comments about abdominal pain in patients with IBS.

So we have already talked about the fact that this is a symptom-based diagnosis. It's heterogeneous, both phenotypically as well as

Page 101

pathophysiologically. And along that same line, pain is also multi-dimensional. And again, as has already been discussed by the speakers that preceded me, there are a variety of different issues that play into pain, and there are a variety of different issues that we can measure in regards to pain.

We'll talk about where we are at the current time, but I think that for this group, regardless of what discipline we're talking about, right now, we're really focused almost exclusively on pain intensity. But I think we'd all agree that for any of us that actually take care of patients, there are a variety of different issues around pain that involve factors other than intensity that are equally important to the patient.

So understanding the impact of different pain dimensions is important certainly to guide PRO development for future clinical trials, as we're discussing today, and define inclusion criteria, which I assume will also be discussed later in this meeting.

Page 102

Now, a couple other things to consider about pain in IBS is that, right now, we lump all patients with abdominal pain and altered bowel habits together as suffering with IBS. Now, I've already emphasized to you that these patients are remarkably heterogeneous from a clinical phenotype standpoint: diarrhea, constipation, and a mixture of both.

Now, we lump all those patients under the rubric of IBS, but realize that the characteristics of pain of not just the bowel habits, but also abdominal pain are different amongst these different subgroups. This is something that we've only very recently started to learn.

This is work from Brennan Spiegel's group at UCLA -- he was kind enough to share a number of these slides with me -- that shows that there are differences, for example, in the mean frequency of pain attacks amongst patients with different subgroups of IBS. So patients with IBSD have more frequent pain attacks than patients with IBSC or IBSM.

Page 103

The proportion of pain attacks, which interfere with work or daily activities is also different. And this is interesting because the IBSD patients have more frequent pain attacks, but the impactfulness of the abdominal pain is greater amongst IBSC patients. In fact, I'll show you data from our own group that we just presented at DDW this past year that shows the exact same thing.

Now, behavior is during pain attacks, so taking medications roughly -- or between the groups. But look at "goes to bed." Patients with IBSC actually behave very differently than patients with other subgroups of IBS in that regard.

And patients will tell you this in clinic. The way that they frequently will respond when they're in a painful flare is they go to bed and try to go to sleep. And defecation on the other hand is a much more common endpoint for patients with IBSD and IBSM, IBSD in particular.

Now, this is the data that I told you about that we just recently presented at our national meeting last year. We haven't published this in

Page 104

full manuscript form yet. But it's an interesting study that was conducted in 71,000 U.S. citizens using a digital app-based platform that we've created at UCLA Cedars and University of Michigan called My GI Health.

This is an app-based platform that utilizes the PROMIS questionnaires to be able to determine the frequency and severity of all GI symptoms. So all 8 of the most commonly reported GI symptoms for which a patient might see a gastroenterologist are assessed as part of this platform using computer-adaptive technology so that the patient only answers questions about the symptoms that they're experiencing.

Anyway, utilizing this, we were able to identify a large number of patients with IBS, and we were also able to stratify between the different subgroups of IBS on the basis of a whole variety of different types of symptoms, including abdominal pain.

It's interesting that, of patients with IBSC, they had significantly greater PROMIS scores

Page 105

for abdominal pain. So their overall PROMIS scale scores for abdominal pain were really quite dramatically higher than the other subgroups or than IBSD in particular. So it's interesting that, again, the frequency in that one study was a bit more in IBSD, but the overall burden seems to be greater for IBSC. Similarly, you can see small differences for abdominal pain severity, which trended towards but didn't reach statistical significance.

But bothersomeness was significantly greater in the IBSC group as well -- in our study, we actually had a small increase in frequency, which you can see is really a very small difference.

At the current time, the standard to measure pain in patients with IBS is an 11-point numeric rating scale. And this has gone through a lot of iterations over the last several decades. I've been involved in virtually every drug development program as has Tony. Tony and I have been developed literally in virtually every drug development program for IBS that's occurred over

Page 106

the last 20 years. And I can tell you that, the way this has evolved, the pain has always been assessed utilizing either Likert scales or NRS.

Initially, the very early trials with alosetron, for example, used a 5-point Likert scale. But over time, we've recognized that the sensitivity of a 5-point Likert scale is probably not great enough to really distinguish between the relatively small differences that can occur in the treatments offered for patients, a heterogeneous population like IBS.

So the current standard is an 11-point numeric rating scale, and that's been evaluated now by several different groups and validated, and we'll talk about that. So the currently utilized numeric rating scale assesses pain over the last 24 hours. The PRO guidance is, you need a weekly average of worse pain to be greater than 3 to qualify for clinical trials. And that is the current standard that's utilized or that is recommended by FDA.

This NRS does work. It's been validated

Page 107

against a variety of different clinical anchors. And it turns out that in a number of different studies now, an MCID of around 2 or a difference of around 30 percent in reduction of pain score is clinically meaningful. And that's actually been validated against, again, whole variety of different factors, including IBS Symptom Severity Scale, the Functional Bowel Disease Severity Index, IBS QOL, EQ5D, which is an assessment of quality of life as well as presenteeism in a variety of individual IBS symptoms.

So lessons learned about pain in IBS, pain and discomfort are different. This is actually work again from the UCLA group, but has also been validated by other groups. And that is that patients really just make a distinction between abdominal pain versus abdominal discomfort. Somewhere around 80 to 85 patients really draw that distinction, which tells you that a smaller proportion actually considers them to be more on a continuum and to be the same, really be part of the same spectrum of the same disorder.

Page 108

IBS pain is multi-faceted. Some pain dimensions drive illness experience more than others. Probably the one that drives it the most is severity, but frequency and bothersomeness are also extremely important in terms of what drives the patient to go see the physician and what also causes them to experience disability, both in regards to their home life and work.

Patients with more intense, frequent, constant, and unpredictable pain have higher illness, impairments, and again that word, "unpredictability," I think is a really important thing that's really difficult to measure. But again, if you talk to patients, patients will tell you that one of their biggest concerns is not knowing when they're going to have a problem.

The multidimensionality of pain should be borne in mind as we develop conceptual frameworks for PROs. So I think, unlike some of the other conditions that have been discussed so far, the FDA has actually released regulatory guidance for trials being conducted in patients with IBS.

Page 109

A couple of themes that come out of the pre-reading documents that were sent to all of you, what I think are important to emphasize is, for many years, the standard in IBS trials was to actually use a global assessment primary endpoint, so Subjects Global Assessment, SGA, adequate relief, or satisfactory relief, so a single question item that assessed global IBS symptoms.

We're actually very comfortable with that in the IBS investigative community for many years. FDA has problems with a global endpoint for a variety of different reasons, I think many of them quite valid. So they've recommended that a single general item asking patients to rate overall change in IBS symptoms as a primary endpoint to support an efficacy claim is not recommended. So that's pretty much out for us in IBS trials.

What they've recommended instead is a primary endpoint that encompasses the main symptoms of IBS, so really consistent with the Rome criteria, abdominal pain and abnormalities in bowel habits.

Page 110

For drugs that are developed to treat a single IBS system -- and we're starting to see more of these, for example, drugs specifically targeting pain, for example. Specific symptoms or signs should be the primary endpoint, but the other symptoms of IBS should be assessed, not necessarily to determine if there's efficacy, if none is expected, but to make sure that you don't exacerbate any of the other key symptoms of IBS.

So what is the actual regulatory guidance right now for the different subgroups of IBS, for IBSC? And the pain response definition is going to be durable across all of the different IBS subgroups, so for pain severity, weekly average of worse pain in the past 24 hours, a score of greater than or equal to 3 on a 11-point numeric rating scale is what's currently used to identify patients who are eligible for an IBSC trial.

As I mentioned earlier, for IBSC, stool frequency is currently the bowel-related symptom that's focused on, so fewer than 3 complete spontaneous bowel movements per week is also a

Page 111

recommended entry criteria for IBSC trials on the basis of FDA guidance.

Now, in terms of responder definitions, a decrease in weekly average of worse abdominal pain in the past 24 hours, of greater than or equal to 30 percent, and that's based upon that validation data that I alluded to earlier in this discussion; and also, an increase of at least one complete spontaneous bowel movement per week from baseline in terms of stool frequency.

Actually it's interesting. I think the FDA very rightfully identified a number of concerns about global endpoints, and they recommended interim guidance in terms that I've just laid out for you. The interesting thing was there was very little validation data of the recommended interim endpoints.

What's I think very gratifying is that there's been publication recently of some work. It's post hoc work, and we have to accept all the limitations of that, that actually validates the endpoints selected by FDA.

Page 112

This is actually data from a post hoc analysis that was published on the heels of the linaclotide phase 3 clinical trial program. Remember, linaclotide is a guanylate cyclase C agonist that's FDA approved for the treatment of patients with IBSC as well as chronic idiopathic constipation.

What they identified using clinical anchors, so patient-reported complaint questionnaires, which assess symptom-specific patient rating of change as well as degree of relief of IBS symptoms, is that a threshold of 30 percent was very consistent with their post hoc analysis from those clinical anchor data.

In addition, they also validated that the increase in complete spontaneous bowel movements of 1 per week was also consistent with the degree of change identified by patients who felt better following drug therapy.

So for IBSC, we actually have at least some post hoc data that validates the thresholds that have been recommended by FDA for responders in

Page 113

patients who are entered into large randomized control trials.

This is also an interesting analysis that was part of this paper, and this looked at varying the number of weeks needed to meet the FDA responder endpoint. Remember that -- I didn't mention this, but I should have -- right now that the regulatory standard is at least 6 of 12 weeks of response to be defined as a responder.

It turns out that the overall accuracy offered by changing the threshold turns out to be greatest right at that 6-week cutoff point. You can see that as you start to demand a greater number of weeks, of course your specificity goes up, but your sensitivity drops quite precipitously. Also, varying the percentage of improvement, that is the weekly average in abdominal pain score, you can also see how that affects the results as well.

For IBSD, it's the same regulatory recommendation in regards to abdominal pain, that is a 30 percent -- or first in terms of just enrollment criteria, a weekly average of worse pain

Page 114

in the past 24 hours of at least 3 on an 11-point NRS.

Here, though, for stool, for the bowel-related complaints, it's not stool frequency. And this is an interesting story, which is that when the FDA initially came out with their interim guidance on this particular topic, they actually wanted a stool frequency endpoint for diarrhea.

Actually, the functional GI community pushed back really hard on the FDA, including producing evidence to show that in patients with diarrhea-related complaints, stool frequency is oftentimes not a good surrogate assessment for complaints of diarrhea, clinical complaints of diarrhea. So I think it was gratifying for everybody involved that they were willing to revise the criteria and base it more on stool consistency.

You can see that patients to be included in trial should have at least 2 days of Bristol Stool Form scale score of 6 or 7, which is loose or watery stool.

To be a responder, the same definition in

Page 115

terms of improvement in overall pain score saw a reduction from worse abdominal pain in the last 24 hours at baseline of at least 30 percent. And then here, we see a responder definition that's really based on improvements in stool consistency, so a 50 percent reduction in a number of days with the bowel movements that is type 6 or 7 using the Bristol Stool Form Scale.

So to summarize, IBS is a symptom-based disorder without a reliable biomarker at the current time. It's a multi-symptom disorder, and it's heterogeneous both from the phenotypic standpoint as well as the pathophysiological standpoint.

Symptoms are largely measured using patient-reported outcomes because that's all we really have to measure at the current time. Pain measurement in IBS focuses right now on severity, but I hope that I've shown you and open your minds to the thought that we may want to think about other aspects of pain other than only severity or intensity.

Page 116

A 30 percent reduction in abdominal pain severity has been determined to be clinically meaningful and is the current regulatory standard as recommended by FDA.

The last thing that I have on the summary slide that I think is also important to think about, and it ties into the comments that I made earlier, is that we may need to think differently about how to measure pain and what aspects are important to pain based upon IBS subgroup.

So in other words, frequency, intensity, bothersomeness, unpredictability are all traits that may differ between patients with IBSD and IBSC, for example. So thank you very much.

(Applause.)

Q&A and Panel Discussion

DR. LEMBO: At this time, we're going to take questions and have a discussion. We'd like to invite the panelists to come back up to the table. And this is supposed to be a discussion, so if anybody has questions, why don't we go ahead and raise your hand. We'll take them individually, and

Page 117

please state your name, of course.
DR. KATZ: Daniel Katz from Boston. So my question is the following. It seems like we've chopped up the body into all these different parts, and if your functional disorder happens to be in your colon, then you get managed by one group of specialists, and if it's in your bladder, you're going to get managed by a different group of specialists. And if it's in your vagina, you get managed by a different group of specialists.
The outcome measures that we've heard about all relate to, well, should it be 2 bowel movements or 3, or 2 urinations or 3, or all of these very kind of organ-specific numbers.
But when I listen to each one of the speakers, it seems to me that there's a common message, which is that there's an underlying proclivity that certain people have towards being sensitive to pain or other stimuli. And if that happens to show up in your bladder, well, then whatever passes through your bladder is going to cause or evoke symptoms; if it happens to be air

Page 118

passing through, if it happens to be a tampon passing through it. But the commonality is that whatever gets in touch with that organ will provoke whatever -- evoke -- the stimuli are.
So it seems to me that a relevant question is, what is that underlying proclivity towards being sensitive to pain, and how do we identify when that is present? How do we diagnose that? And could that be a treatment target? Certainly it is for tricyclic antidepressants and other things, and does that have any implications for how we measure outcome of clinical trials in these disorders?
Are we losing something by just focusing on the superficial manifestation of whatever the end organ is that's bearing the brunt of it, especially since most of these patients, as I've heard, have symptoms referable to multiple different end organs.
DR. WESSELMAN: We should probably go through each speaker because we have four different pain syndromes, which are all -- yes. This is a

Page 119

really key point, and I was thinking about it, exactly about what you mentioned when some of you presented data on each individual pain syndrome, because the drugs that actually had some efficacy aren't the general pain drugs that we also use for neuropathic pain.
So the question is really, as we are trying to identify certain endpoints for a given manifestation of visceral pain, are we going the right way, because this research has actually been going on both in basic science research as well as in clinical translational research for the last 20 years, but we have not really identified any pharmacological targets that have proven to be very valuable.
So we might actually want to step back and think about these overlapping visceral pain conditions to see if we address them from kind of above rather than trying to pinpoint a certain symptom that we want to have as a primary endpoint.
DR. CHEY: I think it is an excellent point. I think there is an investigative way to think

Page 120

about this and a clinical way to think about this. The clinical way to think about this is moving towards multi-disciplinary care models because there is a lot of commonality between these different conditions.
I think that, increasingly, there is evidence to suggest that a team-based integrative approach that embraces not just the GI symptoms, but for example, we have a -- Quentin knows about this -- called the Michigan Bowel Control program, where we have collaborative effort between the urogynecologist, colorectal surgeons, physical therapists, gastroenterologists, behavioral therapist, dietitians. We all see patients together, and I guarantee you our outcomes are significantly better than the patients that are seen individually by just GI, or urogynecology, or colorectal surgery.
So from a clinical standpoint, I think that we're moving towards more of those holistic care models.
From an investigative standpoint, I still

Page 121

think, though, that the peripheral symptoms are important triggers, even for patients who have a central sensitization, a central abnormality in pain processing or perception. So we're moving those peripheral symptoms, which are oftentimes important triggers for that sensitivity. It still makes a lot of patients better.

I think that one of the reasons why the therapeutic gain in the clinical trials is so marginal is related to exactly the points you raise, which is that some patients are more top-down, some patients are more bottom-up. And then there's a whole bunch of patients where it's a combination of both. I think that, mechanistically, we cap out at a relatively low rate with a lot of the drugs that are targeting one specific mechanism.

DR. LAI: From the perspective of IC, I think you realize that it is a heterogeneous group of populations. There are those that are more of a top-down picture with systemic manifestation and with chronic overlapping pain syndrome of IBS,

Page 122

fibromyalgia, vulvodynia, et cetera. But there is also a subgroup within the IC/BPS which is very bladder-centric.

We've shown pictures of people having Hunner's lesion inside the bladder, and they do get better just injecting cannula or fulguration locally in the bladder. So if it's a stomach effect, you wouldn't expect these patients to get better with very localized treatment.

If they are bladder-centric patients with Hunner's lesion in the bladder with pelvic-floor discomfort, they get better with pelvic-floor physical therapy. But they are these top-down people, too. So there's a top-down, bottom-up, and there is overlap because there is interaction between the peripheral end organ to the brain and central sensitization.

As you alluded to, I think IBS literature is clear on that, is that sometimes you tune down the peripheral trigger, and your central sensitization or systemic manifestation do get better.

So I think one of the challenges here is to

Page 123

define the phenotypes, who is top-down, who is bottom-up. And they both need to be included in the clinical trials, but you need to phenotype them, identify them almost a priori, rather than doing it afterwards, because any single drug that targets a single receptor or mechanism, if you just mix in the entire group, and analyze it as an entire group without some a priori power calculation about the different phenotypes, it is very likely going to wash out the effects in clinical trials.

I think this is one of the reasons why a lot of IC clinical trials fail because we think it's a single entity, but it is really not.

DR. RAPKIN: I think by looking at provoked vestibulodynia, we've already narrowed the focus quite a bit. Our patients with generalized vulvodynia in fact behave a little bit more like some of the chronic visceral pain disorders.

With provoked vestibulodynia, unlike emptying the bowel or emptying the bladder, you can avoid contact with the vestibule. And in addition,

Page 124

the contact is generally of a sexual nature, which as we know is very involved.

So this is a little more specific to the PVD population. Otherwise, I agree with everything everyone has said. And again, our most important treatment outcomes do include multi-disciplinary approaches, particularly addressing cognitive, behavioral, and pelvic-floor physical therapy.

DR. PONTARI: So there is data with chronic pelvic pain syndrome that there is up-regulation or there's evidence of central sensitization, both efferent and afferent from University of Washington. I agree. It's hard to avoid some of the triggers. You're not going to not have sex or urinate, so it's harder to avoid those.

One thing we're trying to do with the MAPP and education with the urologist is first to realize that people have other syndromes. I mean, people walk in with just prostate, but it's like, you have to ask about the bowel, rheumatologic stuff. So part of it is training people with all those other -- and then making appropriate

Page 125

referrals, and trying to get the rest of the person treated as well.

DR. LEMBO: I think I echo what everybody says. Clinically, we do look for patients, try to identify patients with the other chronic pain syndromes. As gastroenterologists, we focus on the bowel, but we do treat those patients differently, because we do feel that it does take a multi-disciplinary approach.

We've found that, through some research work, those patients with -- actually, we call it, extraintestinal because we're GI focused -- extraintestinal manifestations tend to have much more anxiety and depression. And oftentimes, at least in some modeling, it seems like the anxiety and depression is driving a lot of the extraintestinal symptoms. So we try to treat those as aggressively as we can. And I think everybody it sounds like is doing the same thing, looking at the top-down or bottom-up approach.

So we tend to use a lot more of the centrally active drugs, but I think it's a great

Page 126

point.

DR. KATZ: So I guess, for Bob and for Dennis, who are managing this meeting, I guess the question is, if we end up in this meeting without proposing some kind of measure for this central sensitization that is common across all these syndromes as a way of classifying the patients when they come into the study and as a way of measuring their outcome, have we really done our job here if we just focus on what's happening in the end organ?

DR. PONTARI: How do you measure central sensitization?

DR. BRUEHL: This is Steve Bruehl. I'm just going to ask a related question to that. So what I hear from I think every single panelist is that every one of the conditions discussed is heterogeneous. There appear to be subtypes. It seems like everybody agrees there are certain subtypes that are top-down and implying some type of central sensitization.

For us, the only effective way to measure central sensitization in the laboratory is temporal

Page 127

summation protocols. And I'm not sure clinically whether that is practical to do that, but I know, under the right conditions, they're reasonably reliable; intraclass correlation about 0.7 or so, which is not spectacular, but it's adequate.

I'm wondering whether, as a clinical assessment, if anyone has ever attempted to do that routinely and looked at what effect that might have. Now, I just would mention as an aside, we've been doing this, been working with an urologic surgeon who's interested in overactive bladder syndrome, which by definition is not pain. We're actually seeing central sensitization elevated in that group compared to controls as well.

I just wanted to throw that out there as an assessment methodology. I'm just curious to see what people's experiences are with that.

DR. PONTARI: I think Pat Fitzgerald did that in IC. Quentin, you can correct me if not. I'm sure that that happens in some patients with interstitial cystitis, that you see that summation effect. That's the only study that I know. You

Page 128

may know some others.

DR. CLEMENS: I think the NIH right now has convened a group to try to examine this I guess from the clinical standpoint, and the concept is to develop a standard review of systems. So one way to do it is something like there's a CMSI, which I think was Complex Multi-System Inventory that was developed at University of Michigan.

I think that's being used as a template where each of the various groups is providing input to that and adjusting it, but the concept would be in a clinical trial or potentially from a clinical standpoint you administer that and can capture in an IBS patient, for instance, a standardized assessment of bladder symptoms or vulvodynia symptoms, et cetera, and use that from a clinical standpoint, and also use it, let's say, in a clinical trial to maybe look for a signal; hey, this drug worked well for my condition and it looks like their IBS symptoms got better.

So that's one potential way as long as it's short enough to potentially be used for clinical

Page 129

purposes to be a surrogate for central sensitization based on the presence of widespread symptoms.

We're doing that in the MAPP. The other thing that we're doing is a body map, which may be even a simpler, more clinically useful way where patients can just put a check mark on where it hurts. And we're examining constructs related to number of sites versus pain severity of those sites and seeing -- in the MAPP, we're doing this -- whether pain severity is an important thing to measure or just number of sites.

But those are a couple different ways where it might be measured clinically and adaptable to clinical care and also of course for research.

DR. PONTARI: And Steve Hart's doing the pain sensitivity test, too --

DR. CLEMENS: The sensory testing.

DR. PONTARI: -- the thumb crusher, where you're trying to see where you are on the pain sensitivity scale, too.

DR. LAI: I think it would be useful if you

Page 130

could use some kind of QST, quantitative sensory testing. It could be something complicated or something as simple as a pin prick to look for a temporal summation as a surrogate measure, central sensitization.

I think it would be useful to investigate, to see if what you see in QST might predict a certain response to certain centralized systemic treatment. I just don't think there's any data for IC or CP at this point to guide treatment. I don't think it's even looked at as a possibility.

The other thing, like Quentin mentioned, there's a body map. People who check a lot of pain sites throughout the body, that could be a surrogate measure of increased sensitivity and perhaps would correlate with what you see on quantitative sensory testing.

So that could be even a potentially more practical approach to identify these people with a top-down syndrome.

DR. WESSELMAN: Next is Bob.

DR. DWORKIN: So if I were to do a

Page 131

pregabalin trial, I might predict that temporal summation would improve versus placebo, and I would predict that pain would improve versus placebo. But would you guys predict that -- and then this is the top-down question -- that pregabalin would improve the urinary or defecation symptoms?

DR. CHEY: Not necessarily, and therein lies the problem. I will say I think FDA was wise to provide some guidance for drugs aimed at single items, because there are clearly situations -- pain is probably the one that's most obvious -- where there will be drugs that largely target pain, and you just want to make sure that they don't make the other symptoms worse.

But no. I think that, mechanistically, there are lots of examples you could come up with where you might only affect diarrhea, constipation, or abdominal pain.

DR. LAI: Doesn't the CP trial show that it improves urinary symptoms?

DR. PONTARI: Right. So again, the secondary outcomes are only people complaining, but

Page 132

it did show that. And I would say, in some people, I would expect it. And the question is, is pain and urinary coming from the same thing? That's the question. If the urinary symptoms are coming from pain, which I think they are in some people, I would expect it to improve urinary symptoms in some people.

DR. DWORKIN: Urinary pain or frequency? Frequency?

DR. PONTARI: Right. When we say urinary, we mean frequency and urgent -- well, then there's painful urgency. So pain, period, urinary symptoms as a separate thing, I would expect, yes, in some people.

DR. LAI: At least with interstitial cystitis, a lot of urinary habits that you're seeing is driven by pain. The reason they go to the bathroom every 30 minutes is because when the bladder fills up and the visceral organ gets distended by the urine, they start feeling pain.

So they're going to the bathroom every 30 minutes, every hour, to relieve the bladder, to

Page 133

decrease pain. So if you can improve the pain, this could be a secondary thing that you could see in terms of improving the frequency or urgency.

That also drives to the point perhaps you need to look at them separately. Using the composite score is fine, but if you track them separately, you may see a stronger signal of what is actually happening here.

DR. WESSELMAN: Next is Andrew.

DR. RICE: Thank you. Coming from a neuropathic pain background, I get a slight hypertensive crisis every time I hear the word "central sensitization." It seems to me it's something that came from observations made by Clifford Woolf and others about a very short-lived phenomenon that occurs in experimental rodents for few tens of seconds and both in the musculoskeletal area and I'm hearing in this area. It almost crept in as a kind of diagnosis, and then people start to infer mechanisms, and therefore probably drug targets from it.

Certainly, in the neuropathic pain

Page 134

community, there's been a lot of kickback against that. One of the most aggressive arguers about it has been Pierre Hanson [ph]. And I kind of go along with his approach, but I don't necessarily take it to the extreme he does.

To define it on the basis of altered sensitivity is not adequate because that can easily be a primary sensitization. And any form of sensory profiling -- and maybe Ralph has a different view -- can't really define central sensitization and certainly can't start to describe neurologic mechanisms to that pain.

So if one thing that comes out of this for your community is some agreement on what you mean by central sensitization, and how you should measure it, and what is reasonable to imply from it, and what is not reasonable, as stated, to imply from it to me would be a valuable thing. But it's a term we hear bandied around a lot, particularly in the therapies community, without sometimes people thinking exactly what they mean by it.

I don't know if I'm being unreasonable or

Page 135

not, but that's my view.

DR. WESSELMAN: Has anybody a comment directly to that? Yes? Go ahead.

DR. BARON: Perhaps I should comment on this because you addressed me personally.

(Laughter.)

DR. BARON: I'm a strong believer that there are some, but very few, measures in QST, in quantitative sensory testing, which clearly are indicative of a problem we call central sensitization, whatever this is. And there are some items which we assess and measure in our testing protocol, which are clearly only present in patients with an insensitivity. And I think if you reduce your assessment tools to these few parts of the protocol, I think you have an idea about central sensitization.

DR. RICE: Are you confident that those tools differentiate between peripheral sensitization?

DR. BARON: Yes.

DR. RICE: You're talking about --

Page 136

DR. BARON: If I'm talking about pinprick hyperalgesia and dynamic allodynia, for example, you can measure these abnormalities in a remote area, even in a patient with visceral pain. For example, in head zones, you can find signs of, yes, desensitization. So it really is indicative of a central process, not a peripheral process.

DR. RICE: Would it be useful, therefore, to define appropriate sensory tests for this group?

DR. BARON: What shall I say? Yes.

(Laughter.)

DR. RICE: They don't have flat skin that we have the luxury of in most of our conditions. They have difficult-to-access areas of the body.

DR. BARON: Well, but I'm talking about remote areas like head zones, where you have remote pain -- any groups. I think even Cathy is doing some QST in these areas.

DR. RICE: I still think it'd be useful to define what central sensitization is in this context.

DR. BARON: Right.

Page 137

DR. RICE: It'd be a big progress.
DR. WESSELMAN: Sharon was first.
DR. HERTZ: I was just interested in hearing a little bit about some of the genetic phenotyping in the vulvodynia. This work that's being done, is there any cross-talk for the different syndromes in terms of that type of work and any common findings?
DR. RAPKIN: I think that just the beginning, most of the large ongoing studies now have been collecting material for genetic study. But the current polymorphisms are usually hypothesis driven. So for example, in some of David Foster's work, what he's looked at, are alterations in the way inflammation related to candida infection has been looked at.
In these studies having to do with oral contraceptives, there are alterations in genes relating to the androgen receptor. So I think unfortunately right now, they're hypothesis driven. But looking overall at gene-wide association studies, not a lot yet.
DR. PONTARI: So in MAPP, we have a genetics

Page 138

group, and we started out and identified a bunch of candidate, SNPs, whatever it was. And then Dan Clauw went to the folks who are collaborating with the University of Michigan and said, "Well, if you have less than 100,000, we're not going to do it," because we really don't have the numbers needed to do that, and that's where we are.
A lot of little studies have identified a lot of interesting things, but we don't have enough patients to do a meaningful study at this point.
DR. WESSELMAN: Roger?
DR. WIEDERHORN: I have a question for Dr. Pontari and for Dr. Lai, and that is that Dr. Lai identified classical IC pain as bladder-centric, which to me means visceral. Painful bladder and for various gradations of chronic prostatitis, is there any somatic component? Is there any way to distinguish visceral versus somatic in either of these two conditions?
DR. LAI: I think that's really one of the things that the MAPP study wanted to look at,

Page 139

because there's a realization that it could be a visceral disorder in some people and the realization that there could be systemic manifestation in some patients. I think we need to distinguish those two groups. I'm not sure you would want to call them two different disease conditions.
I mean, if you're going to say, well, all the risks are going to be called interstitial cystitis and study that way, and all the systems and potential overlapping conditions, it's all going to be called BPS --
DR. WIEDERHORN: Right. No, I agree.
DR. LAI: -- I'm not sure is the right path.
DR. WIEDERHORN: Yes. But is there a component of each in some of these patients, or is one all -- like for IC with a Hunner's ulcer, maybe it's more likely to be visceral, less likely to be somatic.
DR. LAI: I think at least I'm aware of two papers where they compare patients with Hunner's lesion versus one without Hunner's lesion, and they

Page 140

looked at different things. I think there are some slight differences in the urinary frequency, how often they go to the bathroom, and nocturia, how often they go to the bathroom at night.
There are some demographics differences in terms of men versus women. Men are more likely going to have Hunner's, for example. And the ones with Hunner's lesion tend to be older, probably at least a decade older. We're actually doing the same study, hopefully have a third paper in this area.
We also look at systemic syndrome because that really is the question. Do people with Hunner's lesion have systemic manifestation? Do they have somatic symptoms? And I think the answer is yes.
We see, at least in the population that we're studying, and we haven't published, is the people with Hunner's lesion do have irritable bowel syndrome, but less likely, statistically less likely.
So you could make an argument they are

Page 141

statistically or maybe clinically less likely to have systemic manifestation among the ones with the visceral syndrome, but again, it's not going to be a clear-cut line. You don't. It's not zero percent.

DR. WIEDERHORN: But I was thinking mainly of a pelvic-floor disorder as opposed to a prostate disorder or bladder disorder manifesting pain, and is there any way of differentiating that or no?

DR. PONTARI: I didn't used to believe in pelvic-floor dysfunction, but now I do because people get better with therapy. And if somebody can help us with diagnosis, that would be great. It's hard. I mean, you can go in, you can palpate, and things like that.

Clearly, there are people who respond very, very well to pelvic-floor physical therapy. People will come in who we think -- pain with ejaculation, you think, all right, you're contracting the muscles, and this is probably pelvic floor. And pelvic-floor spasm can give you pain at the tip of the penis characteristically, so the answer is yes.

Page 142

I think what's hard is an objective measure of that, which someone knows more -- I'm sure happens in other syndromes, too, I would imagine with IBS and probably vulvodynia.

So do vulvodynia patients get better with pelvic-floor PT?

DR. RAPKIN: Yes, certainly quite a large number do. It's one of the most effective modalities. But the actual measurement of the algesia -- we have a vulvalgesiometer that I know Frank Tu is also using.

Ours has not functioned in the last 6 months, and the individual who created it can't seem to fix it. We have the issue that you have to have some sort of a pressure sensor covered by a glove and go and transduce through a computer system. It's not like using a von Frey hair or something like that.

DR. LEMBO: It's interesting that in the GI world, we've used pelvic-floor therapy for constipation. And there are lots of studies showing that if you do sphincter retraining, with

Page 143

or without manometric or balloon expulsion, documentation, patients tend to do pretty well with it. And it's increasingly recognized in IBS.

Bill, I don't know if you knew of any data with pain associated with it. We know that the defecatory part tends to get better, but I'm not sure about pain.

DR. CHEY: Yes. We published a paper on this, actually, and there are actually two papers now that show the exact same thing. And that is that patients with constipation-related symptoms who undergo anorectal manometry or defecography and have evidence of outlet obstruction constipation, who then undergo physical therapy and biofeedback training, for many years, as Tony alluded to, we've accepted that those patients' constipation symptoms get better. But it's really interesting that the abdominal pain and bloating get better in a subset of the IBS patients.

So if you look at, for example, scores using the PAC-SYM, which is a validated instrument in assessing constipation that looks at bowel

Page 144

symptoms, rectal symptoms, and abdominal symptoms, PAC-SYM scores for all three domains actually get better with physical therapy and biofeedback training.

I think of it this way. I look at it this way, which is that, again, I think we have a tendency to want to oversimplify things, so we think of it as just top-down or just bottom-up, and it's not. Most of the patients have both, and the peripheral problem is an important trigger for the visceral hypersensitivity, if you will.

DR. LAI: Just to respond to Rog, there are some IC patients who have pelvic-floor tenderness, and I think in fact probably 70, 80 percent of the people do. There are some IC patients who do not have pelvic-floor tenderness, and there are patients with pelvic-floor tenderness without any bladder symptoms.

So I think as part of the clinical assessment and moving forward for clinical trial, we need to assess the pelvic floor in patients with interstitial cystitis. In fact, if you took the

Page 145

interstitial cystitis clinical trial and look at it globally, the one trial that will show positive results in a randomized controlled trial is to take the IC patients, identify those that have pelvic-floor tenderness, and subject them to physical therapy of the pelvic floor.

That particular group of patients have positive results in a randomized controlled trial. So that shows you the power of phenotyping, and narrowing your subgroup of patients, and targeting to the potential cause because if you just mix in everybody else, I think that would be a negative trial.

DR. BUTTERFIELD: Noam Butterfield of Vancouver, Canada. This is also a question for Dr. Pontari and Dr. Lai. We hear a lot about the heterogeneity and CPPS and IC/BPS, but also the overlap between the diseases.

So if either of you are a PI, and you have two different studies, and one's an IC/BPS study and one's a CPPS study in males, which study would you put your males into?

Page 146

DR. PONTARI: That's why we came up with the term UCPPS, in all seriousness, because the NIH had the ICDB, interstitial cystitis database. There's a CPCRN. We had one guy at Temple who was in both. It's all symptom based.

DR. LAI: I think the majority of the males will qualify for both criteria. If they're somehow arbitrary, there are two different entry criteria.

DR. PONTARI: Right.

DR. BUTTERFIELD: What are the differentiating criteria that -- if you were designing your own protocol, would you say we can't have two separate protocols because it's the same disease? Because clearly, they currently have different nomenclature and they still have slight differences. But what would you define as those key characteristics that would differentiate between the two?

DR. PONTARI: The entrance criteria for MAPP is men and women. I'm trying to think of the MAPP criteria we showed.

DR. CLEMENS: I think there are some men

Page 147

who -- I think the answer I might give is the presence of significant bladder symptoms if you're trying to differentiate between a male with IC versus classic CPPS.

There are some men who have chronic perineal pain. They have no urinary symptoms at all. They just hurt down there in the perineum. They would probably not meet criteria for IC. They wouldn't. And the traditional thought has been that most of the men with chronic pain have minimal urinary symptoms. And what we're finding in the MAPP study is that the rate of these urinary symptoms is more than we thought.

What we're not getting at necessarily with the analysis is that I think a fair number of these men actually have perineal pain and ejaculatory pain as what's driving the care seeking, and their urinary symptoms of, oh, by the way, but not necessarily something that would cause them to see the doc.

So you could look more closely at that, and really I think it's the severity and bothersomeness

Page 148

of the urinary frequency and potential bladder pain that would discriminate there.

DR. LAI: A distinguishing feature would be like Quentin said, do they have urinary frequency and urgency out of the ordinary? Does the pain that they're describing get worse when the bladder fills up, and does it get at least temporarily relieved when the bladder empties?

So those are potentially the distinguishing criteria. I don't think the pain criteria really sets you apart.

DR. PONTARI: No. The question is whether you want the frequency or not. There are IC trials now where you get men and women. It's the same thing. And you're asking for pelvic pain. You're not asking for actual prostatitis, inflammation, BPH stuff. That's a different thing. It's just whether you want to look at urinary symptoms or not.

I don't even know if we know the clinical significance right now of the guys with the bladder pain versus not because we just found it like a

Page 149

couple years ago. We weren't looking for it for 20 years. So we're not even sure -- actually, there were women, too.

Twelve percent of women in the MAPP, in your study, didn't have bladder symptoms. These are people we might think have urethral syndrome, whatever it is. We think these are pelvic floor. So 25 percent of men and 12 percent of women have no bladder symptoms on the MAPP, but they have pelvic pain.

DR. CLEMENS: I have a question for Bill. I was interested that new Rome criteria get rid of discomfort. So for IC patients, no matter how much we try, sometimes we can't get them to say they have pain. It's pressure or discomfort. There is something causing them to urinate every half an hour.

So I guess from the standpoint of -- I don't know; we talked about hurting at the beginning. We've moved in the opposite direction in the IC world, where we very much encourage the inclusion of discomfort as part of their criteria. This goes

Page 150

along with the concept that we have identified pain and urinary symptoms as kind of separate constructs that track differently.

So what tends to be concerning sometimes is when there's a drug study or any intervention that requires a pain level of a certain amount for IC, you're excluding quite a number of patients who are extremely bothered by what's called non-painful urinary frequency. And I think what we're moving to is the concept of dual outcomes and trying to stratify based on those types of symptoms so that discomfort is allowed, pain not necessarily as long as they have severe.

So I guess to comment about the question, what about a person, do they exist, who defecate all the time but don't have pain, and they don't have any other types of symptoms? I mean, I would think that would be IBS, and yet, by the Rome criteria, they don't meet those criteria. So do they exist or how do you handle them?

DR. CHEY: This is a totally fair question. And I think the answer is just realize that Rome

Page 151

really has created criteria for disorders mouth to anus. So in other words, there's a diagnosis, functional diarrhea or functional constipation, where patients can have only constipation or diarrhea and not have pain criteria that would satisfy the diagnostic criteria for IBS.

I actually agree with you a hundred percent. In all of our trials, we always measure discomfort as well, but we measure them separately, so we measure pain and discomfort separately. On the basis of the criteria, we do have a minimum threshold for abdominal pain. It's not just to fulfill the definition for IBS. It's a practical issue around being able to measure change.

Obviously, if a person has a mean worse pain score of 1 or 2, it's going to be really hard to measure a statistically significant change with an intervention. So what we've figured out over time is that that minimum standard, if you really want to do an assessment for pain, is probably that level of that threshold of 3.

DR. WESSELMAN: Does anybody have a specific

Page 152

comment to this question still, or are these new items to be -- yes?

DR. EDWARDS: Yes, if you don't mind if I follow up on that. Rob Edwards from Brigham and Women's. It strikes me there's an interesting potential disconnect between the pain we think these patients experience and how we measure it. So it sounds like we're talking about largely intermittent and provoked pain for these folks.

The Rome criteria for IBS indicate that I think you only need to have pain related to defecation one day a week or more; correct? With vulvodynia, certainly there are ways that people can avoid provoking pain in the vestibular and vulvar regions. With prostatitis, presumably people can reduce ejaculatory events if they've got major post-ejaculatory pain.

So a lot of our pain assessments do things like what you see up there on the screen and look at the worst abdominal pain in the past 24 hours. So that seems like it would work for some conditions like fibromyalgia, which is one of the

Page 153

things that I study. But for people with IBS or vulvodynia, who are only having intermittent pain episodes and maybe only a couple of times a week, which is all it takes to meet the criteria for these conditions, is this a problem that we're measuring pain in this way; in other words, that we're measuring intermittent pain with questions about average pain in the last 24 hours, when potentially, patients haven't had any of those pain-provoking events?

DR. CHEY: Yes, it's definitely an inherent problem, the conditions like IBS. And probably interstitial cystitis, I imagine, is the exact same way. I don't know of a better way to do it. That's the problem. I take all of your comments. I think you're spot-on in all of your comments. But unfortunately, I'm not sure there is another way that I'm aware of to really do it. People have toyed around with things like, for example, pain on only the days that you have pain, for example.

Like for example, the bowel-related criteria is interesting. One thing I didn't mention to you

Page 154

is it's greater than 25 percent with diarrhea or constipation of IBSC or IBSD, but Rome IV went to only for when you're having abnormal bowel habits. It used to be that you had to have 25 percent overall.

So in a way, we kind of addressed that a little bit when we most recently revised the criteria. But I totally take your point. I'm not sure that I know of a better way to do it, though.

DR. DWORKIN: If I'm understanding correctly, a patient who had 2 days in the past week where their worst abdominal pain was 7, they're not going to meet your criteria because their weekly average is 2. And yet they've had 2 really bad days.

DR. CHEY: That's right.

DR. DWORKIN: So the answer perhaps is what you said, that you only take the average of the days when they had pain, and you don't include in the denominator the days where they were pain-free.

DR. CHEY: You can definitely do that. You can definitely do the post hoc analysis. For

Page 155

example, we routinely will do that. We'll look at that. And the good news is, the drugs that work for that overall pain measure, the way we currently do it, are also responding. You know what I'm saying?

But to your point, there are definitely some patients that don't make it into the trial because they don't meet the threshold. And I think it's largely a pragmatic issue about just being able to, again, measure a change. But I take your point. And the thing is, you can definitely do a post hoc analysis to look at that very point.

DR. DWORKIN: You could if you entered that patient in the trial. The patient who had two 7s and 5 0s, you could show change because your treatment could in fact take those 2 days of 7 pain and, after 8 weeks of treatment, he can now have -- he or she -- 3s.

DR. CHEY: Good point. It's an interesting proposition. I don't think that we have really discussed that specifically, but I think it's a point that's worth discussing.

Page 156

DR. WESSELMAN: Is there a comment to this question?

DR. LEMBO: Ursula, can I comment? I need to take this a little further. So this leads to the question that we're only measuring, as you indicated in your talk, intensity. In this case, the FDA has said the worst abdominal pain, which is kind of an interesting change that occurred for us because prior studies just measured abdominal pain, said rate your abdominal pain.

When the FDA required us to switch to worst abdominal pain, we were actually in the middle of the linaclotide trials going from phase 2b to phase 3.

Now, remember there was a lot at stake for the company by switching that one word, and we were quite nervous about it because we had no idea how people would respond. We had never asked. And it turned out that the responses were almost identical. So it means that most patients just used the word -- interpreted their pain as being the worst pain. That's the way we read it.

Page 157

But the question you brought up earlier, Bill, is we're only measuring intensity. We're not doing frequency, duration. There are all these other components, bothersomeness, that we don't include and should we include.

I'll just say one other thing, which is that, for years, we did bothersomeness. We did intensity and bothersomeness, two big components, and we found that patients actually responded identically, at least the IBS patients. And we actually tried to move the questions apart in different parts of the questionnaire, and it didn't seem to matter, so we dropped it a while back.

Do you have comments? Should we add these to our --

DR. CHEY: I think bothersomeness is definitely more of a global kind of assessment, but I do really wonder about this issue about, for example, frequency, in addition to intensity. And in a way, since you have the diary data and you know whether somebody reported zero, you do have that in a way. It's just not formally assessed

Page 158

because the way we're doing it now is worst abdominal pain.

So somebody might have 4 really intense episodes of pain in a day, but you're only going to capture that one parameter, one aspect of the pain; so actually, reading more about this and thinking about this in the lead up to this meeting, and feeling more and more like we would actually benefit from a little bit more of a deeper dive in regards to measuring pain and understanding how our drugs affect pain, different aspects of pain.

DR. LAI: Do you think the frequency or the intensity of what you call the IBS attack or maybe in IC what we would call flares, would that be a potentially meaningful outcome to look at, how often you have the attacks and how intense is an attack compared to your baseline?

DR. CHEY: Yes, I think it might be. I think that's an interesting hypothesis that remains to be tested. The other thing that we've learned from our recent work is this difference between subgroups. I mean, patients' experience of pain is

Page 159

really different in the IBSC versus the IBSD group, for example, and that's something I think we're just really starting to understand.

The D patients have pain around their bowel movements. So they get pain; they have to go to the bathroom. That's typically what happens. The C patients have this pain all the time. They feel uncomfortable, full all the time.

So it's very different. It's not universal, but from a pattern recognition standpoint, their experiences are really quite different.

DR. TURK: The frequency included in that definition, essentially, your second bullet is in fact a frequency measure, so you have both intensity and a frequency.

DR. CHEY: But that's for stool as opposed to pain, yes.

DR. TURK: Right. But you are looking at the frequency of the symptoms that are related to the condition.

DR. CHEY: Well, the frequency of bowel movements, yes.

Page 160

DR. PONTARI: Is there any IBS equivalent to where you have different locations pain? Are there different locations like low quadrant left, you know, rectal things? Do you guys have that?

DR. CHEY: Yes. So there's been a whole bunch of studies looking at that. And the common theme from the studies is that the largest proportion of IBS patients have pain in the left lower quadrant, but that's probably only about half. And then the other half have pain all over the place, like upper abdomen, lower abdomen.

You'll notice in the Rome criteria, it does not distinguish on the basis of location. What's interesting is if you look at the qualitative work by Brennan and Lin Chang's group at UCLA, it's interesting, though, that patients have these attributions like left lower quadrant pain, they say I have colon pain, or upper abdominal pain, they say I have stomach pain or esophagus pain.

So it's interesting. The patients attribute the location of the pain to some anatomical attribution, which of course may be completely

Page 161

wrong, but that's how they think about it.
DR. WESSELMAN: Kathy, was that related to this topic? Yes.
DR. VINCENT: I have two points. One is that if we're thinking about women with any of these conditions, they often have a cyclicity to their pain no matter what the underlying cause is, and I'm not sure how we capture that. I don't think we capture it very well.
We're currently running a trial looking at gabapentin and chronic pelvic pain with no underlying pathology. And one of the things that we did as we were preparing for that was surveying patients as to whether they were more interested in their average pain being reduced or their worst pain. So was it worse for them, the amount of pain they had in general, or were the really bad days worse? Fifty-four percent said they wanted their worst pain reduced and 46 percent wanted their average pain reduced.
(Laughter.)
DR. VINCENT: So it's an NIH all-funded

Page 162

study. We ended up with going dual outcomes of worst and average pain, but our statisticians aren't very happy with that for how to analyze it.
DR. CHEY: By the way, that's almost identical to the data we have in IBS, like almost identical. So to your point, the majority say that worst pain is most meaningful, but a whole bunch say yes.
DR. WESSELMAN: Was there a correlation to the number of comorbidities or the types of comorbidities for those two almost identical subgroups by number?
DR. VINCENT: That was literally just a quick patient survey on a website in order to answer that question of what we should be using for our primary outcome. When we actually come to analyze at the end of the trial, then we'll look at that.
DR. WESSELMAN: Ian?
DR. GILRON: Ian Gilron from Queens University. There's already been a lot of discussion about overlap among symptom-based

Page 163

conditions like IC, fibromyalgia, migraine. And in terms of trial methodology at least, I know from neuropathic pain, generally we try to exclude competing pain conditions.
So if you're trying to enroll somebody with diabetic neuropathy, and their average pain is 5 out of 10, and they've got osteoarthritis of the right knee, and it's on average 7 out of 10, we would tend to exclude them because of potential for misattribution of pain and pain intensity.
So I'm just wondering, whether for our recommendations whether we need to consider -- or first of all what has been the experience with trial recruitment with overlapping pain. Do you just turn a blind eye? They have coexisting fibromyalgia. Do you just forget about that, or how do we address that in clinical trial enrollment in symptom-based pain conditions?
DR. PONTARI: I've got a question for you. So from an anatomic standpoint, my understanding, from the studies from Pittsburgh, and probably you guys, and whoever else, is that there's more cross-

Page 164

talk between the bowel and the bladder. If you inflame the bladder, you get upregulation in the dermatomes, one above the other, and you get bladder inflammation.
So I'm not sure how you could -- we don't. So the first thing is we don't exclude people with IBS in IC trials or prostatitis trials. So I'm not even sure if that's a reasonable thing to do based on the neuroanatomy as far as -- it's a little different than having neuropathic diabetic neuropathy in rheumatology, but because of the neuroconnection, I'm not sure that's a reasonable thing in this condition.
DR. CHEY: I completely agree. I think the one thing that's happened in IBS trials, because there is data to suggest that patients with significant psychological comorbidity respond less well to particularly peripherally-acting drugs.
So most people have been excluding patients with significant psychological comorbidity, but I'm not aware of -- Tony, do you know of any? I don't think we've excluded any.

Page 165

DR. LEMBO: Not that I can recall, either.
DR. CHEY: Yes. Actually, to your point, though, a more interesting thing -- and this is coming out of this meeting, clearly -- is this idea of deep clinical phenotyping as part of any of the trials that we do. I'm really attracted to this idea of getting a much more comprehensive inventory of not only, in our case, the GI symptoms, but some of the other symptoms that have been discussed this morning.
I know Tony and I have suggested that, and we've met with some resistance because of the burden. Obviously, it's a very practical issue, just the questionnaire burden, and also finding things that you don't want to find out.
So pretty much narrowly focusing has been I think the theme of the day in most drug development programs. But it would be really valuable if we could do a deep clinical phenotype or deeply phenotype these patients to start to understand these patterns a little bit better.
DR. GILRON: Just to get back to that, so if

Page 166

you had someone with chronic widespread pain and maybe that was their predominant issue, but they happened to come upon your trial, which was an IBS or an IC trial, where do they fit?
I mean, should they be included in your trial if their fibromyalgia -- if their upper back/neck pain is equal or more severe than the symptom burden related to their visceral pain? I'm not trying to cause problems.
DR. CHEY: No, no, no. Let me just make a comment on this because this raises a point that I meant mention in my talk, but I forgot to. And that is that one way that academic as well as industry investigators have dealt with this is to put a cap on the amount of pain you can have.
In other words, you'll notice that the guidance said you have to have at least 3. Well, many trials, not most trials, but many trials will then just say they have to be between 3 and 7 or something like that. And so if you're the kind of patient you're referring to as having a lot of pain all over the place all the time, those patients get

Page 167

excluded during the baseline period.
So that's one way of dealing with that because, again, I think that the prevailing wisdom, whether it's right or wrong, is that those patients tend not to respond to certainly peripheral-acting therapies.
DR. WESSELMAN: Chris?
MS. VEASLEY: Chris Veasley with the Chronic Pain Research Alliance. Our organization deals specifically with overlapping conditions, and we have worked with industry. Industry has tried to enroll patients, whether it's low back pain, IBS, IC, vulvodynia, other conditions, that don't have other pain disorders. The problem is that they don't have enough people for the trial. So what most are doing is allowing patients to have other conditions, other pain disorders, but they're not tracking it during the trial.
For example, if somebody with IC has also migraine, and they're on day 15 of their trial, and they have a migraine attack, that's not being tracked. So that's obviously an important

Page 168

indicator, and it's certainly something that we need to deal with because there's a huge literature base showing that the more sites of pain you have, the higher treatment recalcitrance you have.
So it has to be affecting what outcomes are coming out of clinical trials, and I agree with you, we need to have a recommendation as to either how we handle that or how we track it, and understand what the bidirectional relationship with that is, just like we track mood and sleep.
Just going back to what Andrew said, I fully appreciate the comment about central sensitization because it's discussed very much in this community. Really, what we're talking about is pain syndromes that are driven by more CNS, more CNS-driven pain syndromes with mechanisms like central disinhibition and kind of what Clifford Woolf will describe as dysfunctional, a category of dysfunctional pain syndromes.
As Michel mentioned, the two patient-reported outcomes or surveys that are most indicative of having multiple conditions have been

Page 169

some measure of general widespread pain and number of somatic symptoms.

I'm going to push the panel a little bit further because my question to you is, we're looking at whether improvement in organ symptoms, so whether it's bowel function, bladder function, and pain as pain goes down and does that improve, how does that correlate with overall health-related quality of life and psychosocial function?

I'm very familiar with the vulvodynia literature, which is to say that if pain improves that quality of life or function, whether it's sexual function or physical function, does it necessarily correlate with that? And we've certainly seen in that in the general pain community.

I'm wondering, in IC and IBS, has that been studied? So if you have patients whose pain is improving, their bowel function or bladder function is improving, does that correlate with an improvement in health-related quality of life and psychosocial functioning, and is that dependent

Page 170

upon time? So basically patients who have new onset symptoms versus those who have pain for many years?

DR. LAI: I think they do correlate in the IC literature.

DR. CHEY: Yes. And I think as well, certainly not a perfect correlation, but generally speaking, the patients that have experienced a meaningful improvement in their abdominal pain are the ones that tend to experience an improvement in at least disease-specific quality of life.

I must say that, as an interesting point relative to the question that you posed, our trials have been very consistent in showing disease-specific quality-of-life improvements. They've been less consistent in showing general quality-of-life improvements.

So I'm not familiar with specific analyses -- Tony might be -- for a general quality of life, but for disease-specific quality of life, one of the main drivers of improvement of disease-specific quality of life is pain.

Page 171

DR. WESSELMAN: I think first was Frank, then John, and then Mark.

DR. TU: Thanks. Frank Tu from NorthShore Health. So if you go back to the stated goals for the meeting, which seems to center around this idea about defining endpoints for trials for chronic pelvic pain and irritable bowel syndrome, it seems like there's several dichotomies that keep on coming back and forth.

To speak specifically, this current discussion, what Roger said about is there a way to distinguish visceral versus somatic, it seems to be one of the critical questions. And the other one is this point that Ralph and Andrew have talked about, about whether or not we can measure central sensitization as a meaningful, stable construct in patients.

I'm still struck by this idea that we might be able to define subgroups of patients, and that perhaps one of the critical things to do here is to attempt to identify cleaner visceral-dominant patients versus cleaner somatic-dominant patients;

Page 172

clinical exam. Henry's talked about pelvic-floor exam. Andrew's talked about vulvar skin assessment. No one's brought up things like doing bladder distention or anal manometry, but those are your obvious provocative measures that could potentially be used.

Is that within the scope of what we're trying to do here, to actually see if we can define subgroups that would be described as being responders in subgroup A versus B based on some sort of a winnowing test?

DR. PONTARI: Frank, who can you do bladder distinction for?

DR. TU: Your patient who's got bladder symptoms, you simply can do a standardized challenge on them.

DR. PONTARI: You don't mean a hydrodistention. You mean just a bladder fill.

DR. TU: Drink on a consistent basis using a parameter like we use at MAPP. Those are all examples, but it seems like that's a critical question about saying that if a person fails on a

Page 173

vulvodynia trial, for example, is it because they never had vulvar skin sensitivity, and in fact they just had muscle tightness?

These are not always worked out in these broader trials. Typically, we'll use a cotton swab. I think the vulvar trial is a little more consistent in trying to define a pure subgroup.

But a lot of these IBS trials -- it'd be interesting if, Bill, you could comment. Has anyone tried to track and see if you are not a viscerally sensitive IBS patient -- so you have all the symptoms, but you have no meaningful anal manometry results, like you can be distended up to a large level and you don't display bother, are you more or less likely to respond to a given drug? That seems like that would be a critical question on endpoints.

DR. CHEY: Yes. So there's not a whole lot of literature on this, but there is some literature on this. And the bottom line is, unfortunately, visceral distention, like for example, rectal balloon distension or sigmoid balloon

Page 174

distension -- which was really popular for a period of time. I grew up doing distension studies in the lower GI tract as a marker of the biomarker for visceral hypersensitivity, and it turns out it's not a very good biomarker for visceral hypersensitivity, certainly not in terms of predicting response to therapy.

So unfortunately, the models -- for example, people for a while were doing animal studies with visceral distention to sort of predict whether there would be a pain response in IBS patients. And at least to my knowledge -- Tony, correct me if I'm wrong -- it really did not bear fruit. And we've as a community now gotten away from doing visceral hypersensitivity testing with balloon distention.

Now, I'm convinced that you're right. I'm convinced that the issues that you've identified are at the heart of the matter in terms of trying to subgroup these patients top-up/bottom-down. But as I said earlier, I really think that there are many patients where it's both.

Page 175

So it's not going to be as simple as doing one thing to identify one element that's abnormal. Unfortunately, I think that the patients have a number of different issues, mechanistic issues, that are going on, that conspire to cause their symptom experience.

DR. LAI: Do you really think a balloon distension can actually distinguish the central group from the peripheral group? Because I think even the central group is probably going to be showing sensitivity to balloon distention.

DR. CHEY: Absolutely.

DR. LAI: And if the peripheral group does so, I don't know.

DR. CHEY: I can tell you unequivocally at this -- I shouldn't say that.

(Laughter.)

DR. CHEY: I should never say unequivocally. But my interpretation of the data as it's been conducted to date is that, no, it does not help you to distinguish.

DR. RAPKIN: In the UCLA group, they found

Page 176

it more characteristic of women than males to have hypersensitivity with balloon distention.

DR. CHEY: The UCLA group did find a correlation between severity of abdominal pain and sensitivity to visceral balloon distention, but there was a lot of overlap. So the problem is that, yes, there are statistically significant differences, but could you use it as a biomarker to distinguish between groups? Probably not. Probably not.

DR. TU: In follow-up to that, I just wanted to comment on it. It seems like with Andrew and Ralph here, we can't just casually go off this question of what does it mean to have central sensitization in these conditions because that's one of the probably biggest -- whatever that is, if you say we can see it, it smells like central sensitization versus something you formally define with some sort of dynamic allodynia or some other sort of a measure.

I think it's important not to mix up somatic sensitivity, which the Michigan group has spent a

Page 177

lot of time working on, from true measures of physiological central sensitization.

If you look at work by Steve Bruehl's colleague at Vanderbilt, Lynn Walker, amongst children with functional abdominal pain, there's a small proportion that have very high levels of somatic sensitivity.

If you follow them five years later, they still have very high levels of questionnaire-based answers, like a 50-item complaint checklist. And there's something very different about that 15 to 20 percent that has 5 years of unrelenting somatic sensitivity that is not necessarily correlative with what their physiological measures on QST would be, I suspect.

I think it's important not to drag those two under the same group. It's easy to do as a clinician. I try to do that every day, but I've heard enough people say you can't do that to be casual and say, "Somatic sensitivity is central sensitization."

DR. CHEY: By the way -- and I suspect this

Page 178

would be true of the other conditions -- patients that have chronic symptoms, that patient does not have IBS. I mean, think about the definition. It's related to defecation, associated with change in stool frequency or associated with change in stool form.

If a person just has pain all the time that's unrelated to those issues, by definition, they don't have IBS. They've got something else.

DR. WESSELMAN: We have about five more minutes left, so John, you had a question, and then Bob -- John, Mark. So then Bob.

DR. DWORKIN: I don't remember. I think it might have been Michel who showed the slide of the pain scale that Pfizer used in developing gabapentin and pregabalin. And their wording was -- was it you, Mike?

DR. PONTARI: It was a pregabalin trial. It was our --

DR. DWORKIN: Yes, and that was from John Farrar's article. The wording of that question is very interesting. It was, "Please pick the number

Page 179

that best describes your pain in the past 24 hours." They didn't use either average pain or worse pain.

So it seems to me that based on the fascinating data that Katy just presented, maybe we should go back to what Pfizer was using with gabapentin and pregabalin; get rid of these weird words, "average," "worst," and just ask patients what number best describes their pain. And that approach would have satisfied your statisticians, Katy, because then you'd just have one question instead of two.

Actually, I don't know historically how it was we've now ended up with this question about whether it's "average" or "worst" when we started off 20 years ago with "best describes," but it sounds like John's going to answer it.

DR. FARRAR: "Answer" I think is too strong a word. I think what I'd like to do is to perhaps explain that phenomenon, which is that everyone describes pain differently, and we just need to come to grips with its subjectivity.

Page 180

The issue with any scale in a clinical trial is it just needs to be consistently used by the person over the length of the trial. And then if their pain, or whatever, is going to get better, then it will be reflected by that.

The problem is, in conversations with people who study pain in animals or who are very basic science oriented, or who have a very clear understanding of what it means to have acute or chronic pain, or what does worst pain mean, or what does average pain mean -- because it may be that if you ask people about their pain without specifying, some are going to answer with regards to its worst, as in 58 percent, you said, or something, 56, and 46 percent are going to answer with regards to average because that's what's most important.

That doesn't bother me particularly if you think that the treatment that you have would work for both of those instances. I think it might bother you if in fact you have a treatment that just gets rid of the worst pain. And the obvious example of that would be something like trigeminal

Page 181

neuralgia, where you have these acute episodes of very severe pain and nothing in between.

There is no right answer, I think, to that question. I think that you gather information by being careful about what you measure, and then it's important to think about how that affects how you try and interpret the results.

DR. DWORKIN: Sharon?

DR. HERTZ: Laurie Burke used to always describe work that was done. Laurie Burke is somebody who used to work at FDA, and she's on the outside now, enjoying life.

(Laughter.)

DR. HERTZ: But she's described studies in which when you look at people's report of average pain and you look at people's report of worst pain, the reality is, average pain pretty much skews towards worst pain because if you do a patient diary, and on a couple days, they're scoring 2 or 3, and a couple of days, they're scoring 8 or 7, you can average that pretty easily. But if you just ask them to think about a week and average

Page 182

their pain, that too is not coming to mind nearly as much as that 7 or 8.

So the reality is, as long as whatever the processing is -- sort of what John just said. As long as whatever it is they're reporting is consistent, it kind of doesn't matter. But we just need to understand that when we ask people to average their pain, that's not a skill that most people can do that would reflect an actual averaging of pain diary recordings.

DR. FARRAR: That's exactly right, and it gets to this issue, which was said earlier, which I didn't know that IBS actually used a global question for many years, which one can argue that really the question you want to ask is, overall, how bad is your life or how bad is the pain, allowing the person to integrate those pieces into a picture. It drives people nuts if you want to try and dissect it into pieces because you can't.

As long as the treatment that you're focused on is hypothesized to affect both the worst amount of pain that patients have or the average amount of

Page 183

pain that they have, then I think it's fine. And I honestly believe that in most cases, with most drugs that we're using these days, that's the case, so I don't have a problem with it.

It is clear that patients remember worst pain better probably, like on Tuesday I had a really bad pain as opposed to on average. I would argue, though, the other aspect of worst pain is that you get a bigger response in worst pain. But we looked at a nice study, and we did a nice study where we looked at that. And if you actually calculate the percent pain, change in pain with worst and average, it's identical, at least in post-herpetic neuralgia and diabetic neuropathy.

So I'm agreeing with you. I think what we're getting at here is that pain is really a more global measure than we think, and it's not necessarily the connection between nociceptor and the brain that we're actually measuring. We're measuring more than that.

DR. WESSELMAN: With this being said, I just got a note from Bob. So the webinar is going to

Page 184

start in this room at 12:00, which is basically in five minutes, and lunch is next-door. So we would like to conclude this session, eat, and then watch the webinar, and be back here at 1:00 for the next session.

(Applause.)

(Whereupon, at 11:54 a.m., a lunch recess was taken.)

Page 185

A F T E R N O O N S E S S I O N
(1:06 p.m.)

DR. SMITH: We're going to go ahead and get started. So this next section that we have after the lunch break is now going to focus on FDA perspectives and approaches to different both outcome assessments, clinical endpoint development, and also how they think about three of the different conditions, so prostatitis, IBS, and I guess interstitial cystitis.

So we're going to cover those things this afternoon, and then we'll have a discussion panel at the end. For the most part, we're going to save, again, questions, really process kinds of questions, until the discussion period. Dr. Wiederhorn did ask that he saved time in his talk specifically for clarifying questions. So for his talk, we will allow for some clarifying questions.

So the first person I'm going to introduce here is Sarrit Kovacs. She's a reviewer in the clinical outcomes assessment group at FDA. So

Page 186

Sarrit?

Presentation – Sarrit Kovacs

DR. KOVACS: Good afternoon. So I'll be presenting an FDA perspective on clinical outcome assessments, and I will be presenting my perspective.

Today, I'll be presenting four main topics, the importance of capturing the patient voice by encouraging patient-focused drug development or PFDD; the 2016 update to the 21st Century Cures Act legislation, as well as FDA flexibility in getting the patient voice heard; then a roadmap to selection or development of a clinical outcome assessment or COA by focusing first on defining the target patient population and conceptualizing clinical benefit for those patients.

Next, I'll discuss the importance of establishing the content validity of a COA with evidence from qualitative research with patients to ensure that the concepts of interest are being assessed properly. And last, I'll describe some considerations when using COAs to assess patient

Page 187

symptoms of pain and urgency.

So the patient's voice is important to consider when developing patient-reported outcome or PRO tools intended to assess how patients feel or function. Patient-focused drug development or PFDD is about engaging the patient throughout the spectrum of drug development activities. And as part of FDA's commitments under PDUFA V, FDA was tasked with conducting public meetings with patients and patient advocates focused on 20 specific diseases and conditions, and FDA has conducted more than 20 to date.

Each PFDD meeting resulted in the voice of the patient report that capture the patient and patient advocates' perspectives and experiences both from participants who attended in person and those who participated via WebEx.

The 21st Century Cures Act was enacted into law on December 13, 2016 and primarily affects activities of the Department of Health and Human Services and its agencies. And part of the aims of the Cures Act is to increase the involvement of

Page 188

patients and their perspectives in research and in the medical products development process. It emphasizes the need for patient engagement and directs the FDA to include the patient's voice in drug development and review.

Section 3002 of the Cures Act is focused on PFDD. FDA is required to publish guidance documents for industry addressing the topics listed here on this slide, including methodological approaches to collection and analysis of COAs for the purpose of regulatory decision-making. FDA is also required to conduct a public workshop on COA and better ways to incorporate COAs into endpoints.

PFDD will also aid in providing evidence to establish whether treatments are in fact providing clinical benefits to patients. Clinical benefit can be measured directly or indirectly. Direct evidence of clinical benefit is derived from studies with endpoints that measure survival or how patients feel and function in daily life, for example using a patient-reported outcome or PRO tool.

Page 189

Indirect evidence of clinical benefit is derived from studies with endpoints that measure other things that are related to how patients survive, feel, or function, for example surrogates or biomarkers such as endoscopic results. Indirect assessment means impaired justification for its value as a replacement for how patients survive, feel, or function.

So what is a COA? A COA is an assessment of a clinical outcome. It could be made through a report by a clinician, a patient, an observer, or through a performance-based assessment, and there are four types.

The remainder of my talk, I'll focus primarily on PROs, a type of COA based on a report that comes directly from the patient about the status of his or her health condition. PRO tools can be administered via self-report or interview and could include both a rating scale or an event log such as a bowel movement or urinary frequency diary.

FDA has developed a number of tools to help

Page 190

guide the development of fit-for-purpose COAs, and I'll give a brief overview of the first three tools listed on this slide, on my next few slides. And in terms of the last two tools listed here, FDA developed the 2014 Drug Development Tool, or DDT, Qualification Guidance for Industry, which includes information on the process related to CDER's COA DDT Qualification program.

In 2016, FDA compiled a pilot COA compendium as a communication tool for industry in an effort to foster PFDD by collating and summarizing COA information used to support labeling claims in many different diseases and conditions. And it's intended to provide clarity and transparency and to be used as a starting point for early drug development.

I'm sure some of you are familiar with FDA's 2009 Tri-Center PRO Guidance for Industry, and this guidance defines good measurement principles to consider when selecting or developing a well-defined and reliable PRO measure intended to provide evidence of clinical benefit to patients.

Page 191

The goal is to avoid labeling claims or statements that may be misleading or false.

The Code of Federal Regulations or CFR is a codification of the general and permanent rules published in the Federal Register by the executive departments and agencies of the federal government. And Title 21 of the CFR is reserved for rules of the FDA.

Part 314 of Title 21 relates to applications for FDA approval to market a new drug, focusing on adequate and well-controlled studies that include methods of assessments such as COAs that are well-defined and reliable, which is critical for drug approval and labeling.

There must be sufficient empiric evidence to support the COA's use in a target patient population, that the COA is measuring the right thing in the right way in a properly defined patient population, and the COA scores accurately and reliably quantify changes in patient scores over time. This is important in order to be able to confidently attribute patients' improvement to

Page 192

treatment effect.

All types of COAs, not just PROs, can benefit from the good measurement principles described within the PRO guidance. It's important to note that this guidance provides an optimal approach to PRO development. However, both flexibility and judgment are necessary to meet the practical demands of drug development and to ensure data integrity and interpretability.

We know that not every step of the instrument development and evaluation is necessarily relevant or feasible in the context of an individual drug development program, for example pediatrics, rare diseases, and in the spirit of flexibility, we are encouraging drug sponsors to leverage existing data and existing instruments before embarking on developing a novel COA when possible and feasible.

So next I will present the roadmap to clinical outcome assessment, selection, and development. As listed on a previous slide, the roadmap is one of the tools developed by FDA to aid

Page 193

in PFDD, and there's a link on the slide to a more detailed version of the roadmap if you want.

This tool has been extremely instrumental in helping FDA and external stakeholders to systematically think through the issues that need to be considered in sequence before making final decisions on clinical trial endpoints.

So again, we recommend careful consideration of column 1 and first understanding the disease or condition, including the natural history and patient subpopulations. Important for today's discussion, we need to consider the patient perspectives regarding what's most important and relevant to them.

After this, we can move to column 2 and conceptualize a treatment benefit or clinical benefit by identifying the concept of interest to assess and treat as well as identify in the context of use or targeted patient population for drug development and the appropriate COA type. Only after these two columns' activities are performed should one embark on identifying an existing or

Page 194

developing a novel COA appropriate for its specific use, shown in column 3.

Now, we'll move on to establishing content validity of a PRO tool with evidence from qualitative research with patients. As listed on a previous slide, the wheel and spoke diagram is also another tool developed by the FDA to aid instrument developers in developing COAs for use in clinical trials. This is an extremely pared-down version of the diagram.

For now, I'll focus on spoke 2, which is relevant to the FDA's assessment of whether a COA is well defined and reliable, acceptable to support medical product approval, and suitable to support labeling claims.

Empiric evidence should be generated according to good measurement principles to support the content validity and psychometric properties and performance of the COA, and this is regardless of whether one is using an existing or developing a new COA.

Content validity is the extent to which the

Page 195

COA measures the concept of interest, and it includes evidence that the items and domains of the COA are appropriate and comprehensive relative to its intended measurement concept, and the population, and use.

For FDA, the most critical consideration is whether content validity has been established, so qualitative data supporting content validity. And if that's not provided or not sufficient, the agency can't review or interpret any quantitative data that's submitted to support the psychometric properties or performance of the instrument.

So establishing content validity of the PRO tool requires evidence from qualitative research, so focus groups, one-on-one interviews with a sample of patients that matches the targeted eligibility criteria for the clinical trial. The qualitative research should provide evidence that the tool's instructions, items, and response options are relevant, meaningful, appropriate, comprehensive relative to the intended measurement concept and to the targeted patient population.

Page 196

There are some content validity considerations on which we focus for PRO tools. For example, are we asking the important and relevant questions of the patients in the assessment? Do patients consistently define and understand the concepts in the way intended? For example, can patients distinguish among abdominal pain, cramping, discomfort in consistent waves?

Do they experience abdominal bloating and divergent waves? Do some patients experience bloating as an internal feeling of fullness or tightness or, as other patients may describe it or interpret it, as a physical distention or swelling of the abdomen?

Are differences between adjacent response options meaningful to patients? For example, can patients meaningfully distinguish between a response option of a little versus some, quite a bit versus very much, severe versus very severe? In other words, would a one-category improvement from baseline necessarily constitute a meaningful improvement to the patient?

Page 197

I'll now discuss the use of COAs for assessment of pain and urgency in clinical trials. Irritable bowel syndrome or IBS guidance per industry recommends evaluation of abdominal pain using the 11-point NRS, assessing patient's worst abdominal pain in the past 24 hours.

Analgesic indications guidance for industry states that pain intensity can be measured by a numeric rating scale such as the 11-point pain NRS, or Visual Analog Scales, or categorical scale, but that there are advantages and disadvantages to each, and it's preferable to use a scale that's more sensitive to change and more interpretable, such as a disease-specific pain measure.

The BPI short form, item number 3, is an example of an 11-point NRS, and it's a well-documented measure of pain and appears reasonable for use to assess patient's pain intensity.

For consistency, we do recommend that there are verbal anchor descriptors, at least for 0 and a 10 rating. And if the sponsor plans to assess worst pain in the past 24 hours, we recommend that

Page 198

the word "worst" and the recall period of past 24 hours or last 24 hours are both included as part of the instructions and items for the PRO to maximize the chance of obtaining valid, reliable, and consistent data.

In contrast to the NRS, the pain Visual Analog Scale or VAS is a continuous scale comprised of a horizontal or vertical line usually about 10 centimeters in length, or exactly 10 centimeters in length, anchored by two verbal descriptors for each symptom extreme.

There are some concerns that the VAS may be imprecise because patients are required to visually differentiate increments in the line without any label tick marks and there can be sometimes inconsistencies in the length of the line due to formatting issues, especially with paper due to photocopying or printing.

So in principle, electronic VAS would be preferable if the sponsor uses a single electronic platform throughout the study and patients are restricted from zooming in or out or stretching the

Page 199

line.

There are some challenges and considerations to keep in mind when assessing patient's pain. The localization of pain on which the patients should focus should be clearly specified in the instructions and wording for the PRO. For example, should patients focused on lower abdominal pain below the bellybutton, pelvic area, or upper abdominal area? And it's useful to include a diagram of a body with a location of the pain circled in order to focus the patient and increase consistency across patients' responses.

It's critical to conduct qualitative research to obtain the patient input as to where exactly the pain is being experienced. Patients should be interviewed also regarding what the most appropriate and feasible recall period would be for their symptoms, and then that chosen recall period should be clearly stated in the items, instructions, and wording to ensure consistency across patients' responses.

For some conditions, having patients report

Page 200

on the worst pain is appropriate, whereas in other conditions, it may make more sense to look at the average of their pain, and this should be explored with qualitative research with patients.

So when assessing patient pain intensity as a prespecified endpoint intended for labeling claims, it's important to capture also concomitant medication or analgesic use. For example using a patient-reported log at baseline and throughout the trials as capturing these data would better characterize the patient's pain experience and aid in interpretability of the pain data.

In addition, drug sponsors should optimize the frequency and timing of pain assessments in order to capture meaningful data, and consideration must be made with regard to measurement of pain, whether it's an episodic or chronic pain condition.

There are some identified challenges in using COAs for the assessment of urgency, whether it may be bowel urgency or urinary urgency. There are some conditions where patients' experiences with urgency are considered when diagnosing the

Page 201

condition. For example, urinary urgency characterizes overactive bladder syndrome. Pain associated with urinary urgency characterizes as interstitial cystitis. Patient input is needed to better define the concepts of bowel urgency and urinary urgency.

It's important to note that it is difficult to measure your urgency adequately without knowing what level of severity and frequency of urgency is considered to be normal functioning and what's considered normal to the patient. There's also a need for qualitative research with patients to better establish what's considered a meaningful improvement in feelings of urinary urgency or bowel urgency.

Both clinical and statistical significance in findings will need to be demonstrated. Sometimes, small changes in PRON point scores can yield statistically significant clinical trial results, but these small changes in patient scores may not necessarily be clinically meaningful and may not indicate clinical benefit.

Page 202

I have here a few practical considerations to keep in mind when including COAs in clinical trials after content validity has been established. The first is that phase 2 trials represent an opportune time to evaluate psychometric properties in performance of a COA, including what constitutes clinically meaningful within patient improvement in scores prior to initiating a pivotal phase 3 trial.

Second, patient global impression of severity and change scales should be included as anchor scales in both phases 2 and 3 to help determine and confirm what magnitude of improvement may be meaningful to patients.

Third, the COA items and response options included in clinical trials should be the same across phases 2 and 3 for comparability of the data. And lastly, psychometric evaluation study protocols should be submitted to FDA for review and comment before initiating those studies.

To meet the challenges of patient-focused outcome measurement and to ensure that COAs are fit for purpose for drug development, we recommend

Page 203

early consultation and close collaboration with FDA throughout the drug development. This slide shows three pathways for engagement with CDER and to obtain COA review and advice.

The first pathway is within the context of an individual drug development program. Here, we review drug sponsor submissions and provide advice on the sponsor's proposed COA strategy when a COA is intended to support a labeling claim, even as early as the pre-IND stage.

The second pathway is within CDER's drug development tool qualification program outside of the IND pathway, where we can work with instrument developers to create and qualify COAs that meet unmet public health needs and can be used publicly across multiple drug development programs.

The third and final pathway is through the Critical Path Innovation Meeting or CPIM process, where an instrument developer or drug company can have an informal discussion and receive general non-binding feedback from the FDA on a COA in early phases of development, outside of an individual

Page 204

drug development program.

To summarize, the patient voice is important to consider when develop PRO tools, intended to assess how patients are feeling and functioning in their daily lives. There are regulatory standards that need to be followed to determine whether a COA is well defined and reliable and adequate for use in clinical trials. However, FDA maintains flexibility in our evaluation of the evidence, taking into account the evidentiary standards, feasibility, and practicality.

I presented some challenges and considerations to keep in mind when assessing patients' pain and urgency as clinical trial endpoints. Early planning and discussion with FDA is important to ensure that clinical trial assessments are fit for purpose and measure what's most important to patients.

FDA has developed numerous tools and pathways for COA development, review, and advice, and FDA is open to engagement early and throughout clinical trial endpoint development.

Page 205

There are some links.
(Applause.)
DR. SMITH: Thank you, Sarrit.
We next have Laura Lee Johnson, who is the acting director of the Division of Biometrics III at CDER in FDA.
Presentation –Laura Lee Johnson
DR. JOHNSON: Hello, everybody. It's nice to be here today. The Division of Biostatistics III actually oversees many of and helps service many of the clinical divisions that you're hearing from today at FDA, although actually not with Sharon Hertz. She is in a different division that my friend, Tom Permutt, actually is the director of. But we work very closely together because I help oversee all of our patient-focused drug development work across the Office of Biostatistics and in conjunction with several of the other centers.
So actually, this disclaimer is the wrong one because my slides still haven't finished clearance. But they were cleared in different ways

Page 206

for other talks, so hopefully it's okay. So I'm just talking for myself.
This is from our multiple endpoints guidance. The primary endpoint for determining that a drug is effective should encompass one or more of the important features of the disorder. It should be clinically meaningful.
Now, Sarrit was talking about assessments. She talked about measurement. We talk about outcomes. I'm talking about endpoints. So we're going to go a little bit through what is an endpoint, and I'll give you some of our technical definitions for that.
What is the statistical analysis, which is really how I'm going to take all this data, and then how are you going to interpret what comes out of that analysis, not just that you looked at the p-value at the end, but what is it that's really interpretable there? But even sometimes those p-values depend a lot on that analysis.
Then something that's dear to my heart is how do we discuss promotional materials? And I

Page 207

will tell you, at the end of the day, whenever I get fully stuck, what I think about is, what goes into labeling, what is going to be on that Super Bowl ad, and now let me work backwards; that, and then what do I hear all the patients talking about, all the murmurings in the room?
So a lot of my work, I've learned by standing in the coffee shop lines and just listening to the people around me. What is it that they're talking about? What is it that's important or that they are entrusting having changed?
So where a lot of these definitions are found are actually in this living document called BEST. So for a long time, within NIH and FDA and between the different agencies, we had different definitions of things like the word "biomarker."
So we were told we had to sit down and do something about that, and this is on the National Library of Medicines website. You will find all these definitions in there, and as they change, then these slides are out of date.
But we have an assessment, so that's the

Page 208

interpretation or evaluation of the measurement. The measurement is this value you've got using some test, tool, or instrument. And remember, that could be from a lab. It could be from a patient-reported outcome. It could be from a lot of different things.
You have some similar words, but I want to focus on the endpoint. So remember, we've talked a lot this morning. People talked about symptoms. They talked also a lot about really how you kind of diagnose people. You talked about concerns.
The endpoint may not address all of the diagnostic criteria, for example. Now, a lot of people sometimes think that it has to, but in fact, it may not. What is going to actually change? What is your actual question?
Your ability to predict something about a person may in fact not be -- unless that is what you're studying, your question is about, it may not be the appropriate efficacy endpoint that you're interested in.
So there are a lot of different ways that we

Page 209

have a lot of different measurements and assessments, but it may or may not be tied to the endpoint for the question of the study.

So here, I want you to think about the precise definition, types of assessments, timing of the assessments, the tools that are being used. Like, sometimes, I just see, literally, it's like physical function. I don't know how they're measuring it. I don't know what they're measuring, when they're measuring. But they say physical function will change, so that's supposed to be on a hypothesis test. It makes it very exciting.

Anyway, so other details , how the multiple assessments within an individual are going to be combined. And this gets back to this concept, if you're going to ask people about a 24-hour recall and you're going to have many momentary assessments, are you going to ask them to fill out a daily diary every day for how many weeks? Are they just going to do it for a couple of weeks at a few different periods of time? What is it that people are doing, and how am I going to sum this

Page 210

up?

So what Sarrit was talking about many times, I think -- and we focused on it, and we have to because that's how you build in quality, because when I get to the endpoint, all that quality work up front is what lets me know that now I have a chance at actually having an endpoint that's going to be reasonable and useful.

So the endpoint relates to the concept and the measure, but you also have to think about the statistics and that summary. And realistically we also think about this idea of the sensitivity, but is it going to likely predict benefit? And I say that because sometimes you have things that just will never change.

So I work also with a lot of cancer studies. If you've had a surgical resection for prostate cancer and now you have problems with your sexual function, they don't have a way to deal with that. That is not what their cancer therapies are going to be dealing with. They're like, it is important, but their new chemotherapeutic agent isn't going to

Page 211

change that.

So we also have to think about what is it that's actually the reasonable endpoint for you. If it's something that's important, but it's not going to change for 2 years and you have a 6-month study, again, what's reasonable?

Now, as a statistician, I also want to talk about the wrong analyses. So a lot of times, we assume things are continuous. We talk a lot about our 0 to 11, and it's like, yeah. And I'm not here tomorrow, but I know John is, and we've been on panels together, so I know he'll take care of this.

But we think a lot of times about you have the same interval or distance between responses on a scale, and we don't. And there have been various studies, especially in the pain literature, to talk about this. So we really need to be doing more ordinal analyses and fewer continuous analyses at times.

Now, sometimes it doesn't matter. I'll be honest. But sometimes, it really does. And thinking about -- we like the mean. Right?

Page 212

Everyone thinks about means. They think about changes in means, differences in means, stuff like that, but as Sarrit and others have mentioned, that mean difference between arms, it might be easier to show that difference, but it's really difficult to interpret the meaningfulness of the difference, which is why, many times -- and you'll see this in the PRO guidance -- we talk about, yes, you might be testing it at the population or the group level, but then you also want to do analyses that look at the individual level.

So we'll talk a little bit more about that moving on. Sometimes people say the mean total symptom score changes, and, again, what exactly has changed?

Now, what did work for one of the applications that we had was the mean number of symptom-free days. So they actually did have a continuous variable, and we said, okay, great. You have a difference, statistically significant. We don't know what it means. But then they actually had the mean number of symptom-free days and

Page 213

compared that.
So we also looked not only at that, we looked at how many patients, actually, what percent of them had a drop in certain number of episodes. And we had qualitative information. They had done interviews with patients, and they had a nice, representative sample. And from there, they said, yes, this is the amount that matters to us. So with that information, we were able to help make a determination.
Now, these mean symptom-free days -- and again, you've got to talk to people because a lot of times, there are trade-offs in these symptoms, but they may be willing to make a certain trade-off. With the evidence, we're willing to look at that because not all symptoms are expected to go to zero.
So maybe it's not symptom free, but as one of my previous bosses when I was at NIH used to say, she's like, "Listen, my pain was at a 10, and now it's at an 8, and I'm still sitting on my couch, and I can't play with my kids. I don't

Page 214

care. You have not solved my problem."
So when do we use ordinal regression? More frequently than we see it, but we should be using it more and more. But that's not all that interpretable, either, so it's not like I solved a problem too much.
The key thing to remember is that the instrument is not your endpoint. You have to think about the timing, the frequency, and what really should matter for you. So let's think about this PRO that measures symptoms, is used in trials, and this not dealing with the topics of today, but I want you to extrapolate it.
I had one, they're measuring all the symptoms, except the patients in these trials end up in the hospital, they may die, and some of them end up being put on mechanical ventilation. So they cannot answer for themselves, we are told.
So what's the endpoint in the trial? Well, they just wanted to use the PRO. We said, well, now you've got missing data. But it's not missing. I know where these people are.

Page 215

So what my endpoint really needed to be was something about the severity of the disease, which is a combination of symptom information, when I have it, or that they have gotten so sick, or died, then in fact I need to put that as that higher end.
So I saw the RADAR paper, and DOOR, and stuff like that, and those are things that I think can be very useful. I'll be honest, not a lot or all of our clinical divisions -- some of them have probably never heard or thought about it, but I know that Lisa LaVange and other people in our office have done a lot of work in this area and are very interested in it.
But you've got to think about what your endpoint is, that everybody can have measured and what it means. So you have these lovely daily diaries. The problem is, again, what's the analysis? Next to never, do I see people actually take every single day. How many days can be missed? Why are they missed?
So we go down to, like, item-by-item, day-by-day sensitivity analyses when we're looking at

Page 216

these. That's what we're now instructing our reviewers to do, because we found so many issues. And it's something, though, to figure out in your development phase, and as you're doing the ongoing development, where might this be missing and how can we improve the efficiency of our trials?
So responder analyses in the traditional sense, I make a line in the sand. These people have responded, these people have not. Statisticians in general hate these.
One of my colleagues has actually now done some research, but we haven't verified it yet. The problem is, sometimes you might actually get a gain in power. We always say you won't, that when you use a continuous outcome, you have more power. But every once in a while, something weird happens with a variance. You have a bimodal distribution. You might -- it might work out. But the problem really for me is that we never know what the heck this definition is, like next to never.
Now, I also have psoriasis, and the dermatology and dental group in my

Page 217

organization -- and for them, you clear. Okay. Well, when you see clearance and you hold that clearance, then they're considered a responder. Fine. There are some things that you can actually measure. But the threshold, also many times we set it the same across all the patients, and that may not be realistic depending on the baseline for the patients that are coming in.

So what is it that different patients want? But it needs to be -- and we do this in the rare diseases. I was listening this morning, and I was like I should have just put in more of my rare disease slides where we kind of had this pick-your-own-symptom approach. But you have to measure everything in everybody, pretty please with sugar on top, but you usually need larger sample sizes when you're going for responder analysis.

I have seen some of these that are called more continuous responder analyses, where really what they're doing is they're doing their standard actually mean comparison between study arms. They're really plotting a continuous distribution

Page 218

function, but they call it a continuous responder analysis.

What's nice is when people are fighting over -- and you do this, really -- it's not a statistical test. You say this might be my responder definition or this other number. Are my curves separated all the way through? But if you actually do any type of area under the curve type of test, it's the same as doing an ANOVA pretty much, just so you know.

But again, I get back to this idea that chronic isn't stable. Then you have episodic. So does the mean matter? How many days? Like what am I averaging and what's the endpoint for any of these?

This gets back to this idea, if I have 10 relevant symptoms but patient A only has 1, patient B has 3, they've picked their symptoms, what's my endpoint? Do I need them all to resolve?

I have one set of data where they said that they wanted no numeric worsening, to which I then asked, so you've got a 0 to 10 score. You're

Page 219

averaging this on data. Really, if it changes by 0.1, do we care? I mean, maybe we do, but does that mean that we don't approve the product?

So when you talk about no worsening, it's a statistical test to me. Is it non-inferiority? Is it superiority, but going in the negative direction? What is it that we mean when we say no worsening? You have to really define this for the statisticians, and think about it clinically, and think about the relevance of what we're doing.

So another topic that we sometimes see is this time to event. So they'll tell me time to pain progression or time to some type of symptom deterioration, although traditionally people thought about this; just they had an MI, or a stroke, or whatever.

But the problem here is, again, I've got to define what is that event. They never had the event at baseline or that symptom at baseline. This can in fact be sometimes a useful analysis. Do they ever develop it in the time to that? So sometimes, if you're looking at progressive

Page 220

diseases, this might be useful, but you have to be able to define it.

This is what many people have for their missing data plans.

(Laughter.)

DR. JOHNSON: And you can't. So just be aware that, at least at CDER and the FDA, we have instructed all 200 plus of our statisticians that they should assume missing not at random, so not missing at random. A lot of things that come in have the analysis assumption at least of missing at random. This means something to about 4 people at most in this audience, but just know sensitivity analyses don't assume missing at random.

So what's the endpoint? We've got to focus on this. What is that statistical analysis? How does it tie to it? How are we going to interpret it? And how do we discuss it?

So I was excited to see the multiple endpoints guidance sent out to everybody. This is a draft. I don't think we're still accepting comments on it, although if people have comments,

Page 221

go ahead and send them, and we'll figure out a way to handle that.

But it's trying to focus to say our world many times is broken into co-primary endpoints, where you must establish efficacy on all of these primary endpoints. And you now have a multiplicity problem here because if you don't win on all of them, you're out. So that's really essentially one test. But we're testing them individually. It's not a composite.

Then that next bullet is at least one of several primary endpoints is sufficient. So when we say that, our goal there is, we know that there is a lot of heterogeneity, or we don't really know which endpoint exactly it is, but we know this constellation of symptoms and there may be various ways to measure them. We're not sure. We're going to put them all out there. But we're basically saying, if you went on any one of them, and it's clinically significant and statistically significant, that's good enough.

So these are actually pretty different

Page 222

because sometimes when I'm in co-primary land, I'm like, if you don't solve both of these very important problems, no. And this one, I'm like, listen, if you can solve any one of these problems, go forth.

Then I have composites. Composites get misused a lot. Composites were put together thinking about things like MACE, so thinking about these rare events like, okay, you might die, you might have a heart attack, you might have these very big things but rare and important.

So I'm going to combine it together to have one clinical endpoint to be able to count all of these potential events that could occur.

But that's not how it gets used any more. People say, oh, all these things are kind of important. I'm going to combine it together, and usually it's basically one score at the end, a lot of times equally weighted even when there shouldn't be equal weighting to it.

But it can harm you because you may have some things that just never move, and that's going

Page 223

to drag down your power. Then you have other elements here that you have this continue, so people end up making a bunch of responder analyses or responder definitions and one big responder analysis and they really don't know what the responder thresholds should be.

So a lot of composites we see shouldn't be composited into a single endpoint. And then also I have to describe it again. I have to write it out so my grandmother would have understood it, and that doesn't work so well.

Then we also talk about multi-component endpoints and these clinically critical endpoints. They're too frequent as the primary endpoint, but we want to make sure people have them in there. And sometimes, you lose on a primary. Like, you want a mortality. And there are actually some methods that we have in publications that talk about how you can save alpha and recycle alpha, so that you can recycle your type 1 error there and maybe even actually get to that. So George Kordzakhia has those papers.

Page 224

So an endpoint hierarchy, we think about primary, secondary -- we go in some various orders, how we're going to spend out that alpha. But in general, if you don't include your endpoint in the multiplicity plan, then we consider it exploratory. I don't care what you call it. Your protocol can say it's a secondary. If it's not in your hierarchy, we consider it exploratory, and now you have to convince us it's so important that it's going up.

So it's always generally things like that. But if you're not pre- and well-defined endpoint, alphas not allocated, please don't get mad and contact your congressman and Janet Woodcock because we're going to probably say it's not going to section 14.

But we have all these badly behaved endpoints, too. Not all endpoints, even when they're common, are good. Percent change is a big one. It behaves very badly although it is very common.

Change scores in general are troubling.

Page 225

Responder definitions, we've already talked about. Now, there's this great page from Vanderbilt. They put it up. I'm not saying FDA agreed to it, but I think it does a very nice job of explaining in medical school wording what some of the problems are with these.

Now, the other problem is, sometimes you've just got to make a choice and move forward. So if you don't understand meaningful change for your continuous endpoint or variable, don't assume you know it for a binary responder variable or time to event.

These are my conclusions. The problems with assessments will lead to issues with the endpoint. Again, build in quality everywhere you can. The assay, the tool, the instrument, they are not your endpoint. I like endoscopy. It's not an endpoint. You've got to go a little past that. But many endpoints don't match the situation. They hurt your interpretability. And you've got to choose the analyses and conclusions you want, thinking about the whole picture in mind. So thank you very

Page 226

much.

(Applause.)

DR. SMITH: Thank you, Laura Lee. The next person we have is Roger Wiederhorn and who's a medical officer at the Division of Bone, Reproductive, and Neurologic Products at FDA. And he will be taking a few clarifying questions because he saved some time.

Presentation – Roger Wiederhorn

DR. WIEDERHORN: Thank you. I'm Roger Wiederhorn. I'm a medical officer at the Food and Drug Administration in the Division of Bone, Reproductive, and Neurologic Products. Essentially, I'm a subject content analyzer, working with regulatory requirement restraints. They let me out today to talk, though.

Now, from the standpoint of a regulatory perspective on interstitial cystitis, the initial study that we engage, that comes before the Food and Drug Administration for a new drug, is called an investigational new drug application. This is a request for authorization from the FDA to

Page 227

administer an investigational drug or biologic product to humans.

The IND needs to include the following: a protocol, chemistry manufacturing and control information, pharmacology and toxicology information, and previous unit experience with the investigational drug, as some of them have been used before for other indications.

This is reviewed within a 30-day period. We receive approximately a thousand or more INDs per year and approximately we review for the following issues.

Are there risks to the clinical participants in the trial acceptable? Is there adequate safety monitoring? Has the sponsor submitted sufficient supporting data to establish relative safety for the proposed indication? And is the trial design adequate to meet its intended objectives? At the end of 30 days, if the sponsor does not hear from us, they are free to proceed.

Now, as part of the NDA process, or the IND process, rather, safety issues are identified,

Page 228

effective and tolerable doses are established, proof of concept that the drug works is documented, and the sponsor following this may submit a new drug application to the FDA once all drug development activities have been deemed sufficient by the sponsor.

Now, usually the sponsors will meet with the FDA prior to submitting such an application for pre-NDA meetings so that we can reach agreement that these studies are adequate. In other cases, a special protocol assessment occurs where they'll actually show us their phase 3 protocols. We will critique them and ask for improvements prior to submission.

Now, in addition to the NDA review that I just mentioned, the NDA must contain adequate numbers of patients to assure safety. These are minimum requirements depicted in this slide by ICH E-1 agreements. 300 to 600 patients exposed for 3 to 6 months at the acceptable or contemplated clinical dosage is required to detect an adverse event frequency of 0.5 percent to 5 percent. We

Page 229

would like to have approximately a hundred patients exposed for a year for long-term safety monitoring.

The total number of patients treated with investigational drug is anticipated to be about 1500. Now, these are minimal exposures. Drugs that treat chronic, non-life-threatening conditions are really for those. But depending on the circumstances, larger trials may be required.

Not all phase 3 trials are successful, and between 2000 and 2012, the Food and Drug Administration approved 50 percent of 302 new molecular entity applications during the first submission. The deficiencies that did not result in approval include efficacy alone for 32 percent; safety, 26 percent; both safety and efficacy, 27 percent; and chemistry manufacturing and controls or labeling, 15.2 percent.

Now let's talk specifically about interstitial cystitis, which is what I've been tasked to discuss. Dr. Lai has given you a very good presentation of the clinical aspects of that. I'm going to just quickly summarize them again.

Page 230

It's a syndrome. It's characterized by urinary frequency, nocturia, urgency, superpubic pressure, and pain with bladder filling, relieved by emptying. Cultures are negative for infection. There's no precise definition of interstitial cystitis. The etiology and pathogenesis of this disease are unknown. Evidence-based definitions of the disease are lacking, and our understanding of this condition relies largely on expert opinion.

Now, what are the NIDDK criteria? I'm going to briefly summarize those. Admittedly, there's controversy as to whether or not glomerulations are significant at all. But to be diagnosed with interstitial cystitis, patients must have either glomerulations on cystoscopic examination or classic Hunner's ulcer, and they must have either pain associated with the bladder or urinary urgency.

You must have, again, for inclusion in trials that the FDA requires for approving a new drug to treat interstitial cystitis, glomerulation or Hunner's ulcer, the pain that we've talked about

Page 231

associated with the bladder, urinary urgency for 9 months. You have to have a small-capacity bladder based on cystometry, and you have to have an intense urge to urinate at 150 ccs of urinary volume. Daytime frequency of greater than 8 for at least 9 months is required. And again, you're excluded if you don't have or fulfill any of these criteria, if you have involuntary bladder contractions and if you have the absence of nocturia.

These are a little bit out of order. The regulatory reason why we insist on the NIDDK criteria is to define a homogenous population of interstitial cystitis patients suitable for enrollment in clinical interstitial cystitis trials.

These criteria are not to define the disease, but to ensure that in any group studies that adhere to these inclusion/exclusion criteria, the populations will be relatively comparable.

Now, it's assumed that all these patients will present with symptoms of urgency and what they

Page 232

consider frequency. So these symptoms are not included as positive factors.

A rough estimate of the prevalence of this disease using IC NIDDK criteria or suggestive criteria compatible with NIDDK is 0.1 percent to 2.3 percent in the U.S. population.

Now, what are the measurable symptoms of interstitial cystitis? Pain obviously is paramount. We've heard several presentations today on how pain can be qualified. Is it severe? What type of pain? Is it constant? Is it there every day? Is it visceral? Is it somatic?

This really hasn't been well defined. It's been thought about. And I also want to point out that consideration of pain alone does not do adequate service to patients who have interstitial cystitis, and in fact, it may take our attention away from the other features or other facets of this disease.

This is a severe, debilitating disease, and I don't want anyone to make light of it for patients who have it. It's very variable. Urinary

Page 233

frequency is something that should be considered, how often do they get up at night? Urgency, as Dr. Kovacs just showed, is hard to measure. We don't have an acceptable measure or way of reporting it at this point. And Dr. Lai has alluded to flares of the disease. That may or may not be something that could be built into our protocols for us to look at.

Now, Elmiron is a drug that was approved for interstitial cystitis. The endpoints used for approval were really measures of pain. This was approved in 1996, and it's classified as an orphan drug. I'm using that just for an example.

Dr. Kovacs has talked about PROs. And from my standpoint, patient-reported outcomes allow the capture of disease aspects not felt to be previously quantifiable or discernible. And in some cases, it's a formidable undertaking for anyone to undertake the development of a PRO. But in doing so, new aspects of the disease may be discovered in addition to developing a measurement instrument.

Page 234

The virtue of PROs is they come directly from the patients, so we put the patient back in the mix saying what's important to you about your disease?

At the current time, PROs are not used to diagnose IC or as efficacy endpoints, as none have been shown to be accurate and reliable. Now, if you're going to develop a drug or PRO for interstitial cystitis, what are some protocol or design considerations?

Generally speaking, we want two double-blind placebo-controlled studies. You must specify the baseline maintenance therapy that's acceptable because patients who have severe interstitial cystitis, it would be unethical to take them off the medication.

Flares should be defined, and criteria of severity stated as well as what indication you will use in the protocol for rescue therapy. In other words, we don't want patients to drop out because they're unhappy with how they're doing in your protocol, unless your therapy may allow them to

Page 235

continue. That will need to be documented and to be one of the things that's measured, and what type of rescue medication you're going to use has to be prespecified.

Now, going forward, what's needed? We really need -- and Dr. Lai and I are in agreement about this -- non-invasive diagnostic methods for IC. Cystoscopy and cystometrics are invasive. We need biomarkers and other diagnostic tests that can be easily done.

Now, well-defined and reliable measure of urgency is also necessary. Patient-reported outcome has already been suggested, and the FDA currently wants electronic source data in clinical investigations as opposed to paper submissions.

Now, I've given you a very high-level view of what's going on. I have additional slides to talk more in detail about the various steps in development of the drug, or again, I'm not sure I gave you what you were looking for with this, so that's why I left extra time to field questions.

(Applause.)

Page 236

DR. WIEDERHORN: Thanks. Are there any questions?

DR. PONTARI: I guess the point is that you can't develop a PRO without the criteria first; correct?

DR. WIEDERHORN: Without doing what?

DR. PONTARI: Before you can develop the PRO, you need the criteria resolved first. Correct?

DR. WIEDERHORN: Well, no. We have criteria at this time. They're the NIDDK criteria. You may not agree with that, but that would be the basis on which we would at this time ask patients to be interviewed, to characterize interstitial cystitis.

So we have criteria. They're old. At the current time, we're not aware of anything better. Now, I know there are studies showing comparable patient groups might have comparable outcomes, but we would like data-driven outcomes to verify that.

Yes?

DR. BUTTERFIELD: The criteria used are focused on obviously patients with defined IC that

Page 237

have either Hunner's lesions or glomerulations. And as Dr. Lai and Dr. Pontari were talking about, there's obviously the BPS patients or patients that may not have a defined or visible inflammation in the bladder.

So it seems to me that, if you were to use this criteria, you really would be only selecting an IC population. So what would you recommend for people developing drugs for bladder pain syndrome?

DR. WIEDERHORN: Well, at the current time, those are our only criteria for IC. Both classical IC and BPS are a multi-factorial disease with multiple causations. If you stick to the NIDDK guidelines, there may be les possibilities within that group than there is in bladder pain syndrome.

We really don't know what's in the bladder pain syndrome definition. What we have done, however is that we have met with companies who are interested in developing bladder pain syndrome drugs. And what we've said and advised was that you need to have a 3-track protocol, which would include placebo, classical IC patients, and painful

Page 238

bladder syndrome patients, and let's see how they do. Do they behave comparably? Do they behave differently? Because if they behave differently, we may be surprised to find that the drug works better in bladder pain syndrome without an IC or vice versa.

But the point is, we don't know at the current time what bladder pain syndrome is, we don't know how many different causes there are. They're all expert opinion and so we're arguing based on expert opinion that there's less variability within the classically defined IC patients than there is in the painful bladder patients. And that's expert opinion. We don't have evidence for that.

Yes?

DR. DWORKIN: So what you're asking us to do, of course, is onerous, which is a study --

DR. WIEDERHORN: It is.

DR. DWORKIN: -- that's stratified for classic IC and also BPS without you giving them any assurance that, at the end of the day, all of that

Page 239

money and effort is going to pay off.

So I guess my question is, is there anything a sponsor could do data wise in terms of providing you with evidence that would allow you to move forward from the 30-year-old definition to a more modern definition that includes BPS within a kind of expanded IC/BPS, or does it have to be an onerous clinical trial?

DR. WIEDERHORN: I think I would defer -- I participated in the MAPP. I'm participating with the NIH in the Lower Urinary Tract Research Network. But I would defer to Dr. Clemens on that because the idea is, are we going to have data-based phenotypes that we would feel comfortable with?

I'm not aware of that yet, and I know it's onerous, and it's also a factor in why a lot of development hasn't occurred. The interesting thing is when you read the transcripts of the 1988 NIDDK criteria and stuff, a lot of the theory and stuff is very similar to what we're reading now. Yes, there's a few new wrinkles and stuff, but we

Page 240

haven't gone forward very much at all with the field. I agree with you, it's onerous.

DR. CLEMENS: I guess one comment I'd make is that the 1988 criteria were entirely not evidence based. So why is it that we're using 30-year-old non-evidence-based criteria and don't have openness to using contemporary criteria, which are admittedly not completely evidence based?

An example, and the one that particularly bothers me, is the cystometrics. We don't ask people with chronic back pain, see how much you can lift or et cetera. It's a painful test for patients that has been just completely abandoned for IC patients.

So it's more of I guess a somewhat editorial comment. But ultimately, I think there are data that I'll show tomorrow that will provide some evidence that there are differences between patients, and that, at least from looking at treated, natural history, et cetera, do seem to distinguish between patients.

But ultimately, the one we talk about, if

Page 241

what we're saying here is that we need biomarkers for IC, that may never happen. So I think perhaps what we're looking for is there a way that, for instance that -- for instance, could the NIDDK potentially convene another summit and try to put together, do the best we can, what is the current data that we have, but understanding that going into that, some of it's going to be evidence based -- or some of it's going to be opinion, but there's been probably an improvement in the opinion over the last 30 years.

DR. WIEDERHORN: I had a conversation with Dr. Star, just saying what you did, because he asked us, when is the FDA going to change the criteria, and I'm going to say when the NIH has a meeting. It's sort of the chicken or the egg. We stick to this because at least we feel we have a firm foundation as to who's included in the trials. It's not for decision-making clinically, clearly. The idea is we want to have a uniform patient population we're testing.

DR. CLEMENS: I was going to say, perhaps

Page 242

those taking notes for this, it would be nice to have that in the records, that then some of us can maybe talk to Dr. Star and see what we can do.

DR. SMITH: This is probably a great place for us to stop, and we'll take a break and come back at 2:30. Thank you.

(Applause.)

(Whereupon, at 2:07 p.m., a brief recess was taken.)

DR. SMITH: We have two more talks and then the discussion period. This next talk will be by Jordan Dimitrakoff, who's a medical officer at the Division of Bone, Reproductive, and Urologic Products at FDA.

Presentation – Jordan Dimitrakoff

DR. DIMITRAKOFF: Thank you. Thank you for the kind invitation. Thank you for having me. I am another one of the medical officers in the Division of Bone, Reproductive, and Urologic Products at the FDA. The disclaimer is very important, since you will see that some of the things that I will talk about do not really reflect

Page 243

the views of the FDA or the Division of Bone, Reproductive, and Urologic Products, where I work.

So in the interest of transparency, I haven't actually seen any of the slides --

(Laughter.)

DR. DIMITRAKOFF: -- which I think is a good thing, and you will see why. My talk has two parts, and I didn't put this in a separate slide, but the first part really talks about some of the things that we've been talking about the whole day, and the second part actually talks about biomarkers.

So you heard from Dr. Pontari in the morning about prostatitis and chronic pelvic pain syndrome. I just want to make a comment. The talk is listed in the agenda as Regulatory Aspects of Chronic Prostatitis, but I think everyone knows that, realistically, I will actually not be talking about the inflammatory type of prostatitis. I'll be talking mostly or exclusively about chronic pelvic pain syndrome.

So again, Dr. Pontari expertly described the

Page 244

NIH classification, which has been around since 1999. The previous slide actually showed you a nice outline of the four categories of prostatitis syndromes. And what we're really focusing on today is chronic prostatitis or chronic pelvic pain syndrome. This is the category 3 in the NIH classification.

This description is the original description that comes from a letter that was published in JAMA from Drs. Krieger, Nyberg, and Nickel at the time. It was in 1999. And I think it nicely describes the characteristics of this category 3, chronic prostatitis/chronic pelvic pain syndrome, or as we usually call it now, CPCPPS.

Basically, it's another one of the chronic pelvic pain syndromes. It's in males, and that's why, hence, the prostatitis part. And it by definition is characterized by pelvic or perineal pain, which arbitrarily at the time was defined as being present for at least 3 months within a 6-month period with or without varying symptoms, as we talked about in the morning and with or without

Page 245

erectile dysfunction or sexual dysfunction.
So I just wanted to make sure that -- again, I wanted to reiterate some of the points that Dr. Pontari made in the morning, that although CP, chronic prostatitis, and CPPS, chronic pelvic pain syndrome, are usually combined in clinical practice, it is unclear at this time whether it is appropriate to combine them for the purposes of clinical trials, which are intended to support a specific approval of a drug therapy.
So the reason for this, as I understand and this highly educated audience understands it better, is that CP and CPPS may actually reflect different conditions with different underlying etiology and different pathophysiologic mechanisms.
So again, this is a point which has been very well made in the literature by Dr. Pontari, Dr. Krieger, and the chronic prostatitis network at the time, and more recently by the MAPP network, that CPPS, the chronic pelvic pain syndrome, is a vague condition.
The source of the pain is really unclear.

Page 246

And most of the time, when we see a male patient in clinical practice, we categorize it as chronic prostatitis or chronic pelvic pain syndrome, as I mentioned earlier, but the reality of this is that the pelvic pain could actually be originating from a lot of different structures in the pelvis, the prostate, the pelvic floor, the bladder, as we talked about earlier, in the people that probably have a localized pain condition. And CPPS could and probably does represent a number of different heterogeneous disorders.
So there is an urgent need -- and the FDA really acknowledges and understands that, that there is an urgent need to learn more about CP and CPPS and what these conditions actually are. And that feeds nicely into the discussion we've been having the whole morning, going into this afternoon today, that one potential approach to do that is to phenotype patients or to try to characterize them as best as we can clinically and using biomarkers.
So the next three slides -- I always hate people that show slides and say, oh, this is a very

Page 247

busy slide and don't really pay attention to this. So I'm actually going to walk you through the slides. This is just an overview of what the slides are.
But this was really a dream. It was a dream which turned into a hypothesis, which we wrote as a proposal to the MAPP network back in 2008. So I say that I haven't seen these slides, and these slides are actually 10 years old.
So 10 years ago, the idea was, well, how do you actually approach patients with chronic pelvic pain syndrome. And we wrote this grand proposal in response to the RFA, which was issued by the NIDDK for the MAPP network.
So the idea was -- I know that the colors are nice; I hope they project nicely, as you will see on the other slides. But you can ignore the population at the top. This was just a sample population.
If you start with a population of patients that are predisposed to pain or have some degree of pain predisposition, how do you actually tease

Page 248

apart the systemic versus the local factors that might be involved?
I'm just a simple urologist. I don't know as much neuroscience or any of the neurologic stuff as the people in the room. So at that time, I tried to think how would I actually do a study looking at the different factors that might underlie chronic pelvic pain syndrome?
There have been different theories, and the theories are listed on the left side. And these are different theories, at least for, at the time -- that was in 2008 -- 2007, December 2007-2008 -- of what causes prostatitis and chronic pelvic pain syndrome. There's a theory about dysfunctional voiding. There's a theory that it occurs after trauma, after infection, inflammation, nerve damage, and autoimmunity.
Then there are all the other systemic factors that we have been hearing about in the literature that some patients are different. There is dysregulation of the HPA axis. There is dysregulation of sympathetic neurosystem. And

Page 249

again, you can totally tell these apart because this is not evidence based. This was just a hypothesis-generating study, which was required in the RFA.

Then we propose some cohorts that we can use potentially and that we had, that we can look at, and then look at this whole sequence of events, which is totally probably irritating to you because it's, again, not evidence based, and it was based off some neuroscience papers that were contributed by the participants at that time.

I hate to talk about central sensitization because we heard about this in the morning, but again, this is the urologist's view of the life beyond Fifth Avenue.

So this was the phenotyping plan that we proposed in the grant. And again, just to give you a better idea of things that you were asking about in the morning, what I thought at the time was, as a urologist, what do you do?

Well, the NIH at the time in the RFA didn't want us to say what criteria we were going to use

Page 250

to record patients because they didn't want us to get married to the 1998 NIH IC criteria. So the way we decided to go around that is they said, well, we will tell you what the criteria will be. That's why there's an asterisk, and you see it on the next slide. But it says that the patients will be recorded based on the criteria, which is what the MAPP is doing.

I'm not part of MAPP. I don't know what's going on in the MAPP. I haven't been ever a part of the MAPP, so I'm just talking as a person who conceptualized that at the time.

So the idea was, well, if you take patients that meet criteria and don't meet criteria, the other idea was to look at patients with overlapping conditions. And the RFA wanted us to look at patients with urologic chronic pelvic pain syndrome, which is the UCPPS abbreviation, and at least one other comorbid condition.

At that time, again, they were listed in the RFA, and we decided to look at fibromyalgia, CFS, or IBS. That was my list. Again, this shows the

Page 251

limits of my understanding, but these are all the questionnaires I could think of at the time. And of course, the comment was that no one can actually fill out those questionnaires because they were so onerous.

But the idea was, well, if you take all these patients -- one of the major challenges, as you've heard in the field, is there's a belief -- and I understand from what I heard in the morning that it's not so much of a belief, but it's becoming more evidence based.

There is a reality actually in the clinic that we know that there are patients who have -- people call it different things -- pelvic-based disease and people who have symptoms, which are outside of the pelvis.

So my simplistic thinking at the time was, well, how do I know? Well, maybe do a pelvic-floor MRI on everyone and look at the pelvic floor. And if the pelvic floor is fine, we'll just see, well, they probably have systemic disease, and if it's not fine, then they probably have a localized

Page 252

disease. And that's why you have this thing at the bottom, which says pelvic-floor MRI, spectroscopy, and all those different things, which are outside of my area of competence. Then at the same time, the idea was, well, we'll probably do an MRI of the brain, which is what the MAPP is doing right now.

So the idea was, again, taking people who meet the criteria, looking at the symptom domains, then looking at a way to somehow differentiate between people who have pelvic disease versus those who have outside-of-the-pelvis disease, and then in the next step put everyone through -- that's the thing you see at the top.

So then the next thing is looking at biomarkers, and then the biomarker part was, again, not evidence based, but it was an idea of looking at people and looking for evidence for evaluating their hypothalamic-pituitary adrenal sympathetic nervous system, look at different biomarkers in the blood.

This is all from the literature in this area. We came up with an arbitrary score, which

Page 253

hasn't been validated, but it was just an idea at the time, look at the adrenal and then look at the pain memory or system and do some sort of genomic/proteomic evaluation.

So that was just a dream, and that was the way I would still think about a lot of those patients. And here is the asterisk, which says that patients will be diagnosed using the 2008 NIDDK criteria at the time.

So the real challenge, as we talked about, again, is that in these patients, it's actually good to have a biomarker. It's a good idea to have some measurement. And Dr. Laura Lee Johnson gave you a very nice outline of how the FDA actually looks at endpoints and biomarker endpoints.

So I actually wanted just to give you a few bits of the regulatory approach to biomarkers from the point of view of the FDA. This is the definition of a biomarker. It's something that measures and is indicative of a normal pathogenic biological process or response to an intervention. These could be molecular, histologic, radiographic,

Page 254

physiologic.

It's not a COA. It's not a clinical outcome assessment about which you heard from Sarrit Kovacs earlier. But again, the biomarkers may be used by the clinical and the research community for a number of different things.

I think these are very important points when you think about the biomarker, which I didn't think about when I was doing research. But now, looking at this from the regulatory point of view, it's important to have reproducibility of data, to have adequacy of the analytic device, and feasibility of the marker should the drug be approved.

I think we are quite far away from this in the CPPS field now, but I think it's helpful to think about those things up front before designing trials. Again, I think it's very good. I urge you to review this source. It was a very nice public workshop, which was FDA and the Duke Margolis Center for Health Policy.

There is a very nice YouTube video where you can watch the whole public workshop, which was just

Page 255

a couple weeks ago. And it talks about biomarkers in a different setting, but I was listening to this, and I heard so many things that are really relevant to the field of CPPS.

The one thing that I think is very important when you think about the biomarker -- and I sense that is some of the talks in the morning -- is to think about the context of use, which is the second definition here.

The FDA has a regulatory process and requires a statement that fully and clearly describes the way the medical product development tool is to be used and the medical product development-related purpose of the use.

So what this means is that -- and I'm sure I'm not telling you anything new; you are familiar with all those things -- it's important to think about the biomarker in terms of biomarker for research purposes, biomarker for diagnostic purposes, and biomarker for prognostic purposes. And these are all different things, and there are different ways to quantify those.

Page 256

There is an important analytical evaluation and clinical validation process, and, again, those are very well described. We do have a guidance and I've listed this on the last slide. But I will also urge you, if you're really interested into biomarkers -- and I know that the MAPP has a large biomarker group that's probably involved in this.

I urge those of you that are interested in biomarkers for chronic pain, chronic pelvic pain, to review this white paper, which was put together by a consortium, by the FDA and the C-PATH Institute. It's available again at this Duke Margolis Center for Health Policy website.

It's a very nice paper describing all the challenges of how you actually define, how do you develop, how do you discover, how do you define, and how do you actually analytically and clinically validate a biomarker?

So a lot of the things we're talking about in the morning were mostly related to the clinical validation. Again, you have to remember that there is also a process of analytical validation, which

Page 257

is not easy.
So I had to show you this figure, but it relates to my next slide, and one of the people reviewing my slides made a comment that if I show the next slide, you wouldn't be able to understand the context of the slide.
So this is a figure from Dr. Schaeffer's paper in the New England Journal of Medicine. It's a procedure that we used to do in the prostatitis field. It was described in 1968 by Edwin Meares and Tom Stamey. And it was an interesting way to think about this because at the time, they actually described this as a procedure of proving that someone actually has bacteria in their prostate.
So the idea was how do you know, how do you find out if a patient has bacteria in their prostate. Well, what you basically do is, you take a first void urine. They urinate in the morning with a full bladder. Then they stop, and you take a second sample. And then you basically do a prosthetic massage, where you push on the prostate, you get the fluid, and then you make the patient

Page 258

urinate afterwards.
So the idea was that if you have bacteria, the way they described it, in the so-called VB3 sample, the sample after the prostate massage, and if the bacteria are in a higher number than the VB2, then the bacteria are probably coming from the prostate. If you have bacteria in EPS, which is the expressed prostatic secretion, and you don't have the same bacteria in the first or the second sample, they are probably coming from the prostate.
So Dr. Nickel actually published a paper about 10 years ago where he proposed using just one sample before and one sample after, so the VB2 sample is called pre-M, which is pre-massage, and then the urine after the prosthetic massage is the post-M.
So the idea is if you compare the two samples, and if you find something in the post-M you don't find in the pre-M, that means it's coming from the prostate. So my time is up, but I'll just finish. I have one more slide.
So that is what we did at the time. And the

Page 259

dream was, well, if you take pre-M and post-M patients' urine samples from patients with CPPS, you put this in a mass spec, do a programmatic study, and then you use bioinformatics to look at patterns, you would probably eventually be able to come up with a biomarker. But at the same time, if people have systemic disease, maybe you can look for biomarkers in the blood and do the same thing.
So the dream really I think in the biomarker field, and prostatitis, and CPPS, is to have something like this, which is a very old slide again from the cancer field, where you would actually take a patient, not probably do a biopsy, but have like a non-invasive way of doing this, and that's why we're using urine to look at the genes, the proteins, and then come up with an individualized approach to the patient.
So I'm sorry. I probably ran out of time. This is what a biomarker will actually give you, what kind of information it will give you. And just in conclusion, just to say that once we have a good handle on what those conditions are, we can

Page 260

best tailor the inclusion/exclusion criteria for clinical trials.
Again, as Roger mentioned earlier in IC, the idea here is to have a homogeneous group of patients in a trial that may be able to give us an opportunity to detect treatment response if one exists as opposed to enrolling different people having the syndrome and a heterogeneous population, some of which might not respond to treatment, and that would dilute the therapeutic effect.
So this is the final slide, and I just wanted to state that at this time the likely endpoints for trials in CPPS, the chronic pelvic pain syndrome, in men would be PROs, which look at patients' pain, plus/minus voiding dysfunction, plus/minus erectile dysfunction, going back to definition.
There is a PRO guidance that Dr. Kovacs mentioned already, but I'm not aware at this time of any PROs that are validated by the FDA for drug registration trials for CPPS.
So these are the references, and I have to

Page 261

thank three people at least, my team leader, Dr. Kaul; our deputy director, Dr. Gassman; and our division director, Dr. Hylton Joffe. Thank you for seeing this.

(Applause.)

DR. SMITH: Thank you so much, Dr. Dimitrakoff.

We have Lesley Hanes next, and she is a medical research analyst at the Center for Drug Evaluation and Research at FDA.

Presentation – Lesley Hanes

DR. HANES: Good afternoon. I'm Lesley Hanes. I'm a medical officer in DGIP, which is the Division of Gastroenterology and Inborn Products at the FDA. Thank you for inviting me to speak to you today about some of the select challenges in IBS clinical trials and to provide a regulatory perspective in new drug development for the treatment of IBS.

So a lot of what I'll say today is a recap of what you heard from my FDA colleagues, from Dr. Chey, but hopefully you'll be able to glean

Page 262

some new information or have some questions for me during the panel discussion.

I think it does highlight what I'll speak about, even though it's some of the similar things they've heard before. It's because we're collaborating together, so we're pretty much all on the same page.

So I have no disclosures, and these are my views, not necessarily of the FDA or DGIP. Today, this is a brief overview of my discussion, a plan to talk about basic regulations for drug approvals, because this is an FDA talk; select challenges, including pain-related outcomes in IBS trials intended to support drug approval; and the FDA guidance for industry regarding IBS and how it can be used to address some of these challenges.

Since I'm giving the regulatory perspective on the challenges in drug development in IBS, the first couple of slides will review the laws and regulations that guide regulatory work.

In brief, an improved drug must meet each of the following statutory requirements for the

Page 263

proposed patient population. The benefits must outweigh its potential risk. There are specific manufacturing requirements that are required. And it needs to have labeling that is evidence based and adequately provides guides providers and patients to use the drug safely and effectively.

The 1962 Drug Amendments to the Food, Drug, and Cosmetic Act requires the establishment of drug effectiveness as a prerequisite for marketing approval and that the effectiveness is demonstrated by substantial evidence.

So what does this evidence entail? The evidence includes the findings from trials that are designed well enough to distinguish the effect of a drug from other influences such as spontaneous change, placebo effect, or biased observations. And typically two adequately well-controlled studies are required to support a new drug approval.

As you have heard, a key goal of any clinical drug development program is to demonstrate safety and benefit of therapy. How is benefit

Page 264

defined? It's defined as a favorable effect on a meaningful aspect of how a patient feels, functions, or survives as a result of treatment.

It must be clinically meaningful, measurable, and interpretable. And in accordance to the statutory requirements, this observed clinical benefit is described in labeling as a claim or claims using words that represent the measured concept.

So moving forward, we'll focus on how this relates to the development of treatments for patients with IBS. In brief, as you have heard and you're well aware, IBS is considered to be a functional GI disorder, and this group of disorders have also been referred more recently as disorders of gut and brain interaction.

It describes a spectrum of GI conditions in which patients experience signs and symptoms over a chronic time course that can be unpredictable in nature with exacerbations that can be disabling to patients.

There are no known anatomical, structural,

Page 265

or biochemical abnormalities at this time, and signs and symptoms are believed to be related to abnormal intestinal motility perception and brain-gut communication.

Because there are no objective markers such as abnormal colonoscopy or endoscopy, diagnosis is made on patient-reported signs and symptoms. And as mentioned earlier this morning, the Rome criteria is typically used as the diagnostic criteria in functional GI disorders. And this is the Rome criteria. You heard about this this morning.

With the basic regulations in mind and the characteristics of IBS, we have worked with multiple stakeholders, including most of you, including patients, pharmaceutical companies, academia, and professional societies during the drug development process. Partnership among stakeholders can facilitate the drug development in a variety of ways, including assisting in the identification of clinically meaningful, measurable, and interpretable endpoints; assisting

Page 266

in identifying acceptable designs for trials that can enroll and answer key questions; and sharing a commitment to the completion of successful drug programs.

The FDA recognizes that the patient perspective is key to informing the drug development process. And as Dr. Kovacs mentioned, a public meeting on functional GI disorders relating to patient-focused drug development was held in 2015. They provided a lot of information for us.

So this leads to us talking about the fundamental regulatory aspects of the drug development process, which you are for the most part aware of. In this process, we encourage pharmaceutical companies and investigators to meet with us to discuss their proposed trial objectives, design details, endpoints, and current findings at various stages of development, including at the pre-IND or investigatory new drug period; at the end of the phase 2 period when the proof of concept and the treatment dose or doses are determined for

Page 267

future phase 3 trials; and at the pre-NDA or the new drug application stage as well.

So these meetings can help to facilitate mutual understanding, provide guidance, and enable drug development programs to gain further insight into the FDA regulations for new drug approvals. It can also potentially help to prevent harm in the drug development process and optimize the demonstration of efficacy.

So there are many challenges in IBS drug development from a regulatory perspective, and here are some select ones. As discussed in depth by my COA colleagues earlier, we rely upon patient-reported outcomes for the assessments of symptoms and disorders such as IBS, since these can represent direct measurements of treatment benefits regarding how a patient feels or functions.

For conditions like IBS and other functional GI disorders, input from patients regarding their signs and symptoms is essential, but can be challenging to conceptualize, measure, and analyze in itself.

Page 268

So one commonly-faced challenge includes differentiating concepts of abdominal pain and pain-related symptoms for efficacy endpoint definitions and analysis. And this was discussed in depth this morning, particularly the question of, are abdominal pain and abdominal discomfort describing the same symptom?

In the May 2015 public meeting on GI disorders at the FDA, a large majority of patients identified abdominal pain and discomfort as the most meaningful symptoms to them, but as separate entities, as you have heard. Abdominal pain was described as being temporary in nature, crippling at times, and that a variety of different types of pain existed, including constipation pain and intestinal spasms.

In contrast, abdominal discomfort was described as a duller sensation that was pervasive, last hours, and could be perceived as bloating, gassiness, fullness, flatulence, and the sensation of incomplete evacuation.

So how about abdominal distension and

Page 269

bloating? Are they redundant with pain or discomfort? Half of the participants during that meeting identified distension or bloating as significant as well, but they also considered it to be distinct from pain, but could be related to discomfort.

Note this information was gathered from an open meeting format and points to the importance of addressing and distinguishing what and how symptoms and signs are measured as endpoints during the clinical development process.

For the evaluation of treatment efficacy, most clinical research in IBS focuses again on abdominal pain, intensity, or severity as you have heard, as well as stool frequency and consistency. However, there are definitely additional signs and symptoms, such as those listed here, and there are more that are key to the patient's experience.

For clinical trials, it's important to understand whether the concepts that are being measured are intended for use as primary endpoints, secondary endpoints that may potentially support

Page 270

labeling, or as exploratory ones. If they're intended to support labeling, then the concepts, as you have heard, must be clearly defined and measured in a reliable and valid way for labeling purposes.

For example, although straining has been proposed as secondary endpoints in many clinical trials, it has not always been well defined. Patients have equated straining to effort, time, and pain associated with stooling and a lot of a variety of other kind of qualities. So we definitely suggest that qualitative investigations regarding the interpretation and meaningfulness of these concepts are vetted prior to phase 3 trials. We spent a lot of time talking about these concepts in our dealings.

Additional select challenges, here are some of the ones that we encounter during our evaluation of the drug development programs, and I'm going to talk about some potential solutions as well.

These include weighing the benefit versus risk of IBS therapies in the drug review process

Page 271

and continuing to monitor the risk in the postmarketing period, realizing that IBS in itself is not a deadly disorder or disease. And so we want to make sure that the benefit definitely outweighs the risk of therapy.

We emphasize it's important to leverage smaller earlier studies such as phase 2 trials to adequately explore and identify optimal drug doses and potential efficacy endpoints to be used in the phase 3 trials to help ensure program success. The establishment of clinically meaningful and acceptable improvements in signs and symptoms are essential prior to the larger studies.

Instead of comparing the average changes observed across treatment groups with the numeric differences and means, which may not be clinically meaningful on the patient level, we have suggested that investigators assess within patient clinically meaningful changes from baseline and signs and symptoms.

Accordingly, at this time point, the guidance, which I'll get to, recommends the

Page 272

conduction of responder analysis. That compares to a proportion of patients within each treatment group who meet a definition of being an overall responder, and I'll talk a little bit more about this in the future regarding the pain assessment.

We also advocate that trials are designed to target more than one IBS sign or symptom, given the proposed mechanism of the drug, since having a narrow focus can create challenges in itself in addressing the effect and outcome of other relevant signs and symptoms. For example, trials that focus on the primary endpoint on the evaluation of abdominal pain intensity by itself may not adequately assess whether abnormal defecation improves, remains unchanged, or worsens with a specific treatment.

Regarding the placebo response rate, we recognized that this could be high in multiple IBS trials, and this can represent a challenge in demonstrating that treatment effect size, particularly when it's not very large. And we know that in IBS trials that there typically isn't a

Page 273

very large effect size. Therefore, we suggest using trial designs -- perhaps this might be controversial -- with a placebo run-in period for removal of placebo responders prior to randomization.

In addition, adequate trial duration is needed to assess drug safety, efficacy, and treatment durability, particularly for therapies that are intended for chronic treatment of IBS.

So you've heard about and you've read about and rely upon the FDA guidance, and this is the guidance for industry for IBS. It's used to help address and select other challenges in drug development. It was developed in 2002, and it's based upon the Rome diagnostic criteria and published literature.

It includes acceptable and provisional endpoints in trial design recommendation for the evaluation of drugs to treat patients with IBSC, in particular, and IBSD, and provides recommendations for which trial design development can continue to evolve. So we recognize that there is still

Page 274

evolution. This is not hard in stone, necessarily, but to evolve within our regulatory framework.

These recommendations may assist in developing treatments to address the needs of patients while the work of increasing and vetting out reliable PRO instruments for FDA qualification continues.

In the guidance in general, we recommend a primary endpoint that measures the effect of the treatment of two major IBS signs and symptoms in support of the indication for the treatment of IBS.

This is the IBS pain intensity and abnormal defection. For abnormal defection, typically, trials in IBSC assess stool frequency and IBSD trials assess stool frequency as primary endpoint components. And we recommend that these components are evaluated in trials even as secondary endpoints if they are not part of the primary endpoint.

In regards to assessing a clinically meaningful change and pain intensity, the guidance recommends the incorporation of a defined responder endpoint and analysis at this time. This is

Page 275

particularly for the abdominal pain intensity responder, and it's defined as a patient who experiences a decrease in the worst abdominal pain of at least 30 percent compared to baseline in the past 24 hours.

Patient is to be categorized as an overall responder if they achieve a prespecified improvement in weekly or daily response for at least half of the weeks or days of treatment. So if they're treated for 12 weeks, then it's 6 weeks that they have to show that they were a weekly responder.

To note, regarding the 30 percent reduction in pain intensity in comparison to baseline, this was primarily based on published literature of other chronic pain conditions at the time of this publication, such as rheumatic arthritis. So it may or may not affable to patients with IBS, but it sounds like there's been some additional work that has been done and is continuing to be done.

Therefore, we do recommend conducting additional responder analysis that evaluate greater

Page 276

reductions in pain with treatment that may be more beneficial, so perhaps a 40 or 50 percent reduction in pain intensity or more, as well as evaluation of cumulative distributions of various amounts of pain reduction.

We also recommend that sponsors perform qualitative work to see if these and other thresholds can be validated in IBS populations for further use as key endpoints in analysis and trials.

Working backwards a little bit, in light of the Rome criteria and the related components of the recommended primary endpoints that you heard about, the guidance suggests the following entry criteria for patients with IBSC and IBSD, and this has been discussed.

The important thing that we note is that patients who enter the trials, we ask to have sufficient clinical manifestations of symptomatology, whether it's this symptomatology that is chosen for entry criteria or other signs and symptoms, just so that there can be a

Page 277

demonstration of clinically meaningful improvement with treatment.

These are the final thoughts. In conclusion, there are many challenges in the clinical development of IBS therapies. Presented today were just select ones. We recognize that there are definitely more. We encourage early and often collaboration in the drug development process and recognize that collaboration is key to addressing challenges.

My final pearls are that in the IBS drug development, consider leveraging phase 2 trials to optimize programs assessed by clearly defining endpoints, defining clinically meaningful treatment response, and then subsequent effect size versus the treatment arms; identify appropriate doses for phase 3 trials in IBS; and consider the placebo response rate in these trials.

This is thank you to my team leader, who is here, another medical officer, my division director, and the director of the ODED [ph], FDA.

(Applause.)

Page 278

Q&A and Panel Discussion

DR. SMITH: Now I'm going to call all of the people who were just speaking during this session since lunch to come up and sit with us up here for the discussion.

Does anyone have any questions? I see Steve in the back.

DR. BRUEHL: I've got an ignorant question referring to something that was mentioned earlier. So it's clear that some of these disorders have multiple components that are somewhat independent. So you've got -- like an IC, you might have pain as one key component, and then you've also got urgency possibly as a second component. And it's a totally different thing. And clearly, the talks we heard said it doesn't make sense to lump those into one measure necessarily.

If I understood right, I think that somebody mentioned the idea of having co-primaries, so you'd have pain and urinary symptoms as co-primaries. So I'm just wondering, pragmatically, if you do that a priori, and your trial shows efficacy on one but

Page 279

not the other, is that accepted?

I mean, can you still get the drug approved for the one that it was significant for? And then what is the labeling like on something like that? Is it just descriptive to say it may also change these other things, but we didn't show it significantly? I'm just kind of wondering how all that would work.

DR. JOHNSON: I'll start with the types of endpoints, and then I'll let my clinical colleagues tell you what they think is relevant or not. And I think for all of these, it's always a discussion that comes up.

You all are supposed to be thinking about endpoints, so part of that discussion before you let them out the door tomorrow is if you had a therapy that only changed the pain, but never changed the urgency, is that worthwhile for patients? If you had one that only changed the urgency, but didn't change the pain, is that worthwhile for patients?

I say that because that bullet under

Page 280

co-primaries describe exactly that, where you have two primaries and there are various ways that you can handle how you designate the alpha for them. There are a lot of ways you can do it.

But basically it says, if you win on either one of these, you've won. But you still are going to have to report everything. Let me be clear. But on co-primary, it's saying, when you have a co-primary and it would say I won on pain, I did not win on urgency, that means it's a no-go.

So it depends on what the patients think is really important, and what you all are thinking is important, to decide that you've actually made a step. But in writing it, we'll write what it changed, what did not change. That's how you write them out.

DR. DWORKIN: So I have a follow-up question, Dr. Johnson. The kind of classic co-primary that you just tried, there has to be significance for pain and significance for abnormal defecation or urination, otherwise nothing. Right? And hat we just saw was a kind of composite of

Page 281

those two things, pain and abnormal defecation or urination, where a patient gets classified as a responder, and that the analysis is done on that responder composite.

As a statistician, could you comment on the advantages and disadvantages of those two very different approaches, co-primary pain and abnormal defecation, urination versus this somewhat complicated composite responder?

DR. JOHNSON: So you don't necessarily have to also have a composite responder. That's a third way.

DR. DWORKIN: Exactly. So what are pros and cons?

DR. JOHNSON: So I can talk about that, but I don't know if -- do you want to talk a little bit about it first?

DR. WIEDERHORN: The co-primaries?

DR. JOHNSON: Well, about which one we're looking for.

DR. WIEDERHORN: Well, it depends on the particular indication. Like for overactive

Page 282

bladder, we have two co-primary endpoints, which are urinary frequency and incontinence. And the sponsors have to win on both for that. So it really just depends on the specific disease.

DR. DWORKIN: But my question was for, IBS, we don't have co-primary, but we still have two different things that are then combined into a --

DR. WIEDERHORN: And I think a lot of it is the frequency of each of these endpoints is brought out in the review session. And again, I'm not an expert on IBS, so I can't really answer those priorities.

DR. HANES: I could try to take a stab at that. So for IBS, we do recognize that multiple signs and symptoms are important, particularly to include in the primary endpoint. So we think that it's definitely important to look at both abdominal pain, right now its intensity, but could change to frequency or other things perhaps, but we also recognize that abnormal defecation is important as well.

So regarding whether it's a component

Page 283

endpoint, which we kind of call it at this point -- and correct me if I'm wrong, Dr. Johnson, but as component versus a co-primary, it basically depends on the statistical analysis plan and how the alpha is divvied up.

So the question bounces back to Dr. Johnson. But right now, it depends on how it's presented in the analysis plan and what people want to spend their alpha on and what they want to take a risk on in terms of not working out to be in the label.

DR. JOHNSON: So now I will respond. But I think the problem with many ways that people will do the composite, I think when they see the responder definition in a lot of our guidances, they think they have to do a responder analysis. That is not necessarily the case.

Now, some of our guidances do have responder analyses. Also, a lot of the clinical guidances haven't necessarily gone through the statistics office, and we've now changed that process so that they are. But there is some balance and some changing and evolution that has been happening over

Page 284

the years, so that's another thing that I will put out there. But typically, when you are trying to do a composite, it tends to put you from a research standpoint at a disadvantage.

So composites in the fields that we've been talking about today very rarely tend to do well. And I say that because from a mathematical standpoint, a lot of things go wrong. And many times, we're not weighting those well. We end up defaulting to measures that if instead what had happened is that people looked at them individually in their continuous state or their ordinal state, and had done those analyses, and actually figured out how they wanted to set up their testing hierarchy -- and there's some very innovative ways of doing that -- you would have been a lot better off.

But being blunt, I think people say, okay, composite, I've got one test, this solves my problem, because they're not willing to get into the innovative part. The math is not that hard. I think a lot of people have done a lot of work to

Page 285

make it such that the math is easier. But really, the composites for the types of areas you're talking about rarely are how you want to go.

That said, that doesn't mean you don't need to address each of the issues inside of what many people are talking about as composites, but instead just do it in a different way as a multiple endpoint.

That's one reason this guidance we thought was so important. And it took us close to seven years to get this guidance out, to get it through the political type of clearance process because people get scared when they see statistical guidances.

You'll notice that there's some math in the end; there's some Greek symbols back there. And that's actually what we came down to, is actually trying to make it fairly statistical. But the idea there was to say there are many ways to envision this.

But we have really in-depth discussions internally and also with our sponsors to say what

Page 286

is it that you think this substance is going to do? What will it change? And then we have to then think what is the indication.

If you have a disease -- I have one where the fundamental aspects, one was a psychological trait and one was a physical trait. But you did not change both of them, you did not fundamentally do anything about the disease.

So we said, listen, if you are going to study people with this disease and you are saying that you want this drug to be helping the signs and symptoms of this disease, you must impact both. That's a co-primary.

But again, if you have evidence to say, I'm only going to hit one, and people say that it's really important if they can do this; and that's okay if the other one doesn't change or even if I get a little worse, again, you've got to define it, you're going to have to measure everything, but you're only going to allocate your alpha to one of them. Or if you think it might be both, then you have ways of recycling.

Page 287

DR. HERTZ: Sharon Hertz here. I think going back to the original question, why don't you know by the time you're going into phase 3 what your drug is going to do? Not whether or not it's going to be capable of being successful, but why -- to me, there's a fundamental problem, if you're getting to phase 3 and you're not really sure what you're capable of moving, you probably skipped phase 2, which pretty much doesn't exist anymore for I'm sure of financial reasons.

So the idea of going into a phase 3 with a, gee, I'm not sure. I'm not going to have two independent variables and I want either, then how do I label that?

In a population of 100 people, 25 had improvement in pain, but a different 25 had improvement with frequency of a symptom. Well, if it's the same symptom and it's become less severe and/or less frequent, so for instance, trigeminal neuralgia, that could be okay. But if it's two different things, then is it just really a meaningful assessment of the drug, and then you

Page 288

have to start looking at all the details. Did you just miss in some people? Was it just a criterion thing?

So a lot of that needs to be kind of worked out in exploratory mode. Is your dosing a little off? Your entry criteria, are they a little off? Where is the problem?

Ideally, if you do, if there is early work, which doesn't have to worry about putting all your eggs in a basket, prespecifying, and making a guess at the best stab at it, then phase 3 is most likely going to be more -- let me just say, less worrisome. It's less of a gamble, less nail-biting at the end that you're going to have stuff and then you're going to have to figure out whether or not it meets a regulatory standard.

DR. JOHNSON: I do want to emphasize phase 3 was supposed to be confirmatory. And especially the dosing part, a lot of times in the early-phase studies, we see very little work done across multiple doses to get enough information.

Even when you're doing it, you can put in

Page 289

some of these endpoints to get an idea of where you're supposed to be going. But we do, under PDUFA VI, if it's past, have an extra meeting. We have two pilot programs in there. One of them is for innovative studies designs. And while a lot of that is to deal with a variety of other study designs, we will have information.

This is already in the public knowledge. Starting in FY18, we'll be publishing information in the Federal Register to talk about how to ask for meetings to be part of the pilot program. And an important part of that pilot is knowledge of your study design will be made public prior to your finishing any work or going on the market. So it's kind of a tradeoff. You get extra meetings with the FDA, but we'll be doing discussions beforehand.

But I say that because the heterogeneity -- when there are only 24 people in the world, it's very different than when you have many millions of people in the world. So we can borrow a lot from the heterogeneity information we've learned in the study designs and rare

Page 290

diseases, but not too much. And it is hard. When I come down to it -- and Dr. Hertz said it very well -- what does she write in the label, and work your way out from there.

DR. SMITH: Lee?

DR. SIMON: Lee Simon. That's me. The question I wondered about is we've heard multiple times today reference to worst pain, and the worst pain in the last 24 hours, the average worst pain in the last week, or whatever.

Two years ago, we had a discussion at such a meeting about worst pain. And I work with a group of people that do outcomes in musculoskeletal pain, very different than what we've been talking about today. And when we did cognitive discussions with patients about what worst pain means, we could not get any consistency of understanding about what that meant.

So I was wondering if there's actually been work done that actually led to many people at the FDA talking about worst pain in these two indications. And I wondered whether or not there

Page 291

was actually support for using that concept from the context of patients.

Do they understand what it means to identify the worst pain as opposed to whatever pain measure we're going to do, but do they understand worst pain?

DR. KOVACS: I think the context of use in a patient population is the most important, so getting the qualitative research done with those patients that are your target for your clinical trial. So maybe worst pain is most important to them or average pain, but the most important thing is just having the consistency across patients, and in the item, and in the instructions.

So sometimes, we see a 0 to 10 point scale, numeric rating scale, where it'll say please average your pain across the past 24 hours, and then a zero rating, verbal anchor descriptor as no pain versus worst pain, worst imaginable pain. So then you have kind of conflicting information that patients are getting. So you're averaging your pain across 24 hours, but then the response option

Page 292

scale is asking about your worst imaginable pain. So we've seen that.

I think that the most important thing is just having the consistency and what most of the patients, I don't know, 60, 70 percent of the patients, are saying is applicable to them and what they interpret as the most important measure of pain for them

DR. JOHNSON: Yes. And sometimes when you see flares or in the migraine or headaches, we do hear from patients -- when you read through the qualitative work, part of what they're interested in is if they can take something that's really bad -- so if they have some sufferable pain and they can work through it and they can handle that. But it's taking the highs and diminishing the highs to something that they can actually tolerate better, that's a win for them.

So there are some that really do distinguish that worst, and it seems like they can -- but I agree with what you just said. It's very kind of what's the patient group, what's that population

Page 293

you're trying to work in.
There are some things that I thought nobody would ever understand, but those patient groups know it, because they live their disease every day and they've heard enough of that vocabulary that it makes sense for them. But that may not translate to another patient group, and that's something to really think about.
DR. HERTZ: Lee, what did they understand? Did they understand average pain?
DR. SIMON: So it was interesting. So what I've referred to before was the OMERACT process, and we have 52 working groups in outcome measures of rheumatic diseases. And a significant portion of what we do has to do either referentially to pain or actually specifically to pain.
Patients do understand, when you ask the question, how much pain do you have right now, which some of us like, some of us don't like, and then some people are pushing, how much average pain have you had over the day, 24 hours, and then you average that over a 7-day period. That seems to

Page 294

also be understood. But again, this is musculoskeletal pain, and it's not as complex in certain ways as all the other things that are going into these particular syndromes.
So I think that it's really important to be consistent, as we've mentioned, in that population that's being studied. I still think that patients in musculoskeletal pain are most consistent in understanding how much pain do you have right now. And if you want to do that over a 7-day period and average that over that 7-day period, that's possible. But consistently, they tell us that's the one they really understand.
DR. SMITH: John?
DR. FARRAR: I want to ask a couple questions, but a quick comment about that. The problem with right now, especially in musculoskeletal pain, is that the patient has ridden a bus and walked up five flights of stairs. And so their pain right now may not reflect their normal pain during the process.
We talked earlier about the fact that

Page 295

between patients, they don't always understand the question about pain in exactly the same way. But as long as they're consistent about it over the course of the trial, you'll get a reasonable result. So we could talk all day about that or I'd be happy to.
But my question really revolves around this issue of multiple endpoints. And whenever we develop a patient-reported outcome that has more than one question in it, it could be considered multiple outcomes. Right? So the SF-36 is -- well, actually, it's more, but at least 36 separate outcomes, which then get coalesced into a bunch of different subscales.
So there are two aspects to the question, one of which -- sorry. And then in thinking about the definition of the chronic prostatitis or the IC, the definition is constituted of pain and urinary symptoms. That's the combination.
In the MAPP program, in looking at that, there's a nicely published paper that can't remember whether it was Mike or Henry talked about,

Page 296

which said you really can't count on those going the same way with the treatment, so that you might want to consistently treat them separately. It seems to me that what we're talking about is being consistent with the biology and what your drug is thought to do, and then thinking about that.
But I wondered about your thoughts about patient-reported outcome scales. If the GUPI includes things in both pain and urinary symptoms and it's a validated scale, I guess it could be an outcome. But it is combining two things, and I wonder about your thoughts, Laura.
DR. JOHNSON: You can have plenty of different domains inside the same scale, and then you just have a difference. Like, SF-36 doesn't have a total score. SF-36 has 8 different little scores, or you could break it into physical component and mental component, or you could break out the PF-10. But there is no total score there.
I think what's important is, for that tool -- I'm not going to talk without having done the thorough review, et cetera, about how good it

Page 297

is. But let's say it's there, and you have a urinary component, and that pulls together on its own, and you have that pain component, and it pulls together on its own, two separate scores, fine.

It might be that one of those scores, even though you're giving the entire instrument -- so if I'm a patient, I'm just answering different items, right, I'm just going through it. But when you're scoring it, when you're writing the hypothesis and you have a hypothesis-tested analysis, it's just on one of those domains. And it might be that you do a co-primary because the decision is both need to move, or it might be that, really, only one of them is your primary endpoint and another one is a secondary or something like that, like whatever you've negotiated with.

But yes, I mean, it's not a problem. And I think for a lot of these PROs, the mistake that we see frequently is people try to get a total score when there shouldn't be one.

DR. FARRAR: There shouldn't be one.

DR. JOHNSON: But at the same time, then

Page 298

other people go through one, and go item by item by item, and break apart something that really was supposed to be a whole.

So you can make a mistake going both ways, but it really depends on how you develop the tool. And when you go through and do that early quantitative work in the tool development, what that evidence is showing you. So you always have an idea of how you think it's going to go, but I've done tool development, and sometimes you're like, oh, these things are hanging together differently. You get more evidence? Yes, that's the way it's supposed to go. That's why it's science and why we try to learn things.

DR. FARRAR: So it's really taking things and dividing them because obviously even the urinary scale consists of pressure on filling issues related to frequency, issues related to IBS, constipation, diarrhea, fullness. And how you group those, I think what you saying is it's a scientific question that needs to be answered before you bring it to the phase 3 meeting.

Page 299

But no matter what you do, you're going to be combining some set of symptoms, and I think the issue would be -- my advice, at least -- and I'm interested in your thoughts on whether it makes sense -- is that it ought to be scientifically based and demonstrated as being a thing that consistently changes together as opposed to being separate.

DR. JOHNSON: Yes. There should be evidence for that. Otherwise, it's also going to be very hard for your biostatistician to help you design your trial.

DR. SMITH: Hanna?

DR. GROL-PROKOPCZYK: So this question may be as much for Bob and Dennis as --

DR. TURK: Who are you?

DR. GROL-PROKOPCZYK: I'm Hanna Grol-Prokopczyk at the University of Buffalo -- as much for the two of you as for the panelists. But one thing that would help me figure out which of these issues that are coming up that are most salient to the work of this group is to have a clearer sense

Page 300

of what the end product at the end of tomorrow is supposed to be.

So I'm not clear right now whether the goal is to recommend pain outcomes period, or are we going to go into how you measure constipation, diarrhea, and then urinary frequency, and things like that.

Secondly, relatedly, is the goal to sort of focus on outcomes for the various conditions independently and possibly with non-overlapping pain measures, or is there also an effort to try to have some comparability in the pain measures that are used for the various conditions?

DR. DWORKIN: So Hanna, you'll have to take my word for it, but I started to write out an outline for tomorrow afternoon's discussion.

DR. GROL-PROKOPCZYK: I'm getting ahead of myself.

DR. DWORKIN: Your question is on numbers 1 and 2 on my outline. So you're thinking 24 hours ahead, but those are exactly the questions we'll start off with at 1:00 tomorrow afternoon. It's

Page 301

really a group discussion. Are we just going to recommend pain outcomes or pain plus abnormal urination, defecation outcomes?

Are we just going to focus on primary endpoints or also talk about secondary and exploratory outcomes? But that, we'll all decide as a group tomorrow at 1:00.

DR. SMITH: Quentin?

DR. CLEMENS: It's Quentin Clemens. Regarding the primary outcome or composite outcomes concept, I think what the IBS field has done nicely is have defined these subgroups. So conceptually, we could do the same in IC, where we have a pain predominant group, in which case, then, the pain, whether it's endoscopy or some other measure, it could be the outcome. And then we could have a urinary predominant group with a certain severity of urinary symptoms, and then the urinary outcomes we know. And that could kind of be a parallel thing.

What I want to know from the IBS people is what do you do when they have both? Because I had

Page 302

a conversation at lunch, and I think they said, "Yeah, that mixed group is a mess." But we went into that a lot in the course of urinary symptoms. So sometimes when you sub categories like this, the patients don't cooperate and you get there in between.

So how do you all handle that when you're recruiting patients and conducting these studies, where they have both C and D, or a different --

DR. HANES: That's a great question. So typically, there is a spectrum of disease of IBSC/D, and you heard about the mixture. At this point, we haven't seen a lot of studies that looked at IBS or IBS with mixture at this point. Usually, it's typically IBSC or IBSD. And we recognize that on patients with IBSC who are more along the IBSC spectrum are different than IBSD because of their stool characteristics, and what you've heard also about their types of pain, perhaps pain frequency, perhaps pain severity.

When the drugs are being developed typically, they are targeted either for IBSC or

Page 303

IBSD. So pharmaceutical companies, sponsors have come to us particularly for a drug that's for helping the abdominal pain as well as a stool component of constipation or a stool component of diarrhea based on the mechanism of action.

I haven't heard of particular drugs that have come in to look at treating both mechanisms, helping constipation and diarrhea at the same time. But perhaps in the future, there will be something, but right now, typically, it's IBSC or IBSD in the mechanism of action.

DR. CLEMENS: Are most of the IBS patients in the mixed group? I guess that's just a question? Or I mean, is it a very small group that's both? In other words, are these studies kind of ignoring the majority? I suspect not.

But then the other question to follow up is, if you have an IBSC drug and you have a mixed patient, then you just put them in the IBSC, study it, and ignore their D symptoms, and then when a D study comes along, operationally, is that kind of how it tends to work?

Page 304

DR. HANES: That's a good question. The question is about the prevalence of IBS mixed group versus IBSC and IBSD. I don't have that exact answer. I wish Dr. Chey was here. Perhaps my team knows more about the prevalence of the variety of different subtypes.

But what I would say is that, at least in looking at the trials that are proposed, there are entry criteria that further delineate. This was just kind of a brief overview, but there are definitely entry criteria that clearly delineate who's excluded.

So for IBSC trials, those who have a predominant diarrhea or more than, say, a number of stools and diarrhea are excluded. So we try to get a more homogeneous population and not completely heterogeneous with a lot of mixture in it. And on the other hand, the same thing for diarrhea, so there are exclusions in terms of how much constipation there is that wasn't presented today.

So the goal is to really target what the mechanism of the drug can do. So if it's intended

Page 305

to treat diarrhea and pain, then we really want a patient population that has diarrhea and not too much mixture. I think that programs are doing a good job in providing this.

DR. LEMBO: Maybe I could just comment on those questions. So the traditional literature suggests that it's about a third, a third, a third. More recent literature suggests that many of these mixed are drug induced because patients are searching and desperately trying, and they don't get it quite right with the Imodium or some of the laxatives.

Some of this is also induced by the fact of the definitions that we have for Rome, which is not that reliable, because it relies on historical recall from the patients. Even if a physician is reading it, it can be interpreted differently.

The Bristol stool scales, which you used for stool consistency, is notoriously difficult for patients to get that right, because they always say the same thing, "What do you mean? At the start of my bowel movement? At the end of my bowel

Page 306

movement?" And traditionally we have not been able to really guide them. And as some of us have said many times, as long as it's consistent throughout -- and that's not a pun on words -- you do it the same way each time.

That being said, we've seen this several times with the rifaximin trial, where we tried to include mixed desperately because we felt that it was mainly a drug for bloating and not necessary for bowel function. And we're surprised to find so few people that actually fit into the category. When you did the baseline -- I mean, by history, sure, but when you go to the baseline, you take them off their drugs, we found that almost all of them fit into the definition of IBSD.

The other part is, of course, with these endpoints, it's hard to study these patients because you have to pigeonhole them into the CRD, and that's become a bit of an issue as well, although I know I understand that we can do it for pain, but it's a little bit of a problem.

DR. SMITH: Ursula?

Page 307

DR. WESSELMAN: Ursula Wesselmann, University of Alabama at Birmingham. Now, we discussed comorbidities this morning, and I would like to ask the panel how you suggest we deal with them for a clinical trial.

So as we already said, a patient might qualify for the IBS trial, but the patient might also actually qualify to be enrolled in the fibromyalgia trial. And in our own studies, we have sometimes asked patients who had multiple pain comorbidities which one is the one bothering you most, and then to kind of go down the list.

But it's often that several chronic pain syndromes bother them the same way. And the way it develops is usually that they start out with one pain syndrome and then develop multiple others. So it would be really key to prevent that situation, where they have multiple pain syndromes.

Actually, we just published this month a paper where we looked at patients with visceral pain comorbidities and fibromyalgia, and we could show in a clinical sample that if you treat one

Page 308

pain syndrome, then the other one gets better as well. So if you treat fibromyalgia, the IBS or the endometriosis associated, pain gets better and vice versa. And there were a few publications in the literature like that that are also comparing, looking at migraine headaches and fibromyalgia.

So should those be secondary endpoints, or what would be the best way to recognize these multiple comorbidities that you especially see in visceral pain?

DR. WIEDERHORN: Listening to your question, the key thing would be what is the demographics of the population in the clinical trial and what is the epidemiology of the disease, because if you wanted to use it as a secondary endpoint, you'd then have to make sure that you have the proper number, proportion of patients in there to be able to do some kind of stratified or prespecified analysis.

So we would want that, and sometimes we do insist that we have a representative population just to do that. I can't give you a specific

Page 309

example.
Jordan, do you have any?
DR. DIMITRAKOFF: I don't have a specific example. Thank you.
Thank you, Dr. Wesselmann. I think it's a great question because I think, historically, in most of the trials that I am familiar with. I think patients, especially in the process of the CPPS field, patients with severe IBS or severe chronic fatigue syndrome have been excluded.
I think this is an arbitrary definition, and in light of what we have been discussing in your work and the other work in the field, it's obvious that this is probably not a representative population. Again, it depends on the definition of severe, like how severe should it be to exclude it.
So obviously, I cannot speak for the agency or for the division of what might represent a valid secondary endpoint. But I think the reality is that these patients, as you mentioned, have multiple comorbidities. And if you treat one, another one gets better.

Page 310

The other reality, I think listening to everything in the morning, is that we already know we actually classify those patients arbitrarily with a localized or systemic disease. And then one day -- it depends on maybe the day or the time of their life, and how do we know that patients that present, were localized as you mentioned, wouldn't become patients with systemic or other pain conditions.
I think we probably need to go into this whole big data thing, and the things that we're learning about people now that would probably change the way we phenotype patients right.
Sorry. I'm not probably --
DR. JOHNSON: I was going to try to give you about four different answers.
(Laughter.)
DR. WESSELMAN: If I could just comment -- if I can just give one comment in that regard. Like when we started out, for example, 20 years ago with the IC studies, the idea was inclusion criteria were that patients who had other

Page 311

pain comorbidities were actually excluded, only to find out that there's hardly any patients like that. So as you dig into this in more detail and ask the questions, you find out that those patients really exist.
An exception to that is probably vulvodynia, the localized vulvodynia, the vulva, previously called vulvar vestibulitis that Andrea mentioned in her lecture. About 50 percent of those patients indeed do not seem to have any other pain comorbidities, and they don't seem to develop them later in life, either.
DR. JOHNSON: Yes. So personally I find that, from a generalizability standpoint, problematic, but that's what we regularly see, is that people want a clean sample, so they start excluding.
To the extent one of my friends at Cleveland Clinic said to me one day, she's like, "I work at one of those rarified academic institutions, and even my patient population does not look like anything that's ever in a clinical trial." She's

Page 312

like, "Yeah, I feel like I just need to do my own research on my patients once something comes out."
But we do see stratification. So if you know, if you know that there is a comorbidity, we will sometimes preplan those subgroup analyses, and we may even stratify randomization on if they do or do not have another comorbidity or a specific one.
So sometimes, it's constellation of diseases. So it might be that you stratify that they have other pain-related disorders, which means you have to check for it and you have to be able to define it, and sometime groups don't want to do that and some do. But then you can preplan those types of analyses, and you can also power your study to be able to look at different results there.
We do sometimes see groups that will put in endpoints while generally the general rule is you measure everything and everyone every time. But sometimes there may be a specific measure that you're only going to use in that subgroup because they are the only ones for whom it is reasonable to

Page 313

do it. But it depends.
Actually, you reminded me of an NIH study I was involved in where the center -- it was actually a pain center. I want to say we did it at Stanford with Sean Mackey. But they had four different studies going on, so we actually did a biased coin allocation.
So what happened is when patients called in and said that they were interested, we had all the inclusion/exclusion criteria there for all of the studies. So the person was ticking off, and at the end of the interview -- because many of these overlapped, right? So at the end of the interview, it popped up on the screen to say which studies the person was eligible for.
So then they described those studies to the patient and found out which ones they were at least interested in, and then did an allocation. And they weighted it based on if we had different recruitment targets and also so that there was some sense of chance going on. But that's one way, that if people are eligible and have multiple

Page 314

conditions, and it's not a problem to do it.
But again, that was an NIH type of study, and we at the FDA do not control necessarily what comes into us, at least to a very little extent. But I think there is a lot more openness than is necessarily seen because we approve or don't approve what has come into us.
MS. VEASLEY: Just a quick comment. It would be really helpful to have FDA guidance on that topic because, again, there's a huge literature base showing that the more sites of pain you have, the less likely you are to respond to any therapy.
We're working with three companies right now, one on endometriosis, one on IC, and one on low back pain. All three are enrolling patients that have other pain disorders, and none of them are tracking them at all.
So when you talk to companies about this and even, like Sean has a colleague -- he mentioned that last week Dr. Collins had a meeting on developing a public-private partnership with

Page 315

industry and the FDA to develop pain therapeutics. He said 9 out of 10 patients that he treats looks nothing like the one that he enrolls in a clinical trial.
So I think, again, this is the middle-of-the-road approach between rigorous science on the end of doing a placebo-controlled randomized trial, but then also on the other end of, what do our patients actually look like.
Then just a quick question. Not to make something that's already complicated even more complicated, but we're talking about, I think as a speaker mentioned earlier, for every 2 patients she asked, one side worst pain was more important, one said average pain is most important.
All patients are individuals, and all individuals have different preferences. And I appreciate very much the FDA's move and other agencies' move towards patient-focused drug development and patient centricity. But I think sometimes we've gone too far to think that if we've talked to 5 patients, 10 patients, or even

Page 316

100 patients with any one given disorder, especially something so individualized and personalized like pain, that we understand it, and that's just not the case.
There's very big differences between people who have had pain syndromes for 10 to 20 years versus someone who just started having IBS 6 months ago versus somebody who may be in the middle of that spectrum.
My question is, are there other disease areas where patient preferences are being taken into account in more of a sophisticated design? So for example, we're talking about this issue between pain and urgency. So for the patient that has urgency, and that's the primary thing that's most important to him or her, are there approaches that are letting patients identify what their patient preference or most important area is and then tracking them? And if no, how far away from that type of a scientific design do you think we are?
DR. TURK: Who are you? This is being transcribed, so please say who you are.

Page 317

MS. VEASLEY: Sorry. Chris Veasley.
DR. JOHNSON: So at least in some of our rare disease inborn error areas, yes, we do have some trials that are set up that way. They are not frequent, and I will say from an analysis standpoint we're not 100 percent sure that we're going to do the best job.
But it is a new territory and something that I keep encouraging, and I'll be encouraging at the joint statistical meetings again this year. But I think from an analysis standpoint and the interpretation at the end, we need to learn and do more because if everybody chooses a different endpoint, what exactly are they writing into labeling? I think this is something that we do have to figure out.
One other comment you mentioned, Karen Cook and I think maybe Dagmar Ottmann also worked on this. They had a discussion with MS patients, and it was actually about pain. And they were trying to get what are the different questions.
They had shown them this promised set of

Page 318

items and said what's missing? But then they went back and said, okay, does this pretty much capture it? Because a lot of times, we' look at what's missing, but then if you actually ask the patients if it's good enough, a lot of times, they'll also tell you yes.
So I think that's something else that we sometimes need to think about, is that maybe we are covering it well enough. But I do agree -- and we actually have a lot of discussion inside my office about how people, and their perceptions, and what they want change a lot over time and over the course of your illness, and what's important to you changes over the course of your illness.
So those are things that, again -- I think actually it was Lee Simon who asked me the definition of an estimate, if I could say it quickly to him. And Lesley was starting, and so I said hey, wait.
So this might actually be a place to tell you that's really this kind of this what are you supposed to estimate. But part of it is that

Page 319

population. What's the target of the research? One thing we have to be careful about without slicing too thinly everything is thinking about for that target of the research, what is it that they care about, what are the variables, what are those endpoints that we are measuring, what's the scientific question for them, because it may be pretty different for different groups and different places.
But yes, some people aren't going to move, and that's something we also have to consider. The problem is, most companies, they want people who are going to move, and they will very carefully -- and actually a lot of my NIH investigators used to do the same thing, so I can't say it's only them.
But they design a study very carefully and think very carefully about who's going to be in it because they want to see changes. And that doesn't necessarily mean that in the whole global group of people that would then take that therapy, that you're going to see it -- it's a big difference

Page 320

between efficacy and effectiveness when you think about research. That's why PCORI got their money in many ways.
DR. DIMITRAKOFF: I want to go back, if I may, to a point that you make, and I think is an important point, about the importance of how the patient feels. And I think that ties into some of the research about the duration of the flares and how often people have flares.
So I think it's important when you collect all this information about PROs and different outcomes to be informed if the patient is having a bad day or a good day. If you're doing a study, it's important to be able to capture all that information. I think it's possible to do that nowadays with all the mobile technologies and electronic diaries.
But something else I think that you said is very important is that we need to know -- I think historically people have always tried to look at people with newer diagnoses versus people who have had the disease for a very long time. For example,

Page 321

I don't mean to open a can of worms here, but talking about CPPS, the field that I know best, and IC, which I know a little bit, there is, for example -- again, it's not evidence based, but there is some clinical mythology about whether IC exists in children or not.

So if you actually want to find children that have IC, the idea is that if you actually diagnose IC or CPPS in children, maybe you can prevent them from developing the disease later on in time.

So there is a disease, for example, in children, which is called benign daytime frequency, where children start urinating every 15 minutes. It happens only during the day. It doesn't happen during the night. And a lot of pediatric urologists actually believe it's a precursor to IC.

So again, going back to the criteria that Roger was talking about, for example, the NIDDK criteria state that you have to be 18 or older to be diagnosed with IC. So if you think about the populations that we should be studying, the

Page 322

question is can we actually capture specific populations which might be suitable to study both the natural history of the disease and maybe for a specific intervention. I think it's an important point to think about some of those issues when you're designing a trial.

DR. WIEDERHORN: Another point that you brought up was the difference between or relationship to pain versus urgency. Urgency, we have no really good definition for in terms of measuring. Pain may be easier. A lot of patients, however, will say that the severe urge to urinate is pain and vice versa.

That can confound things, and that's why I think a lot of people will look at pain primarily because they think they can measure that better because we don't have an acceptable measure for urgency. And I think that needs to be further worked out because you're right. It's a very difficult distinction for some patients to make, especially when they're interviewed about it. Some patients will say, "I only decide to urinate when I

Page 323

have pain."

DR. JOHNSON: I also want to comment no something from this morning. A lot of people are talking about correlation. I hate correlation with a blinding passion. But one part of this is you do have things that may be associated, that may be related, but when you're thinking about what should be endpoints, if one thing a hundred percent for everybody at all severities ties to another, you only need one of them.

But sometimes things may be associated with each other, but especially on the edges, so for your lower severity, or higher severity groups, or stuff like that, there will be some disagreement. And that's really important. That's also what gets lost when you look at correlation.

But it's okay if you have related endpoints, and we can handle that. That's not a problem.

DR. DIMITRAKOFF: There's a good paper on the correlation of consumption of chocolate and intelligence in the New England Journal of Medicine, probably your least favorite paper.

Page 324

DR. JOHNSON: Storks and babies.

DR. DIMITRAKOFF: Yes.

DR. TU: Frank Tu. I was wondering, would it be possible for you guys to comment on an actual trial that's ongoing, that's on clinicaltrials.gov? I'm just kind of curious about some of the stuff we're discussing, whether it would apply to an actual real-world example if we were then going to try to apply that to guidelines.

Full disclosure, I've done work on an AbbVie ongoing trial that is currently in phase 3 that's listed here for a drug called Elagolix, and it runs into the exact same issues you're bringing up.

And I've looked this over, and I'm pretty sure that anything I'm saying here is entirely in the public domain because I'm on clinicaltrials.gov, so I don't have any sort of disclosures I have to put on this. But one thing we're kind of curious about is this discussion about the applicability of a trial design to a broader population. It comes up on the Alligolex trial because -- this is where I'd be interested in

Page 325

your guidance.
How do they decide on the exclusion criteria? Is that a back and forth with regulators? Because they pick an extraordinarily narrow sliver of patients to study within a chronic -- they study specifically endometriosis-associated pelvic pain, but their exclusion criteria pretty much takes out anything that is chronic pelvic pain that's not caused by endometriosis. And Katy and I were both talking to each other, saying we don't even know how you would achieve that criteria, but that's the actual -- that's the criteria they settle on for this phase 3 trial.
But it results in a trial that we think -- they've got a New England Journal article out on the initial results, but the applicability seems like almost no one. And I was wondering --
DR. VINCENT: The general gynecologist recruited to that, not even people who were very specialized pelvic pain, who would be able to detect a pelvic-floor component, for example.

Page 326

DR. TU: The question is how does the FDA work with a company on trying to decide on ultimately designing a study that would promote an approved drug that would have meaningful utility? And I can say that I understand there's a back and forth because when I was involved in the trial, they did say there was back and forth in deciding how you create these inclusion/exclusion criteria.
But I was just rereviewing it, and then several people brought up this very point about the generalizability of these studies.
DR. HERTZ: This is Sharon Hertz. So it's interesting to contrast that with what Christin was mentioning, where there was really not much of that kind of specificity going on and not even necessarily particularly good description of who's in the study. One might almost argue that that could better provide some type of real-world broader efficacy assessment.
So it goes like this. Sponsors come in, and sometimes they've spoken with you all, and sometimes it's clear they haven't. Sometimes we

Page 327

get very narrow definitions of a population because early work suggests there's efficacy there. And there may not be much information available on a broader segment of a given indication or diagnosis.
So how one weighs the value of an approval for something narrow or pushing for something broad is often something that we don't do. We let the company decide because if I push that company to do a broad population and it fails, then what? Then it means I pushed them to study the wrong population or this drug is no good. It could be either. If they do a narrow population, it works, prescribers don't know how it will function more broadly.
So I tend not to be the one to make that decision. We tend not to be the ones to make that decision. I mean, I think there's some situations in which the definitions are more advanced and broadly accepted in terms of populations. But for the most part, especially within pain, it's pretty wide open.
So now I'm going to ask you, what's better

Page 328

for the public, for the public health, for the population, to have a broad study in sort of a mish-mosh of people and you're not 100 percent sure how to predict who it's going to win in, but it certainly wins in some or to have a narrow population of very well-defined patients that you have a very good understanding of the proportion that will respond, but you don't know the generalizability?
So that's one question. And I'll say that this is a question to consider in the context of indications where there's just not a lot to begin with. So it's not like we have a ton of things where we can make clear-cut decisions.
So I'll turn that back over. I mean, especially in these areas, where I think there's a tremendous amount of unmet need, what is more useful to the community?
DR. TU: My own thoughts on this -- again, this s Frank Tu -- is that, obviously, it's a double-edged sword. What happens in clinical practice is you apply a drug that's been studied in

Page 329

a very narrow sliver of patients to a broad, oftentimes highly comorbid population that has not been allowed entry in the trial. And all manner of side effects erupt that were never seen in any of the initial studies.

Their prior drug, the GnRH agonist that is an injectable Lupron, has these problems, massive weight gain in some patients. They're not described in the initial trials, but those patients who have been vulnerable, things were never allowed in the initial trials.

On the flip side, the drug was used in conditions like bladder pain syndrome and IBS, and a subset of those patients got better, but they were never allowed into the initial trials, either. So from a public health perspective, it's a question of would you rather have more people with horrific morbid obesity, a small sliver of people, or a small sliver of people that benefit from bowel and bladder, unexpected benefits of this drug? I don't know where the sweet spot is on that, but both things actually happen in real practice.

Page 330

DR. HERTZ: But let's take out the nasty really bad side effect, which is going to be true even if it was a broad population because the number of people that are often studied in the context of development is somewhat limited.

So you may not pick that up anyway, even if the comorbidities that are more susceptible were represented. Right? Because maybe there was one case, and that was going to be written off as not known whether it was DOOR related.

So the toxicity piece aside, which is not to say it's not important, but just for the purposes of understanding the efficacy implications, I don't think there's a right answer. This is a preference question. And one could ask patients this, but I think this is very difficult to operationalize.

But what is better in the context of an area of great unmet need; specificity that you're not sure how will generalize or some type of effect and you're not sure how to predict who that will be?

DR. MARKMAN: My two cents -- John Markman, Rochester, New York -- I think that the problem

Page 331

with the mish-mosh is I've often felt like we're condemning ourselves to really trivial effects in terms of analgesic benefit.

As a clinician, I find it incredibly helpful to see a large magnitude effect, and then I know the risks of applying it more broadly, whether that's the use of oxcarbazepine and trigem neuralgia or Toradol in renal colic. I just feel like seeing those huge facts actually really helps inform my decision-making.

So when I see this very tiny separation out of 12 weeks in this very heterogeneous population, I just frankly often struggle with how to match that up with the other 400 options which have trivial effects.

DR. ALTEPETER: Hi. I'm Tara Altepeter. I also work in the GI division at FDA. And I guess I would just add to what was already said, that while it's not necessarily our decision, the population that's enrolled, we try to be as descriptive as possible on the labeling to share that information with clinicians and to allow people to draw their

Page 332

own conclusions and use that information in a way that's most appropriate for you.

So I guess I sort of see that as the balance point. And again, all of us sharing a goal of wanting to help promote the success of programs that do have an effect. And I think the more heterogeneous the population is, the less likely you are to be able to see an effect if it's only going to be working in a particular subset.

So it is a balancing act, but I think, as we move forward every year in our approach to the labeling and information that we're sharing with prescribers is continuing to change, you'll notice, if you look in that section 14 of the labels, compared to what was in there from a drug that was approved 10 years ago, you're going to see a lot more detail.

We're really trying to describe who was that population, what were the different components of the effect; not just, yes, you won, but some of those other components that might be of interest, but may or may not have been part of the primary

Page 333

endpoint, to give as month descriptive information about what we saw demonstrated in the trial and then allowing the market and the clinician's experience as they start to use the drug to then help guide them to whether or not it's more broadly applicable.

But we have seen examples where programs have failed because there was too much heterogeneity, and then we couldn't figure out -- even looking back and doing all these careful post hoc analysis, we couldn't quite point our finger to which one of those 17 confounding things we think might have resulted in a failed trial, even though when everyone really thought that there's a good bio plausibility for that drug being effective.

DR. HANES: I was just going to add, I think that I completely agree with both points, and your question and your mentioning of safety issues. Just to let you know, the drug development process doesn't end at the phase 3 level. It doesn't end at the approval or the non-approval of the drug,

Page 334

that there's definitely a post-approval period. There's phase 4. There are other studies that can be done that are included in the PMRs, or PMRCs, or looking at safety and looking at other endpoints perhaps, or subpopulations of patients that might be critical that are identified later on.

So things continue to evolve. The label continues to evolve as well. And so particularly maybe something that's a salient safety issue could be included to the label at a later point if it's seen in a broader population, because we recognize that there are limitations to the studies. We're not looking at thousands and thousands of patients in the populations, so you may not pick up on those rare severe issues, but they could be included later on.

DR. LEMBO: I think one of the cases that brings this up -- I think the point you're trying to make is the alosetron trials had an entry criteria for pretty severe significant diarrhea, and when it got to the general population of the more milder case, obviously patients got quite

Page 335

severe constipation because it was a little too strong for them and subsequently came out at a lower dose.

So it can be difficult. I mean, clearly, it's not an easy thing to do. And a postmarketing strategy is obviously important, but that was already too late because those events occurred within weeks of people being on the market, where we're seeing patients with very severe constipation.

DR. HERTZ: Let me also just ask a question, and I hope I'm not derailing where you were planning to go, Shannon. So the other question is, then, are there study design characteristics that might help sort through some of this?

I'm going to say something now, which some people in here may chuckle about, but there are certain study designs that might help clear some of this. So for instance, what about enriching the population based on an early open-label phase? This is something I get thrown in my face as the worst thing we've ever done, but I still say it's a

Page 336

fabulous thing.

So an enrichment for responders and then a re-randomization -- or then a randomization, and then what that might allow is looking at who in that open-label period -- when you haven't had enough refinement during phase 2, because this should not replace phase 2, and then describing that.

So if you wouldn't hide that fact from the labeling, a thousand patients were started in open label; 500 seemed to meet the criteria of the following. They were then randomized. And this is what happened to that population. So you create a narrative, and it's useful for a number of reasons, I think.

I have no idea if your indications have ever considered something like that. This is not an area I've worked in directly. But for non-specific analgesics, for whom there's only subpopulations that respond, it seemed to have improved the ability to demonstrate an analgesic effect in my division.

Page 337

So I throw that out there. And is that something that could be operationalized? It's got pros and cons, but I put that out there.

DR. JOHNSON: I'll throw out another one, which we don't see often, but N of 1 studies. Now, depending on the product, an N of 1 study may not work. But those are great types of studies. And if you have episodic, that might be a problem, but if you have the timing and long enough periods -- do people know what an N of 1 study is? It doesn't mean you only study 1 patient. So do people know what a crossover study is?

So N of 1 is kind of like multiple crossovers within the same patient. So you enroll maybe 100 patients, but each one of them, they might be, A; B; A, A; B, B. So they're going to go through each period, but it's randomized what they're getting. And there may be washout periods or not depending on how you want to do your timing and what the medication or therapy is.

But it's a nicer -- and really, if you think about a lot of doctors, when we talk, that's how

Page 338

they treat their patients. That's how they try to figure out if something's going to work or not. That's another method.

But yeah, the enrichment -- open-label run-ins, again, you have to think about -- we are open to a lot of different study designs. You have to decide how well they'll work and what you're going to say at the end. But as a clinician, a lot of times you're going to put a patient on a drug and see how they're going to do. And if they're not responding well, then you change it.

So in that sense, what was just described is not that different than how you practice.

John?

DR. MARKMAN: John Markman. Can I just follow up to that point about the N of 1 studies? Because I was struck this morning about Dr. Rapkin's point about provoked vestibulodynia. And I just think for a lot of these syndromes, provocation is a really important issue. So it would seem to me, for some of these N of 1 type approaches, provocation and then repeat exposure

Page 339

would be one of the models.

So can you just talk a little bit more about how you think about provocation? Is that a temporal thing? Is that a particular activity? And how much definition is needed around that? Because obviously, what provokes these different syndromes is variable, but I think what I've heard this morning with provocation is a really important consideration.

DR. JOHNSON: Clinical colleagues might talk more about that.

DR. WIEDERHORN: Going back to the MAPP study, I don't know if any provocative entities have been really confirmed or observed. I mean, anecdotally, for interstitial cystitis, spicy foods, coffee, various triggers, but I don't think that's been substantiated.

I think the only thing that I'm aware of in the MAPP study was the fact that if you're going to have a flare, you might have that preceded by -- you guys can confirm this -- two or three days of urinary tract symptoms, and then you have

Page 340

the flares.

Is that true? But I didn't know how you could provoke it for interstitial cystitis because I don't believe these dietary indiscretions, so called, really provoke it. And prostatitis again is the same old song. Don't drink coffee, don't eat spicy foods.

DR. JOHNSON: But for vulvodynia, it might be something, yes.

DR. WIEDERHORN: Vulvodynia, yes, but I'm not aware for urologic stuff how we could provoke it.

DR. DIMITRAKOFF: But for the CPSI -- I think, at the time the CPSI was developed, it was mentioned in the morning that there was an internet survey which asked patients about the most common symptoms, which was done. I think this was a paper published by Dr. Alexander back in 1996.

I think one of the consistent symptoms that evolved at that time was pain after ejaculation. I'm not sure how that codes up in the MAPP study more recently. But historically, I don't think

Page 341

people have actually tried to do provocation studies, at least in the CPPS field.

DR. WIEDERHORN: And actually, historically, back when I was training, which was in the Dark Ages, if you didn't ejaculate regularly, you may have gotten prostatitis. And that was true. I was out at sea, and guys on board the ship came down with prostatitis on a regular basis. And the captain of course was very upset with my recommended therapy.

(Laughter.)

DR. WIEDERHORN: It's unproven.

DR. DIMITRAKOFF: They don't know if it's evidence-based, clinical folklore.

DR. WIEDERHORN: It's folklore, right. It's folklore. So that's the problem. I'm not aware of how you can provoke this. I mean, we also said if you sit for a prolonged period of time, you're a cab driver, a police officer, as you sit, you're more prone to prostatitis. Medical students got it because they were in lectures all the time. You know, no proof.

Page 342

DR. SMITH: Go ahead, Jen.

DR. GEWANDTER: This is for Dr. Johnson. I was just wondering, when you suggested N of 1 studies, are you thinking for phase 3 or more of the earlier experimental studies?

DR. JOHNSON: You could probably make the case. I mean, I think N of 1 studies could be used in earlier studies, but if it's a reasonable study design for phase 3 -- I mean, all study designs are open and available for phase 3. They just have to make sense and answer the appropriate question.

DR. MARKMAN: I just wanted to go back to this. There are provocations. The 6-minute walk test and the results on a 6-minute walk test, which have been used in some recent approvals, is as an example to me of a provoked symptom. Obviously, it's not pain as the primary, but is as an example of a provocation.

DR. HERTZ: We use provocation in a different sense. For managing sprains, for instance, if you measure pain at rest, you're going to get nothing. So the pain is often measured on

Page 343

standing or walking. So yes, I guess I hadn't thought of that. So that's something that's simply been accepted in that context because it make sense.

So for instance, if you're going to talk about vulvodynia and you have behavior that consistently seeks to avoid it, it's going to be very hard to study the drug in that person. And part of the criteria and the conversation that has to occur, then, would probably be, this is the -- if you want to participate, here is what we recommend, and then somehow define what the provocation would be that would be acceptable to put in the study and incorporate that into the study design.

DR. DIMITRAKOFF: I think it also goes back to the definition again. It's like what are you trying to reproduce. I think you first have to define it so that you are able to reproduce it. I mean, there have been studies -- I think, again, there was a test in the past, the potassium chloride test, that people used to do, so you can

Page 344

actually put potassium chloride in the bladder, and people used to claim that this is a test for IC.

But at the end of the day, it turns out it's a test of hypersensitivity, and it doesn't really reproduce the condition that you're trying to study. It has huge overlaps with other conditions of bladder sensitivity.

So again, I don't have all the information, and I don't know what's going on in the MAPP. But from what I know, I just think that we're not there yet to be able to do a provocation study just because we're not really clear what we're trying to reproduce.

DR. WIEDERHORN: I think from a standpoint of provocation, you could also study patients who have a recurrent frequency of either prostatitis, or interstitial cystitis, or recurrent flare frequency of a minimum amount, and then study them for a particular drug. That's not provocation, but at least we're trying to treat the acute entity. I don't know if that answers your concern.

DR. JOHNSON: I'm going to think back to the

Page 345

tampon test, I'll call it, that you were thinking, because this has always been a problem when you have some of these more sexually related issues, which is people, they will choose not to have sex or they don't have a partner at the time for whatever reason. I mean, my sister's married to someone in the Marine Corps. There are a lot of reasons she doesn't have a partner at a time. And what is going in. There can be all sorts of different issues happening.

But what you could measure is at the end of each of those periods in that N of 1 study, or in any crossover study, really, is you may have that provocation to see is it still provoking something, so you can standardize that.

But what you can also measure is everything in between. So we talk about concomitant medications. And really, in some ways, there are therapies and there are actions that people take. And so if you notice that people are actually able to have encounters with that -- so like suddenly these women now feel comfortable using tampons, for

Page 346

example, that's something that you want to measure.

So understanding what it is and then measuring it throughout, that can sometimes become your endpoint, like you've realized that this is what's changing hopefully in your phase 1 or phase 2, and now that actually evolves into an endpoint more so than just, okay, at the end of 8 weeks or whatever, I have this fixed measurement.

We've done that in some other pain studies. You'll see that maybe the level of pain remains about the same, but their level of activity changes, or they are no longer using heavier analgesics, they're using less, something like that, where it's kind of the incidental but very important outcome that's there.

DR. MARKMAN: John Markman. I just think this is especially important because we're living in an age when we're told to emphasize non-pharmacological strategies first, and the most important non-pharmacologic strategies for many of these patients is avoiding the provocation, wherever their pain problem is.

Page 347

So when your hands are tied and you're not supposed to use many classes of medications due to adverse events, it's helpful if our designs dovetail with these sort of non-pharmacologic management, which is the cornerstone.

DR. JOHNSON: But I think that there's non-pharmacologic management, but also remember avoidance can have a heavy impact on people's lives.

DR. MARKMAN: Absolutely.

DR. JOHNSON: So there's a wide variety there.

DR. HANES: Speaking in terms of IBS and avoidance of diet, that kind of being studied at this time, for a few years, looking at specific diets that could be used, like FODMAP and things like that. But is restriction -- it could be beneficial, but how hard is it to do and how much of a lifestyle changer will that be, and will patients move away from them and prefer to take a pill instead?

But I think those things need to definitely

Page 348

be considered in light of pharmaceutical development, what they're avoiding. Certain foods could be the answer versus taking a drug in and that there is a way to look at both of them at the same time.

So I think that's a great question that you have, but restriction is not as easy as it seems, particularly when it comes to something that's essential like food.

DR. LEMBO: I'm curious if any of you could comment on the very high placebo rate we see in all these trials. Sometimes the difference between an effective drug and a non-effective drug is simply the placebo was lower for whatever reason.

Do you have any suggestions or have you seen any studies like placebo run-ins or other things that you think might be effective or would recommend to us?

DR. JOHNSON: So I spent more than a decade working at what is now called the National Center for Complementary and Integrative Health. A very large portion is studying placebo. And I say that

Page 349

because we in fact had requests for placebo research. And we have some clinicians -- it was interesting because I had clinicians who would come up to me who said, "Well, if I've now learned that this is safe and I have what is twice the clinically important difference using what is essentially a placebo, as a clinician treating an individual patient, I'm okay with it. As a researcher, I'm not." So we would try to balance there. Sometimes this costs $800. So that's your public health problem versus not.

But there are lots of enrichment designs that happen, and I think, in psychiatry in particular, you will see a lot of variety of those. But the other element is to also think about the fact that people in studies tend to do better.

So there is a placebo rate that's going to happen in your general population, and then there's the fact that they are in a trial, regardless, that kind of elevates their doing better on anything.

The problem with a lot of the enrichment designs are when you say, "I want people to have so

Page 350

many hot flashes before I enroll them," things like that, is you get a different type of regression to the mean. That will happen. Especially when you're looking at episodes, or severity, et cetera, people tend to go into studies when they are doing not so well, and there is a nature ebb and flow, and then they start going down. So this is a problem when you do some of these run-ins.

Another issue that can happen is -- especially now that we have everything on clinicaltrials.gov, which is a great thing. At the same time, people many times know what they need to do to cross that threshold. And we now see increasing numbers of people that are right at those thresholds as they come in.

So there's a lot of balance there, but realistically, I think it's also a lot of our tools and trying to measure what you really want. So when we were using the AUA scale for this one urinary study that I did at NIH, literally the standard deviation was 3 times the clinically important difference that we're supposed to be

Page 351

looking at.

So when you have measurement tools like that, you have a problem. And the placebo rate -- and this includes -- we pulled it off of three large trials for an FDA-approved product, and our placebo rate matched pretty much exactly for that trial, and it was twice. So this becomes an issue. But a lot of it, I think, in that case came down to the tools.

So we talk a lot about populations and all these other things, but many times, it's how we are measuring people as part of the problem. But also remember, placebo is not benign. You're talking to people and, yes, it's there. It's doing something. So you need to get above and beyond with whatever that new therapy is of basically providing empathy.

DR. SMITH: John?

DR. FARRAR: I thought Bob might say this. This group, actually, the IMMPACT ACTTION group, published a couple of articles about attempts to try and control some of the placebo effects in pain-related clinical trials. I would suggest you

Page 352

look at that article because I think it covers a number of things that Laura was just talking about and suggests some ways that might be able to get at it.

Just to reiterate what Laura said, what happens to the placebo group and what we think of as the placebo effect, i.e., the brain-body or mind-body, are very, very different. Most of what happens to the placebo-treated group is the natural history of disease or their regression to the mean for all the reasons Laura was suggesting.

In the mind-body component, there are circumstances where it's relatively easy to understand. So if you're studying a new opioid, it won't happen, but if you're studying a new opioid, we produce endogenous opioid. And people who are in a lot of pain produce a lot more endogenous opioids, so that you could see that they would have a larger placebo effect, even if they weren't in the treated group.

But I think, without getting into much more detail about it, there are several published

Page 353

articles that would be useful in thinking this through. It hasn't been tried in the diseases you're talking about -- well, I don't know that. But we've looked at it with regards to pain and depression.

DR. JOHNSON: I think Ted Kaptchuk has also looked at it in IBS, for example.

DR. FARRAR: Yes. And of course, there's all of Ted's work and all those other things.

DR. DIMITRAKOFF: So there is one paper in the CPPS field, which actually estimated the placebo rate or the placebo effect, that I'm aware of. I think it's also a question of how long that effect lasts, and I think that's what we don't really know.

I think sometimes there are studies where people have estimated the placebo effects. And we're talking about numbers and I think this is an important consideration. But it's also important to know how long that effect lasts, because that's an important consideration when you're enrolling patients.

Page 354

I think in terms of design, I think Dr. Johnson mentioned Ted Kaptchuk's work. And I'm not sure if this is something that he discussed or I read it somewhere else, but I think there is a design where you can do a placebo-controlled trial, and you can actually have a third group where you have patients with a condition, and you promised them that they will eventually get the study drug. And you kind of use them as an active control to your placebo population, because you actually have a way of somehow -- you have some sort of control of treating the patients' expectations or part of the placebo-induced response.

I'm not sure in particular if this is correct, so please correct me if I'm wrong.

DR. JOHNSON: I think there are a lot of different ways that can be approached. But if any of them was a slam-dunk, you'd be able to just to rattle it off, and people would use it, and we'd be done.

DR. WIEDERHORN: Yes. There's another pitfall with placebos that we've encountered, and

Page 355

that was that we did a study with a run-in period. Placebo responders were eliminated. The non-responders were then randomized, placebo active drug. Guess what? The placebo effect was still the same in the repeat trial. So it depends on what you're studying and your measuring instruments. It can be very tricky.

DR. DIMITRAKOFF: I think there is data in the literature that shows it could last up to 6 months --

(Crosstalk.)

DR. WIEDERHORN: I really thought it was much shorter, but you're right.

DR. DIMITRAKOFF: Yes.

DR. WIEDERHORN: It's much more than I thought.

DR. SMITH: Do you want to go?

DR. DWORKIN: Sure. So I'd like to go back to this heterogeneity issues just briefly, which was mentioned with respect to interstitial cystitis and the changing diagnostic criteria with respect to newer approaches to bladder pain syndrome.

Page 356

It seems to me the risks of having two heterogeneous groups of patients in the clinical trial are two that I can think of. One is it makes it hard for the agency to write a label because the label needs to characterize the patients that were studied in the clinical trial.

But assuming that bladder pain syndrome IC can be defined by MAPP in a way that could be used in a label -- and I don't know, but I'm imagining that they will get there if they haven't gotten there already. But it seems to me the only other risk of heterogeneity is the sponsor risk.

So I guess when I think about the fact that the major risk of heterogeneity, assuming that a label can be written, is a sponsor risk, then I'm sorry if I'm beating a dead horse, but I don't really understand the resistance to moving beyond into the present a 1988 definition, because it's mostly or entirely a sponsor risk.

DR. WIEDERHORN: The first comment is there obviously could be better phenotyping even with these large groups, BPS and traditional

Page 357

interstitial cystitis. That's one of the things we're looking for with MAPP results, was maybe we can identify subgroups that are going to respond. But the second thing is who then defines what these diseases are.

In 1988, 1987, the NIDDK criteria, that was done by consensus conference at the NIH. The question is, whose role is it to define or reform things?

The other question of course is fairness to approving drugs. You approved me under this, but now you're going to tell me that my competitor gets approved under something else. So there's fairness to the corporations. That's also an obligation as well as fairness to the American people, trying to develop drugs that are effective for a disease that can be debilitating. So yeah.

As I said before, I asked Dr. Star this question a couple years ago at one of the meetings -- I forget which one it was -- as to when are you going to have another meeting, because at that point, I was looking to the NIH to look at

Page 358

this. Now, these are my views. That's not the views of the FDA.

DR. DWORKIN: So maybe that will be our next meeting next year. We'll have a meeting to come up with a 2018 definition of IC/BPS.

DR. CLEMENS: I'd be happy to attend.

(Laughter.)

DR. DWORKIN: Please go write into your date book next year at this time.

DR. JOHNSON: I'll tell you in a general sense a plan for the heterogeneity plan. If you know you're going to have a high placebo rate, plan for it. I think for a lot of studies, the things that come in, many times our comments are, like this is just wrong; please don't do it. But a lot of times, we're like, you are taking a risk, and by the way, we're trying to give you scientific advice because I can tell right now reading this, you're going to fail because you haven't taken into account what is very clear is a huge problem.

So just plan for it. I mean, most of the time, things are fine, but it's a lot of things.

Page 359

Plan for what you know is a problem. It's very basic.

DR. DIMITRAKOFF: Unless there is a patient with a monogenetic phenotype, and someone comes up with a drug within the next year, and then it would be good.

DR. DWORKIN: But that will be an N of 1 trial.

(Laughter.)

DR. SMITH: Jen, and then we'll go to Rob?

DR. GETWANDTER: So I was actually thinking about the N of 1 trials again, and when I think of them, I think of breakthrough cancer pain, which the whole trial can be done maybe within 7 to 10 days.

So I was wondering if some of the content experts for IBS, IC, and prostatitis could comment on how often patients are having these very cyclical pain flares, how long they take, how long it would take to have another one. So in essence, would it be feasible to be doing these kinds of N of 1 studies for the acute flares in a reasonable

Page 360

amount of time.

DR. LAI: For IC, I think flare happens fairly often. More than 90 percent of the patients will report some kind of flare. It varies in terms of duration. It could last anything from a week to a few days to one days to a few hours. Some people describe minutes' long flares.

So it becomes I think perhaps difficult if you're doing an N of 1 because of the natural history of somewhat unpredictable flares happening. That's one consideration, and washout is another thing you can think about.

It is common, but it does impact the quality of life. We did some studies, what changes during flares. We looked at pain, and we looked at urinary frequency, urgency, and they all go up. You expected them to go up. But at the same time, it affects quality of life.

The longer flares, things that lasted more than a day, maybe up to a week, affects quality of life more than a timely little so-called flare that lasts a few minutes.

Page 361

So there are some qualitative -- and also we did some focus groups, some qualitative, quantitative data on these. So this could potentially be useful as potential things to look for in future clinical trials or a study.

DR. LEMBO: In IBS, I think it's probably similar that there's a wide variation. I'm not aware of any clustering between a specific number of days. But in trials, I mean, we get down to who comes into the clinical trials, that's the intermittent pain that people just have intermittently such as meet the Rome criteria, they almost never come into trials. We don't see them in clinic. Most of the patients have pretty regular frequency.

They do get worse, and there will be some fluctuation based on a variety of factors. We're doing several studies now. Right around the elections, we had tremendous flares with patients.

(Laughter.)

DR. LEMBO: That was the reality. Every patient came by. And women will report worsening

Page 362

systems around menses. I mean, there's a whole bunch of factors that come in. And that's why, as you've said, you need to factor that into the size of the trial, and that's why you have these control groups.

As far as the N of 1, crossover designs, particularly in IBS, are not realistic to do because patients never get back to baseline. And it's been shown over and over again. It may take a tremendous amount of time because the placebo effects were long-lasting.

As you appropriately stated, there are numerous factors that go into why people may or may not respond. One of the things that we see over and over again now is that it's all of the non-specific effects that are occurring.

The classic example in our current trials where we actually are talking to patients afterwards is you got the placebo. Why did you get better? Well, I changed my diet. You said don't change it, but I changed it. That continues on afterwards, so they never get back to their

Page 363

baseline. That's our experience. So we really stay away from crossover designs.

DR. JOHNSON: That's actually one reason I kind of like the N of 1 versus a straight-up crossover because you cycle through multiple times that you get to be in that placebo or whatever your active control arm is. So you kind of get away from the problem that they never come back completely to baseline.

As long as you can have somewhat of a washout of whatever that new therapeutic is -- and again, it doesn't work for everything, but that's one part that's nicer compared to just a simple crossover.

DR. DWORKIN: I'd like to ask everyone about discomfort. Over the course of the day, there were a bunch of mentions of pain and discomfort or pain or discomfort. And my sense is that there's no validated measure of discomfort, so it's not like pain where we all agree that a 0 to 10 scale or a visual analog scale is well enough validated to be a primary endpoint. So does that mean -- and now

Page 364

I'm fast-forwarding to tomorrow's discussion.

Does that mean that, at most, discomfort is a secondary or exploratory endpoint because we just don't have a measure that would allow it to be primary, or did someone know about a way of measuring discomfort that's just kind of reasonably well validated?

DR. DIMITRAKOFF: As I said in my presentation, I'm not aware of any instruments that have been used at least in the CPPS field, but I think part of the reason is that discomfort probably means different things to different people.

DR. DWORKIN: We don't know.

DR. DIMITRAKOFF: So it's probably like validation.

DR. DWORKIN: So the same way that Charlie Cleeland started work on that brief pain inventory four years ago, someone -- maybe not someone in this room, but someone somewhere should start work on a kind of brief discomfort inventory that in 30 years could be incorporated into clinical trials

Page 365

of conditions where discomfort is an important symptom.
Are you volunteering, John?
DR. FARRAR: No.
(Laughter.)
DR. FARRAR: I had a comment, though.
DR. HERTZ: Add fatigue in with that one.
DR. DWORKIN: What?
DR. HERTZ: Add fatigue into that.
DR. DWORKIN: That's right. Fatigue, yes.
MALE SPEAKER: Add urgency.
DR. FARRAR: So the 0 to 10 scale works fine for fatigue, and the PROMIS measures do pretty well, and the frequency as well, and you can count.
The issue about discomfort I think is an interesting one because in our patient population, I see a fair number of palliative care folks who get chemotherapy, and they'll talk about the numbness and tingling they get from the neuropathy as uncomfortable.
The question is, does it bother them enough to need some treatment or not? And I would argue

Page 366

that the same might be said about pain, which is probably the best post-operative measure; is this okay? Could you go to sleep this way, or do you need anything else?
So I think that there is the potential for measuring that. What I was going to say, though, is that discomfort becomes painful. I have had patients who come in and say this is uncomfortable. Yes, I'd like a little treatment. And then when it doesn't work, they say, "This really is starting to hurt now."
So I'm wondering actually whether -- I have never seen a patient who comes in and says this is so uncomfortable, I can't stand it. Now, maybe there are people who do that, but mostly they come in and say their pain --
DR. DWORKIN: Think about itching. Itching is not painful, but it can often be intense enough that you can't stand it.
DR. JOHNSON: There's a lot of itching --
DR. FARRAR: I agree. So I guess what I'm saying is that I think discomfort covers a bunch of

Page 367

different symptoms, like this urgency is so uncomfortable, so discomforting. And I wondered whether discomfort is really a symptom unto itself or whether it's really just part of others.
I honestly don't know the answer to that, but I think it might be worth looking at.
DR. DIMITRAKOFF: I think there is a part -- I'm sorry.
DR. SMITH: Go ahead.
DR. DIMITRAKOFF: There is also a cultural element. I think there is also a way of how people perceive discomfort. Your words actually remind me of a discussion I had with a colleague about 10 or 15 years ago, and we were just talking about various -- I'll just give you a very clean-cut example.
We were talking about dysuria, which is a very clean term, and we know what it means in North America. But then in other parts of the world, it turned out that -- it was part of another project, but he was working with someone who was studying people with dysuria. And the people in the other

Page 368

specific population actually, for them dysuria meant obstruction, difficulty urinating to the degree where they would be uncomfortable enough to call this what we call dysuria here, which is painful urination.
So I think it's a perfectly valid example of what you are saying, that it is actually very important.
DR. SMITH: Hanna, did you have a comment about this specifically because there are a few others -- okay, go ahead.
DR. GROL-PROKOPCZYK: Just really quickly, for what it's worth, when I worked with WHO data from 10 countries, where people are asked to rate pain and discomfort, the two were extremely closely correlated. But it occurs to me that maybe discomfort more than pain is really context-specific because when Lesley Hanes was speaking, it seems like there's ambiguity in some cases whether discomfort can be a synonym for bloating, or could it be a synonym for some other more specific symptom.

Page 369

So it seems like maybe that's something you'd get into before you start designing the discomfort scale.

DR. SMITH: Did you have a comment about this, too?

DR. CLEMENS: Yes. I was interested in Bill's talk, where he showed that 40 percent of the IBS patients didn't report pain. It appears that the IBS field has chosen to exclude that 40 percent, perhaps wanting to have more internal validity or more homogeneity to the population, and the IC world is uncomfortable currently at least with excluding that 40 percent.

I wonder if it's more of the nature of visceral pain or where patients think of pain as burning themselves when they're cooking or something. This is inherently different, and some people just don't view it as -- they're never going to call it pain.

They're going to use another term for it. And the concern is from the academic world that if we could educate them in 30 seconds about what

Page 370

we're referring to, to pain, they'd say, yes, I have pain, because what I have definitely meets that IASP definition.

So that may be where some of the concern is, that at the end of the day, it's perhaps semantics or terminology that's getting in the way of us studying the type of patients that we all would like to study and help. And at least for me, I think that's one of the concerns I have, that it's more of an education thing than it is perhaps -- the patients and the communication thing than anything else with some of them.

DR. LAI: I don't think it's totally a cultural thing at least in interstitial cystitis. There are some patients here that I see regularly who say that it is not pain, but it's pressure. It's very intense bladder pressure, but it's not pain. Is it 0 to 10? Is it pain? Is it 1? No. It's not pain. It's pressure.

So there are people who actually perhaps not get to the point to see you that they complain about pain.

Page 371

DR. CLEMENS: They're urinating every 30 minutes.

DR. LAI: They're urinating every 30 minutes, driving them crazy.

DR. CLEMENS: It's having a substantial impact, whatever it is.

MS. VEASLEY: Chris Veasley. Some women with vulvodynia will say the same thing. They'll say, "I have burning. It's not pain." We'll say, "How is your pain?" "I don't have pain. I have burning."

I want to just quickly go back to something that Sharon mentioned just to answer your question, I think, as yes. We would like to be able to know which drugs are working in different subgroups of patients with these conditions, and we would like to be able to use different methods like enrichment to do that.

The problem is that we know these conditions are heterogeneous, but we don't know how or what we should be using to phenotype. I mean, the whole reason why MAPP began is because all these IC

Page 372

trials failed. So they had to take a step back and say we need to understand this better before we start trying to move forward with larger clinical trials.

I think MAPP and ARP are beginning to really be informative about what some of those different domains need to be and how we can better characterize and phenotype patients. But until we have that information, I don't think we can do what you asked.

Just one quick observation. Having worked in this area with these different conditions and ones that aren't addressed here, it seems to me that there is something different about IBS because IBS has been more successful in getting positive trials and approvals versus these other conditions. There's no approvals for vulvodynia. I see trials keep failing.

I'm wondering if it's because, one, maybe there's more corporate involvement in IBS, are there more trials going forward, i.e., the larger percentage of approvals? Do we know molecularly

Page 373

something different about IBS? There are other targets that are being more specific.

Is there something different about the way IBS trials are being conducted than some other trials in this area? I mean, what might be some of those differences?

DR. HANES: I would say that is hard for me to answer that question. So basically the question is, what is different in the IBS trials versus IC trials or vulvodynia and those type of trials -- I believe that's what you're asking -- that's making the allowance of drugs to be approved in IBS versus perhaps other disease processes?

So I can't speak upon the other disease processes since I haven't worked in them, so I don't know necessarily about their trial structures or what sponsors have proposed and what the discussions have been with the FDA during the process.

I think that every division at the FDA, although there are a lot of similarities, there are differences. So what DGIP, the GI division, might

Page 374

encounter in working with sponsors, and companies, and investigators might be completely different than what DBRUP is doing or not doing, but what has been presented to them from the outside.

So it's definitely a working relationship with pharmaceutical companies and investigators to try to lead to the best optimal drug development process.

DR. HERTZ: I think just to build on, where is the opportunity to explore those questions? So for instance, the drugs for IBS, it sounds like they pretty much affect the defecation pattern, and then there's also an effect on pain.

So are there any products for IBS that are neutral on the actual bowel function but also affect pain, and have those been tried in these other areas, where maybe the target is not local, but central? I don't know if that work's been done, but that might be something for consideration as a way to expand success.

DR. HANES: I definitely agree. I think that it does get back to the basics where the

Page 375

mechanism of action is being proposed. Is it being proposed on a molecular physiologic level? Do they have, say, a receptor that they are targeting, or do they not know how the drug works, and they think it might be a central way that it's working versus a localized intestinal way that it's working?

So I think it definitely depends on the concept of the drug, what has been shown in animal models starting from there, and then what's been shown or what's being proposed to occur in the human.

So I think that's a huge part of it, too. And we have seen proposals looking at one aspect of IBS, so looking at either the abdominal pain aspect or the defecation. And the biggest thing with those -- which we don't discourage, we want to see a variety of different drugs that might help components of IBS because patients need that, and they might not necessarily suffer too much with belly pain, but they suffer more with defecation issues.

So I think that we welcome evaluating and

Page 376

helping with multiple drug development processes, but the issue with that is that we want to make sure that symptoms aren't worsening in the ones that it's not being targeted to or at least that it's being addressed.

So say a drug is targeting abnormal defecation. We want to make sure that at least the programs are really looking sincerely at abdominal pain as well, making sure that it's not worsening, and that they're identifying whether there's any problems with that.

I'm not sure if that answered your question. I'll let my colleagues answer in terms of kind of what they've been doing. But I would say that at least in my experience -- I've only been with the FDA for two years, but there is a lot of contact between FDA and collaborators, and there's frequent meetings with companies. And they definitely present their protocols for the most part, and we try to work with them. So I think that's a good thing. It happens universally.

MS. VEASLEY: I guess just the

recommendation would be, theoretically, we don't think there's a big difference among these disorders, to take a deeper dive in looking at how clinical trials are conducted across these conditions and see if there's lessons learned or better practices that are being applied to IBS that could be applied to other conditions.

DR. JOHNSON: So that might be something that the NIH -- a lot of our reviews are public for at least those approved drugs, and then there are a lot of others. So that's something I think Sharon Hertz laid out kind of a nice plan of ways that you could look at that and see what could come up.

DR. SMITH: Let me be respectful of the fact that it's 5:00. I know there were two other questions. Is it possible for those to wait until tomorrow during our discussion period, or are there pressing questions that you want to get in today?

(No response.)

Adjournment

DR. SMITH: No? Okay. So why don't we end now? Dinner will be at 7:00 p.m. on the mezzanine

level. So we'll see you all then. Thank you so much for today.

(Applause.)

(Whereupon, at 5:02 p.m., the meeting was adjourned.)

B

D

E

F

G

H

I

N

O

P

Q

R

S

T

U

V

W

Y

Z