ACTTION - IMMPACT Research Design Considerations for Randomized Clinical Trials of SCS for Pain

---

November 16, 2018

A Matter of Record
(301) 890-4188

ACTTION

INITIATIVE ON METHODS, MEASUREMENT, AND PAIN ASSESSMENT IN CLINICAL TRIALS

Institute of Neuromodulation
International Neuromodulation Society

Research Design Considerations for Randomized Clinical Trials of Spinal Cord Stimulation for Pain

Friday, November 16, 2018
7:59 a.m. to 2:52 p.m.

Westin Georgetown
Washington, DC

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

(7:59 a.m.)

Presentation - Simon Thomson

DR. THOMSON: So welcome back, everybody, to day 2. Just a little bit of housekeeping, when it gets to the discussion, and we've got a lot of hours of discussion later, could you try and remember to lean forward and announce your name for the transcripters. We sort of lost it last night, yesterday evening, although I think some people's accents are probably recognizable.

The other thing is if you are leaving today, just remember that checkout time is 12:00; otherwise your key card doesn't work. So that's that.

My task is to talk about study execution, and I think you'll find this as a little bit of a revision of some of the points that we started discussing yesterday.

(Pause.)

DR. THOMSON: Spinal cord stimulation, and I know we're talking about randomized-controlled

Page 5

studies for pain. We tend to think of this as a treatment for neuropathic pain, but I think the clinical distinction between mechanisms of neuropathic and nociceptive get a bit blurry, and at the end of the day, it is a clinical differentiation. It's also a treatment for visceral pain, although it has not enough clinical science to support it.

It's a treatment for ischemic pain syndrome, chronic critical limb ischemia, vasospastic disorders, cardiac ischemia, and mesenteric ischemia. Then it's also used in other conditions. It can be used in stabilizing ventricular dysrhythmias in heart failure; spinal cord injury, as has hit the news recently; persistent vegetative states it's been used in; and even augmenting brain tumor chemotherapy.

Now, when it comes to study and design, I use two phrases there, the devil is in the detail and the word "equipoise." And we're going to be talking about these titles. We're going to talk about recruitment, recruitment of centers and

Page 6

recruitment of patients. We're going to be talking about patient information, written website and social media. We're going to talk about randomization and patient education and training in the outcome measures. We're going to be talking about efforts at blinding and reporting on blinding. We talked briefly about programming. We're going to talk about the sham and generally about outcome measures; how do we actually measure our primary outcome?

This is one of Sam's slides on equipoise. He found this, and it turns out it is a drug for horses, but basically, it's a principle of research, genuine uncertainty whether a treatment will be beneficial. And that should be the position of not only the investigator, the staff, but also even the patients.[indiscernible]

Sources of bias, I think this will probably be one of Nate Katz's slides. Subject expectation comes from research staff who are overly enthusiastic about one treatment over another. It will come from looking at the patient information

Page 7

sheets, what's available online about one treatment or another, and even subjects' word of mouth.

I think as we've heard, expectation bias can be very destructive and bias studies to the null, bias one treatment over another, and be at least as large as any treatment effect. And I think there's evidence to show it can be long lasting and in some circumstances even indefinite.

Recruitment in the research center, somebody mentioned about good clinical practice training, and that should be, of course, a standard for anybody. And if they are doing their training properly, they will understand many of the principles about clinical research that we've talked about.

I briefly mentioned how in Cambridge they started doing a spinal cord stimulation trial in refractory angina, but really they had no training in how to do spinal cord stimulation or look after their patients. But what we often don't think about is what about the comparator treatment.

It may not be necessarily another SCS

Page 8

device, but even then, they need to be skilled, that that has different programming opportunities, different lead positions. But it's often if you're comparing against an alternative treatment, have they skills providing the comparator treatment and if it's a more pragmatic study with usual care, can they provide a broad range of usual care treatments?

Similarly, what about the outcome measures? Because they're not just pain scores and tick boxes. It might be exercise, the 6-minute walk test. It might be other outcome measures that are specific to the disease that you're researching.

As I say, the second bit is actually what I've just said.

When it comes to patient recruitment, I think it's important that the referrer -- so if you've got people referring to your center for research, when that interaction they have with the patient might be, "I know just the treatment for you. They're doing this really interesting study on this brand new treatment that's just come over

Page 9

from Europe, and it's wonderful; you should go to that center" and then you get randomized to the non-treatment group.

So it extends even beyond the referral center, and everything is all about trying to manage expectation bias. The care that the patient should receive should not be dependent upon research participation. I think we often find that if they are going to get this wonderful device, then they will only get that wonderful device if they're part of the research study. And it goes further that they may not even get spinal cord stimulation because they don't have the insurance cover unless they're in a research study.

The SCS should be universally available if you are taking them into a study. The patients should be equipoised. They shouldn't come with pre-conceived ideas that the investigator treatment is going to be better than the comparator treatment.

Ideally, the situation is that you're referring the patients to the center for a second

Page 10

opinion, not specifically for SCS nor this interesting new treatment. And as I say, SCS should be available, and patients should be provided with factual and equipoised information, and be indifferent to treatment randomization.

The patient information sheets and the literature available should not just be about the investigator treatment but also the comparator treatments, and that should be explained with rates of success and complications. But what about that that's present in the public domain? What about on the website? What about social media? Should, for example, industry websites be suspended during the time of a recruitment; patient information sheets examined by a third party, and we're going to be talking about a case example soon. Can you control social media? Obviously, we all know we can't.

Here's an example of a patient information sheet, and apologies that it has to be a named company, but this was what was given to patients. The clinical study says that Senza is designed to treat chronic pain in the trunk and limbs at least

Page 11

as effectively and without some of the potential side effects associated with currently available SCS systems. The Senza system is designed to treat chronic pain in the trunk or limbs without the need for a buzzing sensation. So we've got sort of a placebo statement and a nocebo statement.

Then if you turn to the website that was available during patient recruitment, conventional SCS provides considerably less relief for chronic back pain for most patients. So you've just now told the patient that they're going to be randomized to the control group and it's going to be less good.

Parasthesia. In fact, 71 percent of patients reported uncomfortable stimulation on a large survey, a nocebo statement. Process study results, poor back pain relief. The study is about the relief of back pain as the primary outcome.

Although SCS provides meaningful relief for leg pain, back pain relief is still a challenge for most SCS patients. The goal is to provide superior efficacy without the uncomfortable stimulation

Page 12

commonly experienced with conventional SCS therapy. Only the Nevro system delivers the unique waveforms designed to offer compelling back pain relief and avoid the side effects commonly associated with conventional SCS. So you get the idea.

Here's an example of an equipoise statement. Do you remember writing that? I'm sure you thought a lot about it, Richard. The study was presented to candidates as a comparison of two standard, non-experimental procedures, SCS and reoperation, to determine whether SCS should be offered as an FBSS treatment before or after exhausting all reoperation treatment options.

In our own PROCO study, standard SCS uses stimulation frequency between 40 and 100. "In recent years, and SCS device capable of giving frequencies of stimulation as high as 10K has been used with claims of improved back pain relief and without the patient being aware of the stimulation. However, the science to support this claim is not adequate. Furthermore, it is not known as such high frequencies are required to achieve the pain

Page 13

relief."

So when it comes to consideration for our guidelines, I think the important thing is to be transparent. We have to document our efforts to balance research and subject expectation between groups and measure expectation. We can actually measure it of patients both at baseline and the en point for researchers and subjects.

We have to be absolutely transparent with making available the patient information sheets and what was available on websites at the time. So that's when it comes to reporting.

The randomization process, this is probably the one bit that we are quite good at because we do realize that we use mostly computer-based systems to generate some randomizations, although not all as we heard. The recruit has to feel equipoised about what group they've ended up in.

Now, there were reports -- and I haven't got this in the public domain, but having been around a number of the centers, I remember being told of how patients would sometimes be weeping that they'd

Page 14

been randomized to the non-Senza group. So they weren't in different to which treatment they ended up in. We need to also look at maybe surveying the satisfaction patients have with their randomization.

When it comes to patient education, it's difficult enough to educate people firstly about chronic pain; secondly, about this complex treatment; and then thirdly, what are the comparator treatments that are available; and then the outcome measures for a clinical study. It's an awful burden for patients to take on when they're coming into a study, so is sufficient time and learning experiences available to patients when they come in for this study?

There is also the physical burden. In the PROCO study, as you see, we collected real-time pain schools. The patients had to wear this watch for 9 months, inputting this data 3 times a day, every day. Then we talked about blinding. It's true to say that most RCTs in SCS don't have any blinding. And what's worse is not only do they

Page 15

know what the treatment is but also its effect, either positive or negative. So when they end up in that group, there's extraordinary expectation.

As we heard earlier, yes, we all know it's difficult to double blind, but it doesn't mean you can't single blind, at least the data collectors. We've been working on studies where we have clinical teams and research teams, and one is unblinded and one is blinded in order to be able to carry out the therapy, but the people who matter when it comes to data collection are blinded. But then everybody has to maintain that blinding discipline; not least the patients as well.

These are the statements that we've heard, and this is one of Sam's slides; subjects and investigators. So these are quotes the write up of a variety of different recent studies where there's always a statement about the blinding, in other words, why they've not done it. Subjects, investigators, and study site were not blinded to subjects assigned therapy. And that can be even within. So like with the SUNBURST study, it was

Page 16

within device but two different modes, subsection and parasthesia-based programming.

Due to practical considerations, study subjects and investigators were not masked to the assigned treatment group. Given the nature of the intervention, it was impossible to blind patients and difficult to blind investigators during this trial; and this is why we get such a low recommendation from those outside our field.

There are a few studies double blind, not least the PROCO RCT I was involved in. Then people will put forward the Alkaisy study, which although the work was done before ours, it was published after, actually when you read it, a third of the patients had a perception of parasthesia of 1 kilohertz. But as I say, there could be a single blind, and in Jose De Andres' study, they that were single blinded to the observers.

I think there is also this concept of having interactions scripted or monitored. Often with our interactions with the field engineer programmer, there was always a research nurse monitoring the

Page 17

interactions, and then again, that can be documented.

We talked briefly about the programming. And yes, there are advantages of why you want your expert programmer being able to use that technology to the best degree, and I understand that. We talked about how long will that go on and how many visits that will take. And again, I think that should be protocolized to some degree or at least reported on.

Particularly when some of the studies -- so if you're, say, doing an angina study with maybe a 6-minute walk test, traditionally you would always have your research nurse walking with that patient. What might be the interaction that's going on? The patient will know that they've got a device in. So what encouragement is being made to -- "Come on. Let's keep going" is that different, too, at the baseline, for example.

We talked about the dichotomy of programming. Yes, we want the clinical team. If you have the clinical team doing it, you might have

Page 18

suboptimal programming, but we talked about training. Do you use the industry research scientists? We keep banging on about industry, but essentially we got together our clinical scientists leadership from the large industry -- that's just their employer -- but first off, they're actually into the research. They're not marketers; they're into the research. And that's what they tend to provide when they are supporting products in the field, people who are into the research.

So we could make it explicit that that's good practice, but what we don't want is them to be contaminated by marketing objectives.

Or do you use the commercial team? Or as what happened in the Senza study, you had the research scientists, who actually had an office within the hospitals, versus the normal commercial team who might visit every 3 months to reprogram a patient. There should be efforts to control the interaction, monitoring by the research team, and scripting maybe. We talked about the duration and frequency of programming sessions.

Page 19

What about the sham? What's the purpose of a sham? The purpose of a sham I think is twofold, and I'm sure I'll be told otherwise. One is you need this to look for the treatment effects over the non-treatment effects. But also, when it comes to things like inferiority, if you've got both A versus B being the same, does that mean they're both as good or both as bad? That's one of the reasons.

What about sham in parasthesia based? Well, you could do investigatory surgery, and an awful lot of people outside our field would love us to do this, is to be able to look at unilateral treatment for bilateral pain, unilateral stimulation for bilateral pain. Then I think it was Solomon Tesfaye -- I said 1998, but Liverpool, who in diabetic neuropathy, they had a red box. It was an external connected to stimulator, and they had a red box that came on even when it was a sham stimulation. Then they surveyed the patients who all apparently believed that they were having the active treatment.

Page 20

Then there's this idea of the ultra low-dose SCS, one stimulation like they did in an angina study, Zipe's study. But then we heard, "Well, actually, is that an active treatment?" Then what about subperception program? And as I mentioned, that can go awry because sometimes these patients' different positions, if they start to activate, then turn up the amplitude, they start getting sensation.

Also, there are problems with draining the battery because if you're using rechargeable systems, there are fears that if the patient notices they don't have to recharge their device, they know that they've been on a sham. And the issue of whether a patient can tell whether they've been on a slightly different current consumption; if only patients were that clever that they could actually tell, and therefore, "Oh, I know. I've only had to charge once today. I must have been on 1K." That's not what happens; I can tell you that.

Let's get on to the interesting bit on outcome measurements. Obviously, what we choose,

Page 21

it really depends on what is the research question we're trying to answer and who is asking that question. As far as I'm concerned, as a clinician, I'm treating long-term conditions, and I'm not a believer in the pain score as being a useful measure for long-term conditions. I'm much more interested in health-related quality of life and improving that long term. Others are interested in function. Others might be interested in medication reduction.

Of course, we have to select. This is a single primary outcome measure. And we always dutifully collect the secondary outcome measures. And as you'll see, it's important to be able to blend the two because what happens in secondary outcomes might actually explain the validity of your primary outcome.

So when we choose the primary outcome -- and we talk a lot about regulation, but, hey, we're beyond regulation in Europe. It's now about reimbursement. This is a wave that's going to hit you in the U.S. As you're starting to notice, it's

Page 22

not about the regulatory bodies; it's about the reimbursement. It's about CMS.

So what is the actual answer to the question that you want? Is it to satisfy regulatory demands? And they're going to say your device, you say it's a treat pain, tell me about pain. But to satisfy your reimbursement people, they don't want to know about a pain score. It means nothing to them. They want to know about return to work, or function, or quality of life.

Anyway, let's talk about the pain score. We hear a lot of this percentage pain relief. This is what we use clinically, isn't it, in the clinic. We say, "Look, you've had this stimulator on now. What percentage pain relief have you got? 100 is complete, zero not at all." And they give you a figure. And you know and I know that always exaggerates; anything you can measure with an NRSPI difference. That's the other way we do. Sometimes we use a VAS scale, actually a proper VAS scale with a mark on a piece of paper, on a line, or more typically we use NRSPI and express it as a

Page 23

percentage change.

Now what do we do? Actually, what happens in a lot of these studies -- and I'm going to get onto this -- is it's a single point. Looking back over the week, the patient [sic] said, "Here you are, now. This is your data collection time. What's been your pain like over the last week?"

You might divide it up at worst and average or best, and generate 3 scores. You might do that or you might just do one. What is one after? Are you after the worst or are you after the usual? It's often not explained. And the diabetic neuropathy ones, often pain worse at night, they did a day-night one.

Or are you going to measure a mean pain score multiple times of the days over 5 or 7 days? And if you're going to do that, are you going to use a pain diary which notoriously are incomplete, or are you going to use -- one of the ways we got around it was a watch strap, and they had a sliding scale, and it bleeped at them every 8 hours, and they inputted a data point.

Page 24

Then the other thing is we have to have a strategy of what do you do when they don't put the pain score in? We had a paper diary backup, for example. And often the patients with that device, they knew if they had done it wrong, and they just jot it down. Paper diaries, there's data to show that they're only in 11 percent of cases complete. They're often done in the car park at the data time collection points. So often their memory of their pain over that previous week might be unreliable.

As I said, this is what we did with the PROCO, and it does at least take the observer out of it. These are just a little private moment they have with their watch strap as they input their pain score. And it does mean we can monitor throughout several days. But one of the things the Alkaisy study did is they used the pain scores over the whole period, whereas in the PROCO study, we were doing, if you like, our optimization. And it was only in the last 5 days where that was the data collection period, and there's the watch strap we used. That was the PROCO study, which basically

Page 25

went to show that the determinator of outcome was not the kilohertz frequency.

So let's have a look and see if we can be informed as to what might be important how we measure the pain score. This is actually the SSED, which you and McNicol, I think you should somehow include in your literature search because there is so much more data -- I don't know whether you're allowed to do it, but anyway -- in those to explain what's going on.

This is the Senza study at the primary endpoint looking at the responder rate and this fantastic figure of 78.3 percent, which took our field by storm. But they did do a diary. It's not in the write-up. They did do a diary. And there's an 11.6 percent change in their primary outcome measure, 11.6 percent difference in responder rate, just simply dependent upon the methodology that was used for measuring the primary outcome. It was the same in both groups, but if you like, the marketing message has been to say what a wonderful responder rate we've got, and that's been waved in our faces

Page 26

quite a lot.

Let's look in more detail and look at the other secondary outcome measures. Which one correlates? Is it 78 percent? If we look at global impression of change, this is the percentage of patients who would describe themselves as better or a great deal better.

Now, if you had a responder rate of greater than 50 percent reduction, I would expect that to be near the 78 percent mark. So what's better? If the subject expectation is 52.8, 2 percent subject global impression change, is it the responder rate or is it the pain diary? So we need to describe what scores actually best describe the outcome.

Let's look at the ACCURATE study, and here, this is the composite outcome of the two groups. 81.2 percent achieved this composite outcome, which I think was 50 percent pain reduction and no neurological change, versus in the control group, 55.7 percent, actually both quite good results, exceptionally good on the DRG. That was the composite responder rate.

Page 27

But if you just look at the pain scores themselves at 12 months, you'll see that it's not quite so impressive. With a single point looking back, it's only 10 percentage points of the VAS score. If you look at the diary at the mean data endpoint, it's only 5 points on the VAS score using the diary.

Then if we then look at other secondary outcome measures, one would expect satisfaction. How likely would you undergo this therapy again? I know there are other things other than just pain, but look, really, between control, not a lot of difference; not something that I think justifies that big composite pain difference. And then if you looked at other things as well, there's not that much difference.

Then the other problem with the DRG it is actually reported in the write-up, but it's never by their speakers. They never mention this, the adverse events. I always get told by their speakers, no, there were no differences in adverse events between the two groups. And we all know in

Page 28

Europe that's not going to be true. And in fact, even in the ACCURATE study, there is a big difference, statistical difference in adverse events that were related to the implant procedure.

Why is all this important? Let's be realistic. I'm delighted that new products come to the market. Nobody minds that. Well, I don't mind it. And that's the regulators are interested in. Is it safe? Does it do what it says on the tin? But remember, they're always going to be funded by the sponsor. They're often start-ups. They live and die by the study. They're always noninferiority designed. Most of them have been unblinded. I think we're going to have a new one coming which is blinded.

They come with massive expectation bias. Possibly there's observation bias. I commend you to read the SSEDs, summary of safety and effectiveness data. And of course, remember when you go to meetings, the messages one hear are not necessarily scientific, but they are marketing messages. And I think we haven't talked about this

Page 29

today, but Eric Buchser and Sam, we've often quoted the Flacco [ph] thing on randomized studies, and the literature of head-to-head RCTs is dominated by the industry.

Industry sponsored, comparative assessments systematically yield favorable results for the sponsors, even more so when noninferiority designs are involved. And I think if you're a noninferiority design and you're industry sponsored, there's a 97 percent chance that your study will show favorability or at least noninferiority.

Study execution should include transparent methods to reduce expectation and observer bias. We know that, but how do we actually implement that and give guidance? The role of the clinical research and sponsor teams must be documented and managed by the trial management group with independent members. Will a pain score always be the primary outcome? Depending on the question. There are lots of ways of measuring the pain score, and I think we've got to decide which one. Thank

Page 30

you very much.

(Applause.)

So I finished exactly on time, which is pretty good for a blabber mouth. I'd like to now welcome to the stage Jennifer Gewandter I'm going to say.

DR. GEWANDTER: Yes.

DR. THOMSON: Okay. She's going to talk on data analysis, interpretations, and reporting. So hopefully it will be a good follow-on. Thanks, Jennifer.

Presentation - Jennifer Gewandter

DR. GEWANDTER: Good morning. Thanks to Bob and Dennis for inviting me to talk today. With my talk, I'm going to try to talk a little bit about -- follow-up what Nate said about RCTs being the gold standard of evidence, depending a lot on how they're done and how they're reported to the consumer or the reader.

We've talked a lot about different things that can affect the validity of trial results, so I'm going to try to talk about things that are a

Page 31

little bit different from what we've covered already today. Just as disclosure, I'm not a statistician, so I'll try to answer all your questions but might not be able to answer everything.

It really depends on a lot of different things. These are the types of things I'm going to cover today. I'm going to talk a little bit about minimizing type 1 error or the false positive rate by prespecifying and limiting multiple testing. I'm going to talk about designing trials from the perspective of an estimand, which is a relatively new concept in clinical trial design.

I'm going to talk about clinical meaningfulness, the difference between within patient and between group and what that means for the design and interpretation of trials, and looking at the confidence intervals to inform interpretation of non-significant superiority trials. So it's pretty similar concepts to designing and interpreting noninferiority trials, but how we would apply them to superiority trials.

Page 32

I'm going to go through this quickly because I think a lot of you probably know a lot of this already, that we want to prespecify as much as possible, we want to be specific, and we want to keep it to a minimum. Multiple statistical test can inflate type 1 error. If you have an alpha of 0.05, that means that your false positive rate is about -- oh, sorry, that should be 5 percent; sorry, 5 percent.

If you do 8 different tests at an alpha of point 0.05, the potential false positive rate is as high as potentially 40 percent. So this is what we call the family-wise type 1 error. This is really most important to think about when we're defining our primary analysis, but it's also important for key secondary analyses.

This is just an example of a lot of the different things you have to think about. You have to think about what is the primary outcome measure. You have to be really specific about that. You can't just say pain. You have to say pain with a diary. You have to say what instructions you're

Page 33

going to give the patients. You have to decide what's your primary time point is going to be; what statistical tests are you going to use; and what's the model; what are the different factors you're going to put in the model; what should the population be for the analyses, are you going to include all randomized subjects or just the ones who finished; and then what method are you going to use to accommodate missing data?

So all these things should be specified upfront so that at the end of the day you can't make a few little changes, and cherry-pick, and report what you find to be the "positive" quote/unquote or p less than 0.5 analysis.

The other thing I would say is we all like to collect a ton of stuff for RCTs, and that's really great. We want to get as much data as we can from the patient's time. But it's also important to prespecify just a few secondary analyses so that the results of those analyses are actually more reliable, and you don't again do 20 secondary analyses and just pick the few that were

Page 34

positive and support your hypothesis.

I just want to draw your attention to this manuscript. This was led by Dennis on different ways you can adjust for multiplicity. If you really can't choose just one for your primary, you can do things like making multiple co-primary endpoints where you split the alpha. Both of the primary endpoints, if let's say it was pain and physical function, have to reach significance of 0.025 for the trial to be considered positive.

In this case, unfortunately, if only one of the analyses doesn't reach 0.05, your trial would lose and not be considered evidence for the treatment. But then again, if you do get both, you can claim that it does both in your primary analysis.

The other thing you can do is something that we call hierarchal gatekeeping approach. You don't have to adjust alpha, which is great, but you do have to prospectively decide which is more important to you. For instance, let's say we decide that pain is the most important thing, that

Page 35

would be prospectively put as the first outcome, and if that hits at .05, you could then go to do your second outcome.

In that case, as long as pain hits at 0.5, the trial can be considered a success no matter what happens with physical function. But you have to be really sure that you want pain to be your most important because if physical function hits and pain doesn't, you can't call the trial a success.

Then there are other things to think about for the secondary analyses. In general, you want to think about limiting the family-wise or overall type 1 error of the trial. One way to do that is to prospectively decide how much more alpha am I okay with or false positive rate am I okay with for the whole trial?

Let's say you decide that's 10 percent, then your 0.05 would be left for your secondary outcomes, and then you could split that between those secondary outcomes using things like Bonferroni correction or other related step-wise

Page 36

procedures that are a little bit less conservative.

Again, I'm not expecting you to remember all of that. If you are interested, you can read this paper.

The next thing I'm going to talk about is estimands and how we can you use estimands to better design our trials and also interpret what the actual effect estimate really means. From a historical perspective, RCTs would have an objective. You would say, I want to estimate the effect of the treatment compared to the placebo. It's very general.

Then conventionally what we would do is we would design a trial in a specific population. It would have an active and placebo group. We would pick a method to accommodate missing data. We wouldn't really think about what exactly that means for the resulting estimate. Even now, but definitely up until fairly recently, generally that would be things like an LOCF or a BOCF analysis, where you carry forward the last observation or the first observation.

Page 37

More recently after the NRC report, that would be things like doing multiple imputation, some more sophisticated methods; but again, just kind of picking them off the shelf because they are quote/unquote "the better thing to do" without really thinking about what does that mean for my effect estimate.

Because of this, we would decide on the fly, or after the fact, what would we do with intercurrent events? And what I mean by intercurrent events are things like rescue medication usage or maybe even use of disallowed medications. A new push by statisticians is to kind of think about this a little bit differently by using the thing called estimands.

This is a definition from this reference, which is really helpful if you want to learn more about this, about all different things that the estimand includes, things that we already think about like the population of interest; what's our endpoint variable; and what kind of summary are we going to use or statistic for our data. But really

Page 38

the important thing that's different about this way of thinking is that we specify how intercurrent events are reflected in the scientific question of interest. And I think the easiest way to think about this is to just look at some different examples of estimands.

So estimand 1 is we're trying to estimate the effect of being randomized to the active treatment compared to placebo, regardless of whether intercurrent events occur. This is what we call a pure ITT estimand. This is really appropriate when your goal is effectiveness.

One thing that's important to note about this estimand is that it's really impossible actually if you have a lot of dropout that you can't follow up because you can't really impute people's data for what actually happened to them because you really don't know what actually happened to them based on the other people in the trial, which is generally how we impute data in the more sophisticated methods.

Estimand 2 would be the effect of the

Page 39

treatment compared to placebo that would have been obtained if all participants tolerated and complied with the treatment and protocol. This is what I call the efficacy of estimand. This is really for efficacy instead of effectiveness, and it's really assuming that everyone can take the drug or, sorry, use the device, and they are going to do it exactly how you ask them to.

Estimand 3 is the effect of the treatment that is actually attributable to the randomized treatment. This seems very similar to estimand 2, but the difference is that, for example, if someone drops out early for an AE, you're not going to give them credit that the drug or the device worked for them because it's really not effectively attributable to the actual drug because that person was not able to take it anymore.

So these are three different things that you're estimating, and how you handle missing data, either data that you can't find or data that happens after these intercurrent events will affect whether you're asking the question in the vein of

Page 40

estimand 1, 2, or 3.

I want to define intercurrent events for finishing this conversation. There are two types. One, the data are potentially available, and one, they're not available. Data are potentially available for things like someone took an allowed rescue medication or someone took a forbidden medication that they weren't supposed to take. The participant wants to stop using the therapy, but they're willing to provide primary endpoint data anyway. They'll come back.

With these type of data, you have two questions. One, should data after these intercurrent events be included in my trial, and if not, how should they be imputed? For data that are not available, this is things like participants can no longer be contacted, completely lost to follow up, and you have no idea what's going on with them; or the participant decides to withdraw from the study, they're unwilling to be contacted further, and they don't want to hear from you again. Then you might know, hopefully, you've done a good job

Page 41

and you know why they're dropping out. It could be they didn't get efficacy from the treatment, so it's not worth showing up anymore. They don't like the treatment. They have AEs or maybe just completely random, they moved away and they can't come anymore.

For this type of data, intercurrent events, you don't have a choice. All you can choose is how should I impute these data and should it be different depending on the reason that the data are missing? These are your two questions now.

So let's talk again. Let's bring it back to what the estimands are. For the first estimand, you want to follow participants and use their observed data whenever you possibly can because that's really the only way you can actually calculate this estimand.

For estimand 2, you really don't actually need to follow up patients after intercurrent events, at least for the primary analysis, because you're not going to use their data anyway. You're going to impute their data after they're observed

Page 42

and only use their observed data. And it's the same for estimand 3. It's not necessary to follow them up because you won't use those data anyway. And then the question with the estimand 2 and estimand 3 becomes how do you actually accommodate missing data so that you are answering the question in the vein of estimand 2 or estimand 3.

This slide is really busy -- sorry -- and it's kind of complicated. I'm not sure how much you can actually take away from this in a couple minutes, but I just wanted to introduce these terms, and you can learn more about them if you're interested.

When choosing what method to use for accommodating missing data, we think about the assumptions regarding the pattern of missingness. And formally what that means is the probability that the values are missing given the values of the outcomes, either observed or unobserved, and the statistical model.

For missing completely at random, this probability does not depend at all on any outcome.

Page 43

This is something like the person was in a bus crash on the way to the appointment, so they didn't give you their endpoint. It has nothing to do with how they're doing on the drug or whether they had an adverse event at all. The probability of being missing is completely random.

Then there's a term called "missing at random," which I will give you is misleading because this probability depends on observed outcomes but not unobserved outcomes. An example would be at the final visit where you were able to observe them, they report that they're not experiencing great pain relief, and they decide before the next visit that the study isn't worth their time because their pain relief is just not good and they just don't come back.

So you knew at their last study visit that they weren't getting great pain relief, and that is why they decided not to come back. We call that missing at random because you have an inkling of why their data are missing and the probability of their data being missing.

Page 44

For missing not at random, the probability of their data being missing depends on unobserved outcomes. An example would be you have a flexible-dose trial. The participant's last visit, they're getting only mild relief but they're reporting no AEs, and you decide to up their dose. They go away. Actually, their pain spikes, so they decide they don't want to come back, and they never contact you. You have no idea that their pain spiked because before they were getting mild pain relief. And in fact you might assume that maybe they had an AE because you upped their dose. So the reason that their data are missing, you don't have any idea. It's dependent on unobserved outcomes.

Those are the three different assumptions that we make when we make models to accommodate missing data. I'm going to talk about two of those and how we might think about what we choose to accommodate missing data based on the estimand we want.

For estimand 2, again, let me just remind

Page 45

you that the efficacy estimand, you're interested in knowing how the treatment actually works' if everyone can take it and everyone can tolerate it. What I'm showing here is these black boxes, which is the average trajectory for the group that's taking placebo, and this is the average trajectory for the group that's taking active.

These purple dots are one patient or one participant, and the first two of these dots are observed. You know those data. You got them from the patient. This is the pain score. If you impute their data using missing at random, what that means is you base the imputation of these new data on the trajectory in the active group because this participant is in the active group, and they come in slightly higher than the average because they started out slightly higher. You use their baseline data as part of that model.

This is how this patient's data would be imputed with some uncertainty, which is a really important point. You don't want to just impute a single point because that can inflate type 1 error.

Page 46

This is what their data would look like. But if we have the exact same scenario -- oh, and by the way, this person dropped out because they had an AE.

If we have this same exact scenario but we're interested in estimand 3, we want to know what the effect of the treatment is only if they can actually take it, then we would do something maybe called jump to reference, which is a missing not at random assumption.

In this case, again, we have the same exact observed values. The person drops out for an AE, and now instead of imputing their data putting in the model the average from the active group, we use the average from the placebo group, or the reference group, to impute their data. And again, their data jumps up here, and it's a little bit higher than the average placebo person because they started out a little bit higher, and this is how we impute their data.

We've decided a different method to impute their data with a different assumption because of the estimand that we are trying to estimate. So

Page 47

now it makes the decision of whether we want to use missing at random, multiple imputation, something like jump to reference easier because we're deciding that based on what question do we really want to know.

Not only does it make it easier for us when we're designing, it makes it easier for us when we're reporting our trials because we can explain to readers better what actually our estimate means. So I would argue that when you report your trials, you should be upfront and say this is the estimand and this is what our data are estimating. And therefore, we accommodated missing data using this method. So that's all I'm going to say about estimands.

DR. FIORE: May I ask a question [inaudible - off mic].

DR. GEWANDTER: Sure.

DR. FIORE: Greg Fiore. A question is about who makes those determinations of what might have driven the patient to drop out. Is that a statistician who's making that, typically?

Page 48

DR. GEWANDTER: No. Actually, this is a very important point. These information obviously are only available for people who told you why they're dropping out. So unfortunately, people who are lost to follow up for no reason and they don't give you any reason, in the jump-to-reference scenario, estimand 3, you can't put them in the AE group because you just don't know.

ACTTION is actually working on a demonstration of this using some data from the FDA database. And what we ended up doing was if it was recorded as an AE, we put them in the jump-to-reference group, and if it wasn't recorded at all, unfortunately, they had to go in the missing at random imputation.

So I think the moral to that story is, as well as you possibly can collect reasons for missing data, the better off you'll be later when you're trying to impute your data.

Next, I'm going to talk a little bit about clinical meaningfulness and the difference between inpatient and between group, which is challenging,

Page 49

to say the least. This comes a little bit off of what Dr. Thomson was saying.

We think a lot in the pain field. We say, oh, a 30 percent or 50 percent difference is clinically meaningful. Where did that come from? This was I think one of the first studies that did this. It's the one that I am most familiar with, so I chose it to demonstrate to you.

This study looked at a bunch of different trials and compared the changes in the NRS pain scores to a PGIC where patients told you how much improved they felt. And what they found was that -- know also that these are a lot of different conditions, so it's not just one condition, so it's pretty generalizable results.

They found that in about 30 percent, or even less, people said they were minimally improved. So that's where this minimally clinically important difference from baseline for within patient comes from. Then if you go to about 50 percent, you get people who are saying they're much improved or very much improved. So that's where these numbers come

Page 50

from, and that can be very useful for defining treatment responders. You can say that X percent of patients responded to treatment.

Now just as a caveat, my statistician would say you don't really know they're responding to treatment. They can be regressing to the mean and doing all these things, but regardless, that's, if you want to do a responder analysis, where those numbers might come from.

Then there's this idea of a difference between groups. A lot of times people don't want to do a responder analysis for the very reason I just said or also because a dichotomous analysis has lower power. So we often are interested in, well, I would like to do a continuous analysis and get more power.

Let's say I start at 6.5 or at the average participant, the people in the placebo group go down to 3.5 and the people in the treatment group go down to 3. So everyone's had a clinically meaningful difference, or not everyone; like the average person in both groups has had a clinically

Page 51

meaningful difference for themselves.

So the question is now, let's say this is statistically significant, is this meaningful? It's hard to decide that because you have a large placebo effect, but there is a real effect of this treatment. It's statistically significant, but is it big enough? So I would argue that it's different depending on your perspective. So the NCD will be different depending on the risk associated with the treatment.

If the risk is small, then maybe we don't need to see as big of a difference between the two groups as we would otherwise, especially because there's a lot of variability. These are averages. Also, it will be different depending on the perspective of the interpreter.

If I'm a drug or device developer and its early stages, and I have a small difference that is statistically significant, which may be because you're really good at picking a super homogenous population, you might be skeptical to take it forward because you know that once your population

Page 52

becomes more heterogeneous, your difference is going to get even smaller, and it might be hard to show a difference in a trial.

If you're a policymaker, let's say you're someone who is writing treatment guidelines, or you're a payer deciding whether you should pay for this drug, you might want to see a bigger difference, and you might want to see a bigger bang for your buck. We don't know.

But if you're a clinician or a patient who has tried everything else, nothing works for you, and this treatment has low risks, you might look at this be like, hmm, this looks pretty good to me. I don't really care how much more of a benefit the people in the active group got over the placebo group; I want to try this.

So I would argue that -- I know it's contentious; I can see people going like this. I would argue that this takes a little bit more thought than slap 20 percent change on it and go with that, especially when you're powering a trial, because at the end of the day, whether we like it

Page 53

or not, we all know we're supposed to power trials on what's important. At the end of the day, you might not see a significant difference, so are we doing anyone any favors by underpowering our trials?

I know you guys are device people. I'm mostly thinking about drugs, so maybe these large trials are not feasible. I'm just laying it out there. You can take what you want from it.

I thought about you guys, and you might want a bigger difference for things that are permanently implantable. As a patient, I'm going to be like, "Oh, I'll try pregabalin for 6 weeks and see what happens; whatever, no big deal." But you're going to undergo surgery, and you're going to have this thing permanently in your body, so you might want a little better chance that it's going to actually work well for you or there's a bigger difference between placebo and active when this type of treatment is being used. And Bob has written a lot about this. If you're interested, you can take it up with him and read these.

Page 54

The last thing I'm going to talk about is confidence intervals to inform interpretation of non-significant superiority trials. I think we've thought about this a lot because we've been talking about noninferiority trials. Oh, sorry. Did I say noninferiority? I think I did. I meant superiority.

All I'm really going to say about this is if you decide a superiority trial and at the end of the day, you don't get p less than 0.05, you cannot say that things are similar; you just can't. You can use the confidence intervals, though, to comment in the discussion, not in the results, about how likely it is that these data actually support a true negative, meaning one treatment is no better than the other, or the data are inconclusive.

So I would just say, let's say you've decided what does MCD is, and we're not going to debate that anymore, so this dot is the effect estimate, so the average mean between groups or whatever statistic you're using, and these small

Page 55

bars are the confidence intervals.

For this top one, I would argue that even if you didn't plan this study as a noninferiority study, it gives fairly good support to the fact that these two treatments, whether it's placebo or two active treatments, are fairly similar in activity, because the reason you got p greater than 0.05 isn't because there was just a ton of variability and it's all over the place, which is this scenario, where you really can't rule out a treatment that's in favor of active or of control.

I'm just giving a plug for if you do a superiority trial or you're reading a superiority trial where the p value is greater than 0.05, don't just assume that means they're the same. You really have to do some due diligence.

I didn't put the data here in the interest of time, but we did a review looking at how authors interpret these confidence intervals, and they're very rarely actually interpreted by authors, at least in 2015, so you as consumers of these papers need to really think about it yourself.

Page 56

My conclusions, to ensure the RCTs provide the gold standard of evidence, investigators, authors, and readers must pay attention to many trial design details. The topics in this presentation represent only a few important aspects of RCTs that we have to consider when we're designing and interpreting our trials. Paying attention to these details will increase the reliability of our results, and thus their acceptance by important stakeholders, including regulators, policymakers, and payers. That's all I have. Thanks.

(Applause.)

DR. THOMSON: Okay. We're doing a fabulous job keeping to time, so our New Yorker is going to hustle us through.

Presentation - Brian Kopell

DR. KOPELL: First of all, again, thank you for inviting me. The last day and a half has been just really terrific in terms of the quality of the back and forth discussions. I really love the opportunity to interact with people on sort of a

Page 57

more personal intimate manner as opposed to some of these bigger meetings, where you don't really get to have an in-depth conversation, a nuanced conversation.

As everybody surmised from my discussions yesterday, my interest in this particular realm in terms of what this body could do to recommend for trial design centers around economic outcomes and cost effective analyses, mainly because I think that when we begin to see how difficult it is to absolutely determine efficacy in a convincing fashion and so forth and plus the cost of these devices, we really are going to run into a situation where we're not going to be able to provide this really life-changing therapy for our patients, and that's going to be a real shame.

Rod is going to give an update on this. It's probably better than I will be since two of the papers that I'm presenting are actually his. And not surprisingly, most of the cost effectiveness data comes from our European colleagues. We don't really in this country do a

Page 58

very good job at looking at cost effectiveness. It is partially a cultural thing, but it's not something that we have the luxury of ignoring any further.

These are my disclosures.

Probably in his greatest book, Tom Wolfe, "The Right Stuff," attributes a quote to Gus Grissom, one of the Mercury 7 astronauts. And Gus was basically remarking that without any sort of real funding, as cool of a project might be, you can't do it unless you have the money. He said very eloquently, "No bucks, no Buck Rogers." And I think that when it comes to neuromodulation and neurostimulation technologies, which are undoubtedly cool, undoubtedly have incredible potential for our patients, it costs a lot of money, both on the development side and the deployment side. And we don't show the milieu the bucks in this, we won't have our Buck Rogers.

I'm going to start by just doing a quick review of the literature and then give some thoughts about economics in trial design not to

Page 59

prescribe what I think that we should recommend per se, but maybe hopefully stir some discussion for our discussion period in just a little bit.

Review of the literature. What's interesting is that the average U.S. patient changes health coverage every approximately three years. Most of the measures in the United States studies basically are measures that involve break-even points beyond the three-year mark, and that may not be very attractive to payers because they want to see the payoff right away. So I just want to kind of put that in your perspective.

Again, in my opinion, probably the best measurement is what we've already referenced, the so-called quality measurement, the quality-adjusted life year. And most of you probably know this, but whoever doesn't, the concept is a year where somebody is in essentially perfect health is considered a quality life year. Then you can begin to take a look at a treatment and determine its costs effectiveness by determining how much does it cost for one year of this perfect life.

Page 60

If you look at the literature -- and this was a really good review of this from the New England Journal a few years ago -- there's this rough sort of agreement across the board in various different fields that about $50,000 per quality-life year is a very reasonable number to attribute to any sort of therapy, whether it be a pill, whether it be a device, 50 grand for every quality-adjusted life year. That's probably the best way that we could probably dive into this.

There are other ways, obviously, we could probably show this: reduction in physician visits; reduction in hospitalization and ER utilization. Obviously, this kind of goes into it. If you're having a perfect quality life year, you're not going to the doctor. That makes sense.

Perhaps the lowest hanging fruit is the reduction of medications. This is definitely something that we see for DBS for movement disorders, that the number one reason why DBS for movement disorders is absolutely cost effective is that we can reduce the meds. It's very simple, and

Page 61

it's probably going to be the same in this realm as well. I don't have to tell this audience that drugs are expensive. Many patients are on 5, 6 drugs per month. That's a lot of money over the course of a year.

This is probably the lowest hanging fruit where we could potentially show the impact of where spinal cord stimulation can be cost effective. And remember, if we all believe that spinal cord stimulation has a large effect on our patients, we should be able to reduce their medications. We should be able to do that.

Just going through some of the studies in a chronological order, in 2002, this was a study looking at spinal cord stim, and this is basically demonstrating cost effectiveness breakout point at 5 years post-implant. Now sure, that's great, but that determines whether -- or that's predicated on whether a patient stays within one payer for 5 years. If they have commercial insurance, Aetna might kind of go, "Yeah, that's okay. I guess," so just consider that.

Page 62

Rod, this is one of your studies from, I believe, 2004, Spinal Cord Stimulation Cost Effective Within Three Years Across the Board for Failed Back Syndrome, Angina, and CRPS. What is interesting is most of the cost effectiveness data are for indications that are not available in the United States, basically, but there is some for CRPS and failed back syndrome. And again, the reduction is in, not surprisingly, cost of drugs, physician visits, and hospitalization. Here, this study found that for the CRPS indication, it's about $22,500 per quality, well under the 50,000.

Richard, your paper here demonstrated the difference between crossover from spinal cord stim to surgery, and obviously crossing over to surgery adds costs, and if we can prevent that sort of situation, by definition we're going to reduce costs.

Again, in 2010, Rod, another one of your papers, An Advantage of Spinal Cord Stimulation Over Conventional Medical Management to the tune of about 3500 pounds for CRPS. And overall, spinal

Page 63

cord stimulation is cost effective at $30,000 per quality. What's also interesting is that the rechargeable IPG tends to be more cost effective than a primary cell that last 4 years or less. So there's another example of how some of the specific technologies can become cost effective.

Here's one that I found that basically demonstrated that spinal cord stimulation is not in fact cost effective. And it's very interesting. It's mostly for failed back surgery syndrome for workers compensation. Failed back surgery may be the most difficult patient population to demonstrate this cost effectiveness. And ironically enough, that's probably the most indication in the U.S.

I think that this to some degree touches on what you said yesterday, which is when you're sent a set of records, you can almost kind of read the story and kind of determine this is not a good candidate for this particular type of procedure. And in this particular group of workers comp patients, spinal cord stim was by far more

Page 64

expensive treatment.

What was also interesting across the board for this patient population, whether it be spinal cord stim, conventional management, optimize management, only 10 percent achieved any significant pain relief, so that says something about the patient population, doesn't it? That's all it is.

In 2013, this was another study looking at various spinal cord stimulation patients in Canada across the board for failed back surgery, CRPS, angina, and PVD, and again, throughout this, once again, very cost effective treatment.

In 2017, my colleague Ash Sharan did a meta-analysis of 21 studies looking at cost effectiveness of spinal cord stimulation for back pain. What's interesting in this study is the large majority of the spinal cord stimulation studies demonstrated cost effectiveness. So it really does beg the question why can't we do this in a way that really makes our payers compelled to pay for this?

Page 65

As I said, there's this increasing resistance from third-party payers despite FDA approval of these devices. Year after year when I go to these meetings, when I sit on boards of societies, one of the biggest things that come up over and over and over again is the fact that we're getting such pushback from payers basically for these therapies.

A hallmark of this -- I alluded to this yesterday -- is in 2005, VNS for depression was approved by the FDA. It was approved. And then in 2007, CMS basically said it was not necessary and won't pay for it. The result of that is that nobody pays for it or very, very few pays for it. So I would just submit to you, what's the point of having a trial that gets regulatory approval that we can't get into our patients? It's literally the most Sisyphean, basically, tasks that we could possibly do. We'll just roll the rock up the hill for no darn good reason. So we have to basically demonstrate that in addition to reducing pain, we are unburdening the system of this huge economic

Page 66

cost of pain.

Another example, DRG-stim, which has come on the market, doesn't even have its own code. We have to kind of call it spinal cord stimulation. It is sort of spinal cord stimulation. I guess in some sort of philosophical discussion, you could probably determine that. So many have deemed investigational. What good is it if we can't get it into our patients?

How many here are still from industry? Raise your hand.

(Hands raised.)

DR. KOPELL: Okay. How many of you are aware that under the current Medicare payment scheme, not one hospital can actually break even on your devices? Not one company offers a spinal cord stimulation that basically Medicare will give a lump sum and that they can at least break even on? Everyone loses money. How many of you are aware of that?

(No response.)

DR. KOPELL: None of you. Obviously, none

Page 67

of your companies would be in business if you can't make a profit. We want you to make a profit because we want you to continue to develop these things. But in the same way, the hospitals are not going to be able to continue this way. It's simply impossible. It's just simply impossible. Not my law; law of the universe, basically.

So we have to address this. If there is high upfront cost, there has to be a payoff for this system. Maybe the hospitals won't see these patients in ER down the line for the next several years. Then all of a sudden, it makes sense for a hospital to have at least that initial investment into these patients. Otherwise, I don't see how they're going to be able to do it.

The last point has been touched and probably will continue to be touched on. Most of the pivotal trials are essentially 510(k)s; they're not PMAs. So that's the lowest hanging fruit of regulatory approval in the U.S. And then on top of that, they're noninferiority studies.

Now again, I'm very happy to let the

Page 68

companies come to market as quickly as they can because I want to see better devices. I'm a technophile. I like to see this. I want them to thrive. But if you're going to have that low-hanging bar to get into the market, you have to do something for our patients and us to show that your devices aren't just continually burdening the system. That's not going to work.

As I said, what's the rationale for third-party payers, or even CMS for that matter, to pay for treatments that reach market merely by superseding a noninferiority threshold and don't demonstrate any incremental, true economic benefit? What's the point of it? They're there for profit. Whether we like it -- we could have a debate about whether for-profit insurance is the right thing, but we do have basically a for-profit insurance milieu in this country. They're there to make money. So if you basically offer them a device that costs them more money and doesn't do more, I'm sorry, you will run into the laws of economics, and you will see what will happen eventually.

Page 69

So again, if the vast majority of the literature suggests that spinal cord stim is cost effective, then why aren't the payers tripping over themselves to pay for this? Well, probably it is because probably the quality of the data that we have is probably on the poor side. It's retrospective. It's observational. And I would argue that if our endpoints, perspectively when we are trying to come to market, include this cost savings data, not only would this facilitate reimbursement without the need for postmarket studies, I believe it would accelerate the actual growth of this field because it will become very apparent that this is the right way to go and the right thing to do for our patients.

So again, in summary, I think that one of the most important things that this body might be able to do is make a recommendation of how we can get some of this data into these pivotal trial designs so that we don't have to continually fight this uphill battle, and that's it.

Rod, you're going to do a much better job

Page 70

than me in about two seconds.

(Laughter.)

DR. THOMSON: Thanks, Brian.

DR. KOPELL: You bet.

(Applause.)

DR. THOMSON: As Brian has alluded to, we have an extra session to the program. I think Rod Taylor is going to be talking about cost effectiveness from somebody's perspective; payer perspective, obviously.

Presentation - Rod Taylor

DR. TAYLOR: Thanks, Simon. This is sort of unplanned. And I've got to say, Brian, that was a great presentation. I think everything I've said, you've captured. This was a piece of work that I think I managed to twist Bob's arm to agree that we might do. The usual thing, I'm up here getting all the glory, but the person who's done all the hard work is a guy called Rui Duarte. Rui's actually in Rwanda, of all places at the moment, so he can't be here with us, but Rui's helped me do this.

Effectively what I twisted Bob's arm to do

Page 71

is to do an updated review of the evidence for cost effectiveness for spinal cord stim with a particular focus on methodology, which is what this meeting's all about. So I'm not going to present any results. Brian's done that and gave you a flavor. But what I want to talk about is are there any particular recommendations we might want to put in our paper around not just a collection of clinical data but also the collection of economic data and how we might use that data to make economic decisions, if you like, around SCS.

So again, we've heard a little bit about this. There are actually three formal systematic reviews of SCS cost effectiveness, the most recent one by Hoelscher and colleagues. But what's interesting, including our own, is that none of these reviews have really focused on methodology. They've been really all about know, what are the results. And as Brian was saying, results are all pretty positive, but is the reason that the result is positive because of poor methodology?

So what we aimed to do in this review was to

Page 72

focus on the research methods employed. We've also collected the outcomes. We've also collected the results. And actually, we will hopefully have a separate publication on this review, but I think it's the methodological piece that Bob was particularly keen we focus on.

So again, Ewan, a huge thanks. We've piggybacked on Ewan's review, and Ewan and his team ran some additional economic search terms. If I understand it right, you basically just grafted [ph] these on the SCS generic terms. But I think what's important here is that we didn't just limit ourselves to what are called trial-based analyses because many analyses in this area are what are called model-based analyses.

Sometimes with clinicians, I get a little bit of skepticism about model based because essentially clinicians say you've taken the data from a real trial, and then you make it up, you extrapolate it, you manipulate it, you've bastardize it, and then you come up with some other figure.

Page 73

Well, I'd put it to you that's maybe a bit cynical about models. Models are very, very helpful. And one of the reasons that they're helpful is that by definition, when we do trials, our follow-up is finite in our area for maybe up to two years, typically in a randomized-controlled trial.

How long does pain lasts for, chronic pain last for? Well, the answer is it's lifetime. So we've got to take the results of clinical trials and then extrapolate them over what is the appropriate time horizon for the patient. If we were dealing with an acute disease, that would be fine, two years, but we need to think in a longer timeline.

So what we did here -- and as I say, two reviewers reviewed all the titles; it was really one. Rui's done all the hard work. I've been kind of carrying his shopping bags for him, but doing the checking. But this is just to summarize what our inclusion and exclusion criteria were, and I think nothing more to say other than that they had

Page 74

to be full economic evaluations. So studies like the BUD [ph] analysis that Brian showed initially, which was just a cost analysis, we excluded.

How do you assess quality here? This is one of the best named quality checklists you can come across. This is the CHEERS checklist. CHEERS is a group, the Consolidated Health Economic Evaluation Reporting Standards. And if you like, this is CONSORT from cost-effectiveness studies. Everybody knows CONSORT.

This is a good group of people called the ISPOR Collaboration. ISPOR is very predominant. I guess they're the IMMPACT of the health economics world. They published this guideline led by Don Husereau, but then the second author here, Mike Drummond, is sort of the grandfather of the whole area of health economics, so that's a good pedigree.

Twenty-four questions, and what this focuses on is the quality of reporting. Remember when we look at quality, we do have a challenge that what we can only do is judge what we see people writing

Page 75

in the literature. Jane and I had a bit of a discussion about this yesterday. We do need to look at quality, but effectively, we're looking at the quality of reporting.

How many cost-effectiveness studies full economic evaluations in spinal cord stim are there? And the answer is 14. At this moment in time, there are 14, and this is them. I've summarize the population, population RA, for instance. So Andres, the first line, RA's refractory angina; FBSS; CRPS; CMIC [ph], complex, reflex, pain syndrome; CLI, chronic limb ischemia; and then DN, diabetic neuropathy, so quite a few different indications here, but you can see by far and away, the most evaluation has been in FBSS, and of course that is an indication recommended on both sides of the pond.

Comparisons are interesting, aren't they, that you can have anything, believe it or not, from -- I think as I mentioned yesterday, in refractory angina comparing spinal cord stim with coronary artery bypass grafting, Simon talked

Page 76

yesterday about percutaneous myocardial revascularization. But again, a lot of these are, again, what we call CMM or conventional medical management.

I just wanted to focus on the fact that you can see that although some of these analyses are trial based -- in other words, the time horizon of the analysis is exactly the same as the trial -- a number of them are model based. But just to pick up on Brian's point, I thought I might, particularly for the model-based analysis, look at where their primary source of effectiveness or efficacy data came from. And I think just picking up on Brian's comment, how useful is it to have a cost-effectiveness model where the data is effectively predicated on non-randomized literature? I would perhaps say that it's going to be less helpful, and that will be one of the comments I'll come back to.

This is a death-by [ph] slide, so this is just to show you that Rui really did do some very hard work here. These are all the studies. This

Page 77

is the first 8 or 9 questions, so these are the various questions that CHEERS asked. And then we looked to see did they fulfill that question; did they fulfill that criteria. If so, a yes; if not, a no. And I think these are all yeses, but as you go through later, some of them are nos, or some of them are actually not applicable, but 24 questions across 14 studies.

What was the headline here? If you total the CHEERS scores, as I said, normally it's 24. So the denominator here is 24, so if you fulfill all the criteria of reporting, it will be 24 out of 24. The slight wrinkle is that up to 3 questions here are not applicable. So if you're doing a trial-based analysis, one of the questions says, "Was your model an appropriate one?" Well, clearly that's not applicable to our trial.

So the denominator here for some analyses is as low as 21, if you're following me so far, and what I've therefore done is just to express how many of these studies achieved the reporting criteria. And I think what is quite

Page 78

interesting -- and I need to be careful here because a couple of these are mine, so you might say, "Well, yes, of course you said that, Professor Taylor." But actually, the quality of reporting here looks pretty good, doesn't it, by and large. Most of them are fulfilling.

A couple didn't do quite so well, so I just thought I'd pull out what the main issues were. The Andrell study didn't state the perspective. Perspective means did they look at it only from a healthcare perspective or did they take a broader societal perspective and, for instance, look at the cost of return to work. They didn't state the discounting rate.

Handling uncertainty is really, really critical, not only in clinical trials as we heard from our previous speakers, but also an economic evaluation clearly also dealing with heterogeneity, and we can do that by presenting cost-effectiveness ratios for subgroups. Then again -- and I'm guilty of this as well -- many of these analyses are funded by industry, and that's okay, but we need to

Page 79

be explicit about those conflicts and who's in there, and that hasn't always happened with at least a couple of these analyses.

A couple of maybe suggestions, Bob and others, to what this might mean if we're going to use any of this information in terms of our write-up. I think if we use the CHEERS quality of reporting checklist, I think, as you hopefully agree with me, the quality is actually generally quite high. But I would say the real big caveat here is we're looking at the quality of reporting, but that's not necessarily the appropriateness of what they've done.

So you might say, "Well, what do you mean by appropriateness, Rod?" Well, for instance, did they choose the appropriate cost? You didn't look at that. Did they make the appropriate modeling assumptions? Was their model structure appropriate? Did the model structure reflect the clinical disease process? None of those questions are tested here. This is just purely about reporting.

Page 80

Just to make the point I guess is we've got the same problem with the Cochrane risk of bias, too, haven't we? Basically, it's just looking at quality of reporting. So we shouldn't beat ourselves up, but at the same time I think we just need to be cautious that this is a caveat here.

What would I say my recommendations are? And I'll blend these in with Brian's. I think we should be really trying to encourage, wherever possible, any analyses, particularly model based, to be based on randomized-controlled trial evidence. And know, going back to Nate's point, not every randomized-controlled trial is one that we might want to talk about, so they've got to be those -- what were we calling them?

DR. DWORKIN: Level zero.

DR. TAYLOR: Level zero, our new term. That's got to be in the manuscript. Bob doesn't like that. You're going to have to work it out, guys.

Clearly, the other issue we need to be careful about is when we are doing modeling, it's

Page 81

very easy to cherry-pick, isn't it, the trial that shows the best results clinically, and then you pull that through. But should we be doing that? Well again, like any area of evidence review, we should be basing that on a systematic review and not cherry-picking.

Then I think, clearly, as I've mentioned, there are a couple of analyses where we could do better, clearly being explicit about perspective, why we're choosing certain comparators, discount rates, and the time horizon. And I think probably the biggest one that we have here is this issue about -- FBSS, for instance, is a tremendously heterogeneous population, and I think it's one of the issues for actually trial reporting as well as economic reporting, is can we start to differentiate the baseline characteristics of those individuals and even begin to power the study so that we can begin to examine whether the treatment effects are consistent across different baseline characteristics.

That was my additional tuppence worth,

Page 82

Simon, so thanks for your attention.

(Applause.)

DR. THOMSON: Thanks very much, Rod.

I think our last talk is going to be from Sam Eldabe.

Presentation - Sam Eldabe

DR. ELDABE: Thank you very much, Simon. I've been drafted in to talk to you about special issues to do with programming and the sham stimulation. It is perhaps a serendipitous choice, as I'm one of these physicians who still insist on programming his own patients in the clinic. So it is a subject I'm familiar with.

For the purpose of this, I will take you through a number of studies, and I'll show you what is being reported on programming, how it's being reported, and how much is being reported. In order to do that, I decided to look at the studies in terms of whether they are an effectiveness study that's looking at the effect of spinal cord stimulation in any pain indication, or whether it's

Page 83

a study that compares devices, or whether it's a study that investigates modes, parameters of stimulation, or is it a sham-controlled trial. We'll look at those, and we'll see how the reporting differs from one type of study to another.

We will then -- as Nate asked me yesterday -- aim to come up with a list of possibly what we may recommend in the manuscript that can be done about the issue of programming, so to make your life easier.

Let's delve a bit under the bonnet of programming. The question is why does it matter? It matters because the outcomes are dependent on programming. If you implant the device, and you don't program it, you get absolutely nothing, or theoretically you shouldn't. The Alkaisy study shows that this is not indeed the case, but there you go.

There is no consensus on what constitutes appropriate programming. There is no standard for programming SCS in a conventional way. We don't

Page 84

have a standard algorithm. We do have a partial standard algorithm for programming high frequency, we do have the same for burst, and we may have something for other modes of stimulation, high density or the subparasthesia stimulation.

It all sounds far too complex when you look at it from the outside, but actually, the parameters we're playing with are simply three parameters: frequency, pulse width, and amplitude. My premise is that if you know that these are the three parameters you're playing with, you can program any device. It does look a bit like a complicated story, but it isn't.

The other parameters that we need to look at is the position of the cathode that we are importing this current to, and as you can see here, I've shown you an x-ray, and the cathode is at T9. The other thing that is sometimes reported is the position of the lead. And the position of the lead is inferred from the position of its tip. So here, that lead is at the T8-9 disk.

Let's take a look under the bonnet of the

Page 85

device programming and see whether we end up saying, yes, there is an engine all right or whether we find something like this. Before we move on to look at how programming is happening, we'll take another look as to why it matters. I think you are a very enlightened audience. You probably know this more than I do.

This paper summarizes the placebo and the nocebo effects in neuropathic pain and is a very interesting publication. What the authors come up with is a list of the predictors of a high placebo response. I just want you to look at this bold one, which is the number of face-to-face visits predicts a high placebo response. Now, programming is a face-to-face visit. That is never mentioned and never talked about in trials.

You've seen this slide before. This is a typical pragmatic approach in reporting programming in a trial. They are programmed by a separate technician and so on and so on. This is the programmer that I started with when we started doing neurostimulation. It was a suitcase that you

Page 86

had to carry around. I don't think you need it to plug it into the wall, but I think there existed the parallel to that, that you needed to plug into the wall.

Why does it matter? I think we mentioned that it is usually delegated to industry representatives, and sometimes these industry representatives are operating in isolation from the clinical staff and in isolation from the study staff. So you'll get a rep who comes in, takes the patient, goes into a room, spends a couple of hours with them, and you have absolutely no insight into what happened there, what conversation occurred, what information was given, and that is part of your intervention in the study.

There is also the issue of the reporting of SCS failures and explants. If you look at our trials in the general, you will find that the rate of explants of SCS in the trials is zero percent. There are very, very few patients who are explanted because of failure of the therapy within a trial, whereas when you look at the real-world data,

Page 87

you'll find that for most of us, it's around 20 percent. Half of this is because of lack of efficacy.

So you do ask the question, what actually happens? What's so special about patients who are included in the trials? And the answers must be, at least partially, that because we have no insight into this interaction and the patient is not going to request the device be removed until someone tells them that we've run out of options.

If you lock them up in a room with an industry rep whose job it is to come up with another option, we may go on forever, and that's probably what happens, and that probably has an impact on the failure rates that we report in trials.

There are very few reporting of frequency of the programming across groups, and of course these frequent visits focused attention can have an impact on pain and satisfaction of the patient. The programming reporting sometimes happens in the methods, sometimes happens in the results, and we

Page 88

look at that, and nobody thinks about whether these visits have a cost attached to them, nobody actually appropriates that cost, and it does not appear in cost evaluation.

Just to give you an example, this is an extract from the Senza study, and that is what the Senza study tells you about programming in its totality. Have a read of this. This is an extract from the methods section. We're saying that for traditional SCS, subject parameters were adjusted to optimally overlap parasthesia and so on and so on, until you get to suddenly hear, you're getting results. This is no longer methodology; this is results.

When you read it, it is quite difficult to comprehend. You need a translator to understand what these figures are, and I haven't actually worked out where they are yet. What is the meaning of an average and standard deviation of the minimum and the maximum program parameters frequency, blah, blah, blah? What does that mean? I have no idea. I took it to mean that you have two groups where

Page 89

you have high values and low values, and they group the high values together and the low values together, and this is the mean of all the programming values throughout the study. It could be that, but it could be anything else as well.

This is from the study by Jose De Andres colleagues, and he gave us a very good way of programming devices. Even with conventional stimulation where we say there is no way you can produce an algorithm, he gave us an algorithm. I want to go into that in great detail because it is somewhat technical. But it goes to show that if you put your mind to it, you can produce an algorithm to program a device in conventional stimulation, and there it is. And he applied that in his study, and it worked.

He also gave us a very clear approach to reporting on programming, or a better approach to reporting on programming, in that he told us the devices were programmed in a session run by a staff physician and an industry representative; that systems were reviewed at all the study assessment

Page 90

points and if the patient reported a change in parasthesia. However, he failed to tell us how often did this report change in parasthesia happen. So we don't have an idea of what is the total number of visits that occurred to do with programming.

Now these are the effectiveness studies that I told you about earlier on. What I will do is just take you through who reports what and in which part of the manuscript. Here is the population: critical limb ischemia, complex regional pain syndrome, failed back surgery syndrome, diabetic neuropathy, and the refractory angina.

As you can see, the programming method is reported in one study, and here they give us a glimpse of what the programming methodology may be, but the rest do not, much to my shame, including one of my studies. Surgical procedure, every study describes the surgical procedure in great detail, but no study describes the programming methodology in such a detail.

The programming results. Now, this study is

Page 91

an old study by comparison, 1995, but they do a fantastic job of reporting on the programming results. They actually had a full technical publication in 2000 where they reported on every single piece of detail on the programming and how it was done, the lead position, the cathode position, everything that we've looked at.

Apart from that, there is only a report from the PROCESS study 2007 about mean and standard deviations of the values, and that is in one of the appendices of the paper. Otherwise, most tell you that we adjusted amplitude to suit patients, and that's about all you're getting. How many tell us about the frequency of the programming? None. Who's programming? None; nothing.

Conclusions. Effectiveness studies do not report on programming. There are no reports on who's programming, and the reporting is variably presented, sometimes in the methodology, sometimes in the results, and sometimes there's a mixture. This no doubt affects the quality and the reproducibility and the generalizability of these

Page 92

findings in the long term.

So when we think about it, this is a complex intervention. The outcome of this complex intervention is dependent upon a lot of factors, including the competence of the implanter, the competence of the person programming the devices, the instructions given to the patients, and their adherence to these instructions.

So despite this complexity, there are very few studies that have accounted for the potential variability of such a complex intervention. We have no idea what the impact of that variability may be. Because of the concept of the rep in a closed room and a patient, we have no insight in one, two, and three.

Nobody has actually given us a glimpse that this intervention needs to be quality controlled. The quality control on the surgical part of the intervention is fantastic and is present in every study. The quality control on this part of the intervention, the programming, which is the long-term one, which is the one that matters, is

Page 93

absent.
This is a bunch of device comparison studies that you have seen before. A lot of these you've seen before. And basically what I'm doing here is showing you who's reported on programming method, who's reported on programming in the results section, whether the frequency of programming was mentioned at all, not reported, and whether the personnel was mentioned.
You'll find that the picture here is a little bit better. In Senza, I've showed you what was reported. In ACCURATE, there is a table that gives you all the values of programming with mean standard deviations and ranges. In SUNBURST, it's not so clear. We were told that the frequency of programming is as needed.
De Andres' study I've talked about, and it does give us a very clear view of what happens. This is a smaller study that compares conventional stimulation to burst and low burst, which the authors initially called sham, but when it turned out to perform as well as burst, they called it

Page 94

something else. And what they do here is they tell us that we used a broad range of parameters, multiple programs, and they mention some ranges. That's as far as they go about telling us what actually happened.
Conclusions about device comparison studies, you have marginally better reporting. The reporting is mostly in the methodology. There is very little reporting in the results section where it is reported. It is mostly reported as mean and standard deviation and ranges of the three values that we play with. There is very little in the way of reporting of cathode position and non-report on personnel or frequency of programming.
Moving on, these are your parameter studies, and there are quite a few of them. And again, you will find that in the methodology here, we have quite a bit of reporting because it is the essence of this study. You can't do a parameter study without reporting on the methodology. However, the story becomes quite different when you get to the results section.

Page 95

So you're telling us you intend to do X, Y, and zed. In the methodology, did this actually happen? There is always a parameter that is variable that you can increase or decrease, so you might as well give us what is the average or what is the range of this parameter. And you'll find that the reporting in the results section is by no means uniform, but is much better than what we saw earlier on. The reports on the personnel remains quite poor. The reports on number of visits remains completely absent.
Conclusions about parameter studies, you have better programming method reporting. The results section reporting remains quite poor, and the commonest reporting is mean standard deviation and ranges of the values of amplitude, frequency, and pulse width. The majority do not report all the programming parameters; they report only the ones that are of interest to that particular study, so you may find there is quite a big emphasis on the frequency, whereas when it comes to amplitude and pulse width, you're left guessing as to what

Page 96

actually happened here. And none provide a report on personnel or the number of sessions.
So if we were to think about recommendations, I'll give you some potential recommendations. You may agree or you may disagree. I may end up in the position of Mrs. May in Parliament, and you may shoot me down.
(Laughter.)
DR. ELDABE: -- but I'm hoping not.
So we would recommend that in an RCT, programming is an integral part of the intervention and should be quality controlled. The personnel programming in an RCT should be provided with study-specific training; that industry does utilize best practice algorithm, and there is no reason why they don't share these with investigators. They do share them with us in clinical practice, so it shouldn't be any different in an investigation.
Best practice algorithms are currently applied available for HF10, for burst, for HD, and for subparasthesia. There are scripted programs for conventional stimulation as used by Jose De

Page 97

Andres. The investigators should make attempts to strip the programming where possible, so it shouldn't be a free for all. I recall one of the reps who used to turn up in our hospital; buy two coffees, one for him, one for the patient; lock himself with the patient in the room for 4 hours, and that was it. Once he left, our results nosedived.

Investigators also should ensure that the site staff training on the programming script does occur. Where site staff training is not feasible, industry representatives may program the study according to the study algorithm in the presence of site staff. We have to acknowledge that the idea of site staff programming is not going to be feasible across everywhere. But if you were doing a cognitive behavioral intervention, you would not actually release it in the study without quality controlling it.

The questions that we need answers to are who, where, how often, and for how long. Other recommendations about the reporting of programming,

Page 98

if you are describing a programming algorithm in the methods sections, what do we want from you? We want what do you intend to use? What's the script that you're intending to use? What ranges are you going to? Are you intending to put your cathode? Who's going to program it, and how often do you intend to do that?

If you're reporting this in the results section, we want to know what is the mean standard deviation ranges of the three values. Where did the cathode positions actually happen or where the lead positions were? Who programmed it? What was the frequency and intensity? What was the setting of the programming? I'm not sure that you can insist on an outcome in every case, but it is very helpful to know what was the outcome of that particular session.

If we move a little bit from programming to the placebo-controlled trials in spinal cord stimulation, we have a few. And this is a very interesting part of spinal cord stimulation. It's one where most of my interest lies; how much of the

Page 99

impact of this intervention is actually what we call nonspecific effects?

This you can buy from Amazon, Zeebo. It's a very good product. When I clicked on it, there's frequently both together, and you can see that you obtain this for princely sum of $32. It's very cheap.

What are the barriers to placebo-controlled trials in SCS? The placebo-controlled trials in SCS are a little bit more complex than in pharma. Why? Because patients feel paresthesias with conventional stimulations. They also carry this thing about, which is their hand-held programmer. If you want to give them sham stimulation, you're going to have to do something about this or you can take it away. But if you take it away, you have to give them a mechanism to switch off their device in case of a problem. All of these are an issue.

Patients who have a rechargeable device will need to recharge their devices, and they have a certain frequency, and if you're recruiting patients who have had spinal cord stimulation for a

Page 100

while, they will have a fair idea of how long it takes them between sessions to recharge and how long it takes for their battery to deplete.

So if they are in the sham group, they will know instantly that they're in the sham group, unless you program their battery to leak the current somehow. There are consent issues, which we're not going to talk about. Device programming, it's a question of what do you program in order to get to a placebo?

These are most of the studies that have used a placebo control in spinal cord stimulation. Here, the top three are refractory angina studies, and they are all a bit from the last 10 years. What they have done here is they have used a placebo against a parasthesia stimulation, which is quite difficult, but they seem to have achieved it.

This gentleman here programmed patients on the placebo arm to 0.1 volt, and he informed them that they may or may not feel stimulation within his study. Gaetano Lanza and his colleagues, they did something similar but in a sense very

Page 101

different. They set the device to produce stimulation for 1 hour a day; whereas here, it was point 0.1 volt for the whole day.

So these guys assumed that a very low threshold stimulation makes no difference. Here, the assumption is a very low threshold stimulation delivered for a short period of time makes no difference.

This is the largest of the three studies and is quite interesting because they came up with, again, a different placebo. Their placebo was super-threshold stimulation, so the patients would feel parasthesia, but that was delivered 1 minute in 24 hours, which is an interesting concept because if you tell the patients that you may or may not feel paresthesia, for 1 or 2 minutes a day, or 3 minutes a day, they will feel parasthesia, and that's fine. Interestingly, this study found no difference between what is high stimulation and low stimulation as they call it.

I put this one in yellow because this is a study that was done during the trial period, and

Page 102

the setting for this is much easier than if you have someone who's implanted already with a device, so these guys were not implanted yet. This is a study that we did with Professor Buchser there. This is the first one that was to be done in a rechargeable device, and what we did there was to make the device in the sham period discharge current. And because it was a crossover, you had to make the device discharge current at the same rate as it was using current in the previous session.

So how did we do that? The patient came in. Whether they were in the sham arm or whether they were in the control arm, we actually measured their perception threshold in various positions. When we measured their perception threshold, we knew how much the current utilization would be. We then rang the company who gave us how much current leakage.

So if they went into the active program, we programmed them to the active program. If they went into the sham, we programmed the current

Page 103

leakage that was based on the parameters of active programming.

Therefore, it was quite difficult for a person crossing over to discern whether they're current utilization was different or not. And Eric Buchser had the foresight to ask the company to give us a time stamp on whether the device had been switched off by the patient, because this is an issue with these. It's a question of who gets a patient programmer and when.

Here you can see that, in these, there is no comment on whether these patients got a patient programmer or not. The question is how do you manage to maintain people at 0.1 volt when they have a programmer that tells them that they can't increase the current? It stands to reason that you must take away their programmer, but the manuscript doesn't tell you that. You just have to conclude it yourself. It's the same here. We don't know anything about the patient programmer; same here.

Most of these manuscripts also rely on the perception threshold, the perception at which the

Page 104

patient starts to feel paresthesia. Because of the movement to the spinal cord, we know that to be different in different positions. So a patient's perception threshold in the supine position is at its lowest. You stand them up or sit them, the difference can be about 5 volts. Yet, most manuscripts tell you we program these patients to the perception threshold, but they never tell you which position was that perception threshold detected in.

Here, in the comments section, the manuscript tells you whether the patient was programmed in a particular position. Here for example, the perception threshold was done in the supine position. The authors tell you that we took away the patient programmer. The patients were allowed to switch off their device using the charging belt. They can do that.

Here, for example, in this study they took a very interesting approach to how to deal with a patient programmer. They gave the patients their programmer, but they gave it to them in a sealed

Page 105

envelope. And that sealed envelope, when it was opened, you knew that the was unblinded; so a very interesting approach to that.

Here you can see in the sham there are a number of comparisons, and most of the comparisons are with device switch off except for this study that we've mentioned before, where they did burst at 0.1 milliamps. Again, the position wasn't mentioned. This is an interesting study because it's the largest study that has a placebo control. Again, it's a crossover, and the placebo here is device off.

In a subparesthesia stimulation, device off is very easy to work through. In a parasthesia stimulation, if you want to convince patients that they are getting some paresthesia, then you would need to look at some of these modes.

Conclusions and what can we recommend? As you can see, sham stimulation in spinal cord stimulation studies is a variable entity. Sham is not the same thing across all studies. That is because of the fact that some patients feel

Page 106

parasthesia and others do not. In parasthesia stimulation, as we've seen, the sham can consist of very low amplitude continuous through stimulation, short duration of superior threshold stimulation, or no stimulation at all.

All of these carry and inherent risk. The risk in numbers 1 and 2 is that we don't really know that stimulating people for 1 minute at 0.1 or super threshold stimulation for 1 minute today does not have an effect on the central nervous system, and the authors in this particular refractory angina study did argue that the two groups were equivalent because their sham was not a sham. The same applies to this one. When you run no stimulation at all in a paresthesias study, you run the risk of unblinding your patients.

Your sham complexity increases with the use of rechargeable systems. Therefore, in champs stimulation, you need a rechargeable device that accounts for the risk of unblinding by virtue of the patient finding out how often are they needing to recharge their device. So in this study, we set

Page 107

the maximum as twice-daily recharging.

If you have a patient who is on the higher frequency and is recharging twice daily, they move to the sham, and the current leakage is lower. They find out they're recharging once a day, they might guess. So they had to have a frequency that allowed them to recharge twice a day.

Self-discharging IPGs need to consider the current use in the active arm, particularly in crossover studies. The patient programmer status needs to be reported. What have we done with the patient programmer? Do they have it? Did they take it home or has it been taken off? Without that, you can't have a sham.

We need to confirm adequacy of the blinding. I think most studies have done that. Most studies have asked the patient, which group do you think you're in? The studies that are using perception threshold need to report on which positions did you do your perception threshold measurement in.

With that, I think that's my last slide. Thank you very much.

Page 108

(Applause.)

DR. THOMSON: Well, thanks to all our speakers. I think we've got a break now, and then I think we've got a whole afternoon of discussion. So that will be the good bit. Remember to lean forward and announce your name. It's easy.

(Whereupon, at 9:59 a.m., a recess was taken.)

Group Discussion

DR. KATZ: If all of the speakers from this morning could come up and sit on the panel, please: Simon, Sam Eldabe, Jennifer Gewandter, and Rod, please, as well. You spoke, you qualify. Thanks, everyone. We still have more people sitting in the audience than on the panel, so it's okay.

(Laughter.)

DR. KATZ: I just counted. It's close, but I don't think we need a recount.

First, I want to just say that I thought the presentations this morning were really fabulous, were really particularly lucid and important for the goal that we're trying to accomplish. So I

Page 109

want to thank all the speakers, again, for their effort to bring these complex ideas to a place where they're lucid and understandable.

Since everybody took the trouble to fill out that survey, I thought I would show what the results are, and just take a few minutes on that. Then once we're done with that, I think it will help set up the rest of the day's conversations. Then once we're done with that, we can then go through questions and answers about this morning's presentations.

You'll recall that you filled out a survey yesterday. It was a free-hand entry. The results were really incredibly convergent. People as a group really felt similarly about most of the key issues, which I was a little surprised and gratified to see. I did my best to compile them. Of course, with a free-hand survey, let me start out by saying not everybody's handwriting is that good --

(Laughter.)

DR. KATZ: -- my own included. So I had to

Page 110

take some liberties in deciphering what people were actually trying to write. And then, of course, people used different words for the same thing and the same words for different things, as you could expect. So I had to take some liberties and understanding not only what people were writing but what they actually meant.

So you'll forgive me. This is not a precision survey, but I think it will give a rough idea of what people felt about the key issues.

DR. NORTH: How many votes for Al Gore did you find?

(Laughter.)

DR. KATZ: He's still getting votes.

The first question that I asked is what is the key scientific question that you think needs to be answered? As I mentioned yesterday, there's no point in talking about study design until you know what question you're study is trying to answer.

Again, the other comment I want to make is that I don't think we should take these votes too seriously because actually things that were

Page 111

mentioned by only one or two people I think were critically important. The difference between a 4 and a 3 is not really relevant. This is just to give people a sense of what the ideas were.

The most common idea was we need to know the efficacy of any waveform versus sham. That came through pretty clearly, virtually or equally represented were comparing one waveform to another, preferably in the context of a comparison to placebo. So those ideas came through front and center as being key scientific goals.

Then there were things that were really more focused on impact on the patient, impact on the totality of the patient over time. That came through in terms of a lot of mentions on measuring impact of pain and function; pain reduction; long-term cost effectiveness; efficacy on multiple clinical domains; pain function, quality of life was mentioned. Here, it looks like only one, but it actually came through a lot, in a lot of people's comments.

Another thing came through, which is

Page 112

predictors of benefit. Now, I took the liberty of calling a lot of different things predictors of benefit. Some people said biomarkers. Some people said phenotypes. Some people said baseline characteristics. But it all amounts to the same thing. How do we know in an individual patient, whether spinal cord stimulation will work for them? And if it does, what kind of spinal cord stimulation might work best for them based on what disease they have, what kind of person they are, whatever it is? So those were the key scientific questions. That was item 1 on this survey.

Then I asked about what study design elements do you think should be utilized in order to accomplish that goal? I'll share all these slides with everyone, too, if anybody's interested. Those will be posted with everything else, and it's all anonymized and will be posted.

Of course, it's a little bit silly to ask people for a smorgasbord of study hypotheses, and then ask for a smorgasbord of study design elements because there was no easy way for me to match one

Page 113

to the other. But I just thought I would throw out what the key study design elements were that kept on coming up.

Number 1 was sham controls. Some people recommended crossover and some people recommended parallel. But I think that there's actually meaning in those differences that I'm going to get to on my conclusion slide. So don't take this as an endorsement of crossovers is parallel. It's more about what kind of design would be best supported by a crossover study and what kind of design would be best supported by a parallel.

Double-blinding was mentioned. Standard of care controls were mentioned by a number of different people. That's what SOC means here. A number of people spoke about the need to explore the usefulness of doing a trial; what kind of trial. Should it be a sham-controlled trial? Should you try multiple waveforms? Do you need a trial at all?

So people mentioned trial in a few different a ways. There were a few interesting comments

Page 114

about the need for pragmatic designs or potentially the value of registries rather than randomized-controlled trials. So I don't want to ignore those comments.

I asked about what people thought the primary endpoint should be. Interestingly, quality of life was number one. Of course, that that would never fly from a regulatory perspective in the United States. If you want your thing to be indicated for pain, your primary endpoint has to be pain. Sorry. That's just the way the world works. But in terms of what actually interested me, it was quality of life. It was function. That came through again and again and again. Pain of course was still mentioned.

I think the reason pain wasn't mentioned more is because it was sort of taken as self-evident, and these comments were really meant to say, gee, don't just think about pain; think about these other things as well.

There were a number of comments about the usefulness of composite endpoints, if a patient has

Page 115

great pain control, but they've had terrible safety events. For example, do we really count that as a success of therapy? Or if their pain is down by 1 point on a 0 to 10 scale, but globally they don't feel like they're any better, is that really a success?

Or conversely, if someone's pain is not discernibly improved, but they've improved in their quality of life and their function and their global evaluation, and they haven't had any major safety events, should we not count that as a success of therapy?

So the limitation of focusing only on pain came through in a number of different ways, including in an interest in composite endpoints, global response and safety.

I asked people, well, what other scientific questions you think are important, and not surprisingly, I got a smorgasbord of different ideas. I'm not going to go through all these, but comparing waveforms to another came through. There was a lot of emphasis on long-term -- some people

Page 116

said I want to know more about long-term benefit. Some people said I want to know more about long-term risk. Some people said I want to know about long-term benefit and the risk, and I collapsed that all in this single line.

Cost effectiveness came through. Again, predictors of response, who gets better, who doesn't, who's going to get better on what kind of spinal cord stimulation came through.

There were a lot of comments about comparative research, but comparative what? And what I want to say is that most people thought that the key question, in terms of comparison, was comparative waveforms, and that came through in a variety of different ways. What waveform works best for this. What waveform works best for that? What waveform works best overall? However, there were a few people who were interested in comparative product information, so I just want to draw that distinction.

Durability, if it works, does it work for a long time? There were a few methodological

Page 117

comments. How do we balance these nonspecific factors that I spoke about yesterday and others have spoken about; Simon spoke about this morning. How do we improve recur of design, conduct, and monitoring? Something about opioid use. I think those were the main secondary scientific questions that came through.

I just thought I would pull out one quote, which I really liked. There were a lot of interesting quotes that I liked, but one question that somebody asked was, "Can we measure whether the extent to which the treatment brings the patients reasonably closer to their life condition prior to the onset of their pain condition?" I thought that was just a beautiful way of talking about what we're really trying to accomplish with these treatments, but that we never really ask about, at least not in a formal way.

This is my conclusion slide. When you take all these various sorts of comments that people made, it really does boil down to a few things; that the critical scientific questions were

Page 118

comparison of any waveform to sham. And if you're going to do a trial comparing a waveform to sham, and you're agnostic about which waveform you're most interested in, then obviously the non-perceived waveforms make it easier to blind a study like that. And I think Rick has been trying to push us towards that kind of a study yesterday.

Set comparison of waveforms to each other; long-term efficacy and safety and cost benefit and predictors of benefit. Those were really the four key scientific questions that everybody converged on. Then you might say, well, gee, obviously it's not going to be the same kind of study that's going to answer all these questions, so it makes things much more intelligible to think about what kind of study design would help address this question and what kind of study design would help address that question. And that really falls into two groups, and maybe I'll just focus down here.

So whether you like a crossover or whether you like a parallel design really depends upon the goals of the study. And if you're interested in

Page 119

the questions about does any waveform work compared to sham, or how do waveforms compared to each other, one could imagine a relatively short-term crossover study, potentially with multiple periods, 2, 3, 4, periods.

That's probably the most efficient way of answering that question; whereas if you're interested in this other big picture or question, long-term safety, efficacy, cost benefit, et cetera, then a long-term parallel study, where harms and benefits are carefully ascertained and quality of life and health economics issues are measured, that's probably the optimal way to answer that question.

It doesn't preclude pasting an optional crossover at the end of a long-term parallel study, but the key here is long-term follow-up. And of course the longer you have to follow a patient, the less practical a crossover design becomes. If you need two years of follow-up to answer that question, you're not going to do a crossover study. Even a year would be probably impossible or at

Page 120

least very difficult.

Once you kind of get what the key scientific questions are, and then you start marrying that up to a clinical trial design, then all of a sudden the importance of various key study design elements become much more clear, blinding, et cetera, et cetera. The only other point I'll add is that people did emphasize the need for transparency and balance of these nonspecific factors regardless which question you're trying to answer or which study design would be most appropriate to answer that question.

So that's what you guys said. I'm just distilling it down and presenting it back to you.

Does anyone have any comments or questions on that before we go into discussion of this morning's presentations?

Rick, do you have any questions or comments about that?

DR. NORTH: I think that's a very nice summary. By the way, the program says Ali Rezai. Ali has not put on weight. It's Rick pinch-hitting

Page 121

at Ali's request.
I'm a big fan of blinded, randomized-controlled trials, and until recently, we haven't had a good way to do them. I think it is really incumbent on us now that we have parasthesia waveforms, for which major benefit is claimed that we demonstrate that at last.
I think that conventional SCS, which is unblinded, can go along for the ride in a crossover study; that is it can be one of the waveforms that is tested in the usual fashion. And its comparative results, however they compare, even if inferior, allow the conventional to remain on the menu. So I think it's good for everybody.
DR. KOPELL: Richard, I'm curious, and unfortunately our American regulatory people aren't here. But to the question that you said, for years, when I was training to do spinal cord stim, it was like, gosh, if we could only have a parasthesia-free, we could finally do that randomized-controlled trial. And all of a sudden, you had a company come to market, or

Page 122

[indiscernible] to come to market, how did the FDA allow them to do a noninferiority first?
I mean, when they sat up here and they were like, "Hey, well that's not our purview." What is your purview, then? If it's not going to be that, then just be a safety body, basically.
DR. NORTH: Can we speculate?
DR. KOPELL: Yeah, sure.
DR. NORTH: Nevro is not represented here, and the FDA is no longer represented.
DR. KOPELL: Right.
DR. NORTH: And no one was talking about it at the time, nor have they since to my knowledge. But what I inferred was going on was that the FDA responded to a simple argument that you have a device that is grandfathered in and/or approved, and how can you keep us off the market if we show that we are noninferior? It's that simple. So while the FDA might have asked them, politely even, to do a sham-controlled trial, I think it would be hard for them to compel it.
DR. THOMSON: I think we're a little

Page 123

over -- I think if we can use subperception as a way of proving efficacy, everything is one thing. But I think we're a little over-obsessed with the waveform thing. And I think there's more of a need -- I'm a predictor-of-benefit man because I think we've got to make it a much more simple pathway for referral through to treatment.
I think we are -- with the trial period, which we've actually never shown to predict long-term outcome, it costs a lot of money, so it increases the overall health economic cost. And what we should be concentrating on is basically improving our selection criteria so that we are almost certain that we're going to have a responder.
But where we're held back is the economic system of healthcare delivery in the U.S. because we've already heard that it's uneconomic for most hospitals to implant devices, even with the two bites of the cherry, the trial and the implant. Organizations like Kaiser Permanente and our British NHS, which actually looks more at the cost

Page 124

of the whole healthcare system, would be quite attracted to taking cost out of the delivery of the thing.
The other problem we have is the non-believers, NSCS, the people who we need our pain colleagues to refer, they're very unclear as to what patients to refer and what patients do benefit. Then I think there is some exciting stuff going on with cytokines and other potential biomarkers, which is an area of research that I think we should be in.
DR. KATZ: Great. I'm not sure if we should depart from this topic yet or not. Any other comments about these, like what are the big-picture scientific questions and what are the kind of big frameworks of study design for how we would address those questions before we dive into more detail on this morning's presentations? Sam?
DR. ELDABE: I think you nailed it in your presentation. The central question that we need an answer to is how does our original intervention compare to sham. And it's not impossible to answer

Page 125

that question. As I've showed, some studies have done that. And without the answer to that question, we are floating in a sea of uncertainty.

That's a problem. And whatever comes after that is irrelevant because, as we've seen from the sham-control trials, if your effect size of SCS is that big, about two-thirds of it is non-specific effect; in the short term, I hasten to add. I don't know what it's like in the long term, but in my mind, unless we answer the question, how big are the non-specific effects of SCS and how long do they last for, we will never be taken seriously.

DR. NORTH: Sam, do you think that one of the reasons that no one has done a sham-controlled trial of their wonderful new parasthesia free waveform is that foreseeably, there will be a substantial placebo effect, and there's reluctance to bring that out.

DR. ELDABE: Well, you're absolutely right. An industry would not have a vested interest in doing that, but we would have a vested interest in doing that. So that's not --

Page 126

DR. KOPELL: We sort of do. Right? From a scientific, we have a vested interest, but here's the domino effect. You have a parasthesia-free device that has now a superiority labeling. FDA has given that to standard spinal cord stim. If you now say that your superior treatment is no better than placebo, the entire industry gets decimated overnight because you basically have now proven --

DR. NORTH: Conventional therefore is worse than placebo.

DR. KOPELL: That's exactly right? So then what do we do? I mean, then we're really up --

MALE VOICE: I'll stop doing spinal cord stimulation.

DR. KOPELL: Right.

DR. ELDABE: I think you're asking a very good question, and you can only answer this question if you're about two or three years away from retirement.

(Laughter.)

DR. ELDABE: I don't expect you to answer this question, but I should.

(Laughter.)

DR. KATZ: Howard Fields?

DR. FIELDS: Since I'm 11 months post-retirement, I can definitely give my unbiased opinion. I came across a paper. I've been searching around for a little science here. And this is a paper that was published in neuromodulation July of 2015. The title of the paper is Effects of Spinal Cord Stimulation on Pain Thresholds and Sensory Perceptions in Chronic Pain Patients.

Now, this is a beautiful paper. Now, I haven't seen the paper; I just have the abstract here. But what they did was they had people who were already implanted, and they tested for changes in pain threshold both in the area where there were parasthesias and in areas outside where there were parasthesias. There were consistent and significant increases in pain threshold.

Now, there's no reason why you couldn't easily do that prior to entry of a patient into a

Page 128

study. And then you could say, well, if they don't get an effect on pain threshold with spinal cord stimulation, we don't enter them into the study. The other thing you could do is you could adjust your parameters based on whatever ones gave you the greatest analgesia, and then you don't have to reprogram once the person's entered into the study.

So I would say that there is a value to actually doing science, and here you have it. So everybody in this room has convinced, and I'm persuaded, that spinal cord stimulation does produce analgesia. So it gets back to which patients benefit the most. It gets back to patient selection and how do you design a trial. And I feel like there is a way to design a trial to get rid of many of the pestering problems of reprogramming.

DR. THOMSON: So if you were able to look at the early view in your modulation, you'd see that actually we've just done a QST looking at pain pressure thresholds, and we found that there is a predictor of success with spinal cord stimulation.

Page 129

DR. FIELDS: That was in that paper, too, that it correlated with success. So you have a way. It seems to me you have a way out of the woods. You have to have a device to do the quantitative stimulation, but once you've made that investment, you have it, and it's an interesting biomarker.

DR. NORTH: Years ago, Mark Sendu's [ph] group reported using the R100, which is sort of H-reflex measure, to identify patients who are going to respond to SCS.

DR. FIELDS: But if it's a reflex, then you don't know whether the effect is motor or sensory. That's a problem.

DR. THOMSON: Right.

DR. KATZ: Just a quick housekeeping announcement. People, when you start speaking just say your name, then it's --

DR. FIELDS: Howard Fields.

DR. KATZ: Thank you, Howard.

DR. FIELDS: San Francisco.

DR. KATZ: What I didn't fully understand is

Page 130

that this meeting is being transcribed. So when people say things, the transcriptionist wants to put their name, so you can have some idea who's making what comments. So otherwise, the transcript will be completely unintelligible. So if you could say your name, that would be great.

Yes?

DR. VAN DONGEN: Could I make a comment. Robert van Dongen from Nijmegen, The Netherlands.

DR. KATZ: Don't forget to say your name.

DR. VAN DONGEN: Robert van Dongen, yes. Considering the QST data, we did a lot of QST measurements in Holland on patients for chronic pain, seeing in the outpatient clinic, and we would follow them during spinal cord stimulation. We did beforehand, during, and after. But the main problem we saw is when you decrease the medication due to the efficacy of the spinal cord stimulation, you also change your QST measurements.

So it's a dynamic process, which you don't really know what you're tweeting. So your QST measurement might be objective to prevent or to

Page 131

predict which patient might respond, but on the other hand, you change your medication during their follow-up and it also changed the QST measurement. So it's not as simple as we initially thought, and especially when they use opioids or anti-neuropathic pain drugs, then you get a different effect of the QST measurement afterwards.

DR. NORTH: Back in 1977, we tabulated reductions in medication, and it was our assumption that a reduction in medication attributable to the pain relieving effects of the stimulator was a good thing, and that a change in medication couldn't possibly confound the results of the trial. But that of course is simplistic.

DR. FIORE: Nate, a comment, if I may? Greg Fiore.

DR. KATZ: Please.

DR. FIORE: I sit and think about the validity of sham as a comparator. When the data I think is widely accepted, the sham is an effective treatment. But it's not really a treatment that is available to these subjects or to these patients.

Page 132

So why is sham such a valid comparator in that context? Because there's nothing better potentially?

DR. THOMSON: Simon Thomson. I kind of agree with you because when I hear people say we'll just have to stop doing it, I have to think of all of those patients who clearly benefited from these interventions. I don't know a better way of turning on a sham for this particular patient group, is really what I'm saying. And I think that's what you're saying.

DR. NORTH: Send them to your competitor who does a technically inadequate job --

(Laughter.)

DR. NORTH: -- and they benefit anyway.

(Laughter.)

DR. KATZ: Brian, I have a question for you. So why don't we start moving into talking about the specific presentations. And, Brian, maybe I'll start with asking you a question since you're very focused on the payers, and it seems like that's going to end up being a driver of more rigor in our

Page 133

research.
I guess my question is what type of study design do you think would be -- if that's our stakeholder group that we're focusing on, at least for the moment, what type of study design do you think would most persuade that audience that the treatment is worth paying for?
DR. KOPELL: Now, I understand that there's going to be problems with what I'm about to say, but it would seem that, like other device studies, a reduction in medication is the lowest-hanging fruit in terms of showing reductions of costs; the way it is in DBS for movement disorders, and they are absolutely depleted with dopamine. So there is no doubt that they need dopamine, and that is the way that we have demonstrated, essentially, cost effectiveness of that therapy. I think you're probably going to be forced to do the same here.
DR. KATZ: I do want to say that I've been involved in a lot of health economic studies of chronic pain populations of one kind or another: back pain, osteoarthritis, neuropathic pain. In

Page 134

the U.S., medication costs tend to be a relatively small percentage of their overall costs because things like Vicodin and Percocet and ibuprofen, and now generic gabapentin and tricyclic antidepressants, those are all generic now. So even if you add up two or three of those medications, at least from the data that I've been involved with, they tend to be relatively small.
DR. KOPELL: It's still a reduction in cost, though.
DR. KATZ: Yes.
Andrea and then John.
DR. TRESCOT: Andrea Trescot. We talked a little bit about this last night, too, and I didn't go forward with it. But it occurs to me that what you're saying is absolutely right. The medicines are the least -- if you listed all the costs of a low-back pain patient, the medicines are the least of them. So what we need to be looking at is a much more global -- the ER visits, the physical therapy visits, the return to work, the overall consumption of resources.

Page 135

The problem comes in is that right now the insurance companies are not prepared to do that. I actually was threatened with being removed from Blue Cross from an economic credentialing because they said, "My goodness. You see these patients 11 times, and your competitors see them 3." And I said, "Well, my competitors are doing procedures in the surgery center, and I'm doing them in the office. I bring the patients in for a ketorolac injection instead of them going into the emergency room. We're doing the neurodiagnostics here in the office." And the response was, "Well, those are all different pockets and different buckets, and we can't look at that."
Now, that was several years ago. I didn't get decredentialed, and that's where something like the National Health Service or Kaiser or one of the self-insured groups would be I think -- and the military is another place where we might be able to look at this in terms of DoD, where they're covering all the costs in pretty much one bucket. But to my mind, it's absolutely clear that if we

Page 136

can't show to the insurers that this saves them money instead of being a cost setter, they will not do it.
DR. KOPELL: I love the idea of being able to show reduced other utilizations of healthcare. I now think -- because again, I've had two years of experience as a chief medical officer of a public medical device company, and when I think about it, I how long would a trial have to be to demonstrate what you just said? See, the thing is, they still have to run their trials. So if all of a sudden you're saddling them with a five-year trial, that's a nonstarter, too, guys. So you've got a good baseline and that's
DR. TRESCOT: A year, if you've got a good baseline. And that's the key. You have to be able to start with patients that you can identify what their baseline costs are.
DR. KOPELL: I'm just saying that's a challenge in a relatively short-term trial. Now again, I'm not an expert in quality analysis, but can you do a quality analysis that's valid in a

Page 137

one-year study?
DR. TAYLOR: So simple answer, yes.
DR. KOPELL: Okay. So that's what I would recommend, basically.
DR. THOMSON: And indeed, I think one of the things that was striking to me -- so I'm a health economist who normally lives in a cardiovascular cave who's come out and played with some of the chronic pain area. The thing that struck me as -- and it's true in all of chronic pain, that the levels of disutility that people with chronic pain have are equivalent to end-stage cancer and NYHA4 heart failure.
Now you might say, well that's a problem for the patient. Actually, it's an opportunity. Your ability to improve quality of life because the baseline effect is so low is huge. So going back to Brian's point, one of the reasons that payers have gone here is that the quality gains you get, because of the improvements in quality of life, from randomized controlled trials with follow-up, using the EQ5D up to 12 months, are stunning.

Page 138

So you don't have to demonstrate cost savings. For the extra cost to the healthcare system, because of the qualities you gain, you're less than the magic $50,000 per quality or 20,000 to 30,000 pounds per quality in UK, and that's why you've been successful; just an observation.
I think, Nate, going back to your question, what would payers be wanting in terms of a design, I think actually the PROCESS study is not perfect, but I think was a very important exemplar to help UK say that they wanted to fund spinal cord stimulation. If we hadn't had PROCESS and we didn't do modeling, we wouldn't have got, I think, approval.
I think that the cost [indiscernible], if I may very quickly now, is that payers are beginning to get a bit suspicious, and they're hearing about this placebo problem. And because we've now got the technological ability to do the sham trial -- for instance, NICE recently looked at Nevro. The guidance will be out on a website near you very soon. I can't tell you what the guidance

Page 139

says, but what I can tell you is there's a research recommendation saying we need a placebo-controlled trial here. So they're not saying we won't cover it, but they're putting down a marker for the future.
DR. KATZ: Thank you, Rod.
DR. THOMSON: Simon Thomson. Just to add, I think derived from the PROCESS study and the economic evaluation, we did show that the sensitivity analysis, as far as what were the main drivers of cost and how to reduce them, one of them was being able to reduce, in the non-treatment group, the medication -- Oh, no, in the treated group, the medication. And the other is device longevity and initial device cost, and of course complications. I think those are the four things which warrant research, basically.
DR. KATZ: John?
DR. MARKMAN: John Markman, Rochester. I wanted to make two comments; one risky and one safe. The risky comment is risky because I'm going to disagree with Brian. I think that with regard

Page 140

to value in economic outcomes, it's hard to be in a position in a field where the measure of quality of care for many of the practitioners is multimodal care: the physical therapy visits, the cognitive behavioral therapy, the drug therapy, the integration of those with the procedures and their devices.
I think that if our position is we're going to show value by showing how we take away or reduce the need for the multimodal way of delivering care, which they're is the highest standard of care in this country, I think there's a tension there that we would have to resolve because all of our lead organizations are constantly touting the fact that the more multidisciplinary care there is, and the more of it you give in a more integrated way --
I think this is a point of contrast with Parkinsonism, as an example, like in movement disorders. If you cure someone with DYT1 from their dystonia, nobody cares about multidisciplinary care. They care about the fact that the dystonia is gone. And I think for pain

Page 141

care, there's a cultural acceptance that it's going to have to be multidisciplinary because there's an intractibility to it no matter how good your therapies are, and I think we're in a weird tension to be arguing against that.

My second point, less risky because I'm agreeing with Howard Fields --

(Laughter.)

DR. MARKMAN: -- which is always a safe place to be.

(Laughter.)

DR. MARKMAN: There are a quarter million or more cardiac pacemakers implanted in the United States. Obviously, third-degree heart block and tachy-brady syndrome are bulletproof indications for those devices. You cannot say no. If you want to put those in, it's because we have a biomarker which says this is a biomarker which matters. We have this rhythm. And if you have this rhythm, you get the device.

The closest we could come to that is a QST signature just like the one that was described,

Page 142

which would be if you have this pattern, then you get this device. And again, that would be something which we could articulate and say, this is our understanding of the nervous system in 2022. Let's be honest. We've got a new drug for transthyretin and amyloid neuropathy, which is $450,000 a year; $450,000 for one drug, for one neuropathy. And those patients don't just have neuropathy, by the way. When you have an amyloid-related neuropathy, you've got a lot of other medical problems, and the drug is over $400,000 a year.

So the idea is there is a lot of merit to really refining the neurobiological basis for why you're putting this in because it completely recalibrates the value proposition.

DR. KATZ: Just a quick plug. In terms of biomarkers, for those of you who do research, I was just at an NIH meeting -- what day is today, Friday? -- on Wednesday, on biomarkers. The new NIH HEAL Initiative, they have $500 million that they said needs to be completely committed by

Page 143

September of 2019, $500 million dollars. They outlined the major focuses of their funding plans, and one of the major focuses is biomarkers for treatments for pain.

So if anybody's interested in doing studies on biomarkers to predict efficacy of spinal cord stimulation, now's your chance.

Brian, do you want to comment on John's comments?

DR. KOPELL: I hear what you're saying, but I am a big believer in multidisciplinary care. Even doing DBS for movement disorders requires multidisciplinary care. But still multidisciplinary care, by definition, is more resourceful than streamlined or very simple care, basically, by definition. So if we use less of it, it's less of a burden.

If we have a patient -- I agree with you, a complex inpatient requires multidisciplinary care to get the best treatment. That I will never debate you on. But if that patient is the same way 10 years from now versus a patient that doesn't

Page 144

need that multidisciplinary care anymore, I would argue that patient that doesn't need it anymore, that's the success.

DR. MARKMAN: No. But I don't think this is -- in my hands, at least, this is not a technology which -- if I don't see the patient for 10 years and I put one of these in, my assumption is not that I've cured them and they're just coming back for a battery replacement in another decade. My assumption is -- I think someone said this -- they've probably gone somewhere else, because I don't see the ablation of pain. I don't see this as the treatment of DYT1. I think there's efficacy. I just don't think it's that.

DR. KOPELL: Yea, I get it. So then as a payer, if I'm a payer, and I'm a for-profit payer, I'm going to look at that and say, "Well, gee, is it really worth us to be putting in a few hundred grand into this patient if we're no better off than where we were before?" I'm sorry, but they're going to make that their argument, and they're going to win that battle.

Page 145

DR. LOESER: I don't think they will.
DR. MARKMAN: I don't think they win it. And we don't win it in an age -- the unmet need --
DR. LOESER: It's a chronic disease --
DR. MARKMAN: Yeah.
DR. LOESER: -- and chronic pain is a chronic disease. And we need to use the chronic disease model, which is not one of an episodic care, and then you're done with the patient. And I think that's an important point we need to keep in mind when we deal with those who pay for care, that this is a chronic disease, and we do not have a cure for this chronic disease.
DR. KOPELL: Okay. To that same token, any chronic disease will have a certain slope of cost for that care. If you do not change that slope in any intervention, it doesn't matter whether it's chronic or acute, the payers are going to say that's not worth it.
DR. LOESER: The slope of care cost needs to include not only the cost of the healthcare but the cost of the disability the patient may have. And

Page 146

if you are capable of restoring someone to gainful employment, you have made a huge contribution.
DR. KOPELL: Sure. I give you that, but that's hard to demonstrate in a small trial, basically.
DR. KATZ: I have a quick comment on that, on the debate between John and Brian. And then I need to go to Jane; and then I need to go to Howard; and then I need to go to Andrea. So there is a method to this madness.
In terms of the cost effectiveness, I do want to say, just again from my own experience with looking at cost data and chronic pain, and looking at where costs savings do actually come when they're assessed in patients based on, treatments for pain, when you do see cost savings, it's usually not from reducing -- mostly from reducing medication, so there's certainly a component.
It's usually not from the multidisciplinary care either. It's not because you're reducing physical therapy or acupuncture. If you do save costs, it's because of things that the patients

Page 147

don't want. It's emergency room visits. It's hospitalizations. Often those ER visits and hospitalizations can't even be tied directly to what the patient's pain syndrome is, but yet you still see them there, and they're attributable because they're accessing the pain patients compared to a control population.
So I think there is a common ground here, where if an effective treatment does reduce ER visits and hospitalizations and additional surgeries, and management of complications and things like that, then it's a win-win for everybody, especially if you combine that with what Rod said, which is that if we're improving quality of life, then the pressure on actually decreasing dollar cost is less potent.
All right. So that's my comment.
Jane, it was your turn.
MS. SHIPLEY: Thank you. Jane Shipley from Baltimore. First of all, thank you all. I learned a lot this morning, and it's not every day when I learn a lot, so I appreciate your presentations.

Page 148

One thing bothered me, however -- since we're talking about cost effectiveness and study design -- and that is that two presentations mentioned the Hollingsworth cost- effectiveness study.
I would like to point out that that was based on clinical results from Judith Turner's study. There were those of us who considered Judith Turner's study so fundamentally flawed that we suspect that it might have been a deliberate attempt at policy-based evidence making.
MALE VOICE: I think that's not acceptable to make that statement.
MS. SHIPLEY: I just said some of us think that -- I didn't say that it actually was. I certainly have --
DR. KATZ: Jane, do you want to explain why?
MS. SHIPLEY: Well, the sponsor was involved in which patients got the therapy. They presented it as if it were an RCT, and it wasn't an RCT. We've written about this. We've written a letter to Pain. We've presented on this. We've made our

Page 149

objections to the study public.
At any rate, my point I'm trying to make, that aside, is that if you're going -- the CHEERS checklist Rod talked about, and he said we can't cherry-pick, but we also have to remember that junk in is junk out. And if no one is looking at the validity of a study with clinical results upon which another study is based for cost effectiveness, and they're taking the results of the cost-effectiveness study without determining whether the clinical study could be, in fact, considered accurate, and beneficial, and something we should pay attention to, then I think that's a mistake. So I just want to put that out there.
DR. KATZ: Rod, do you want to comment on that from a methodological perspective?
DR. TAYLOR: Yeah. I'll keep it apolitical. I think there are a couple of problems with the study that at least I'm aware of the scientists, regardless of the setting of Washington State that did the study. One was just the population. Brian said that it's worker's compensation, and I think

Page 150

there are so many perverse incentives there that really biases what the outcomes might be. It's almost to the point that could we rely on the outcomes at all.
That's the big issue, Jane, isn't it? It's the patient selection. Then the other one, as you said, it was a non-randomized study. So yeah. I do cock-a-snook, as we say in Scotland. Cock-a-snook? Have you heard that?
DR. KATZ: No. What does that mean?
DR. TAYLOR: If somebody says to you, "Oh, cock-a-snook," what they're doing is they're ignoring you.
(Laughter.)
DR. TAYLOR: So I cock-a-snook to that study.
DR. KATZ: Andrea, I think you were next.
DR. TRESCOT: I think it is important to look at chronic pain the same way that we look at diabetes, in that early intervention has the potential for preventing long-term consequences. And unfortunately, we get the patient when they

Page 151

have gangrene, and then all of a sudden, it's surprising that we're not able to salvage these patients.
So perhaps looking at these stimulators not in failed back, which is -- we've already said a totally amorphous group of patients and not a diagnosis, but rather in those patients who have a more defined pathology earlier in their care, not as an end therapy but as a modulator of care, comparing them perhaps to an untreated group, which I'm not sure that is actually ethical -- but the idea of this early intervention, the fact that if we're using spinal cord status stimulation as a salvage, we are going to have to accept much smaller improvements than if we looked at it as a primary therapy.
DR. KATZ: So let's actually pivot on that topic for a moment because we have to come up with the research recommendations. We have on that slide a list of key questions that we all brought forward, but what we didn't get at in that survey is what pain syndrome would be the best one to test

Page 152

these hypotheses on. I think maybe there's an assumption because there's been a lot of discussion that lumbar radiculopathy in the setting of a failed back surgery syndrome, since it's the most common indication, might be the best place to do that, but we haven't really talked about that explicitly.
So what do people think would be the best pain syndrome to test these hypotheses on?
Simon, do you want to start us off?
DR. THOMSON: Yeah, I'll just kick off with that because we say chronic radiculopathy, as you said, in the context of failed back surgery syndrome, we're working on it, trying to redefine the phrase And again, a lot of this is all to do with the reimbursement.
In the UK, we manage to make it so that SCS is available for all patients with refractory neuropathic pain. It's not a requirement that they have to have had surgery before you can prescribe spinal cord stimulation. But I think in the U.S., it is a requirement, is it, that they have to have

Page 153

had some spinal surgery before.
DR. NORTH: Not an absolute one, but it is customary.
MALE VOICE: I don't think so, not in the state of Washington.
MALE VOICE: No, it's not.
DR. ELDABE: Can I make a comment? I think given what Rod has told us about the baseline EQ5D of the population, the baseline EQ5D of the population Rod was referring to is specific to the failed back surgery syndrome population. It is not applicable to patients with low back pain. You go to low back pain, you'll find the baseline EQ5D around 0.4. If you go CRPS, you'll find the same.
So if you want your best chance, it is with failed back surgery syndrome patients. There is nothing else that works for them. But the question is, which failed back surgery syndrome patients? It is such a heterogeneous population, we have to do a better job at defining these.
DR. NORTH: Simon -- it's Rick -- picking up on your use of the word "radiculopathy," just as I

Page 154

did yesterday with neuropathy, that word means something wrong with the nerve root. And many failed back surgery patients have had nothing more than a fusion based on degenerative findings on imaging studies and whatever other incentives might be built into the system.
But if you are careful to select patients -- and I'd go so far as to say you would want to have a spine surgeon review all the candidate's prior records to see whether there was a compelling case for their first surgery, that they had nerve root compression, even if they don't have it now, so there still is plausibly at least something residual still wrong with that nerve root or roots, that's the subset of people that we should have in this idealized study.
DR. THOMSON: Whether they've had surgery or not.
DR. NORTH: Well, I think if we limit ourselves to people who have surgery, it's a more straightforward homogeneous group; not to say that SCS is not a good choice, as an alternative to the

Page 155

first back surgery, if you will. But we do have one RCT that shows it's superior to repeat back surgery. It's very hard, by the way, to do and to repeat such a study. We tried with support of one of the companies, but couldn't enroll patients.
DR. KOPELL: So you're basically saying limit this to patients that have had lumbar surgery, that had appropriate lumbar surgery in the first place, meaning reviewed by a spine surgeon and limit just to that patient population. When patients come to you with four or five surgeries, you'd have to vet every one of those five surgeries. In other words --
DR. NORTH: You'd have to say that at least one of them was for nerve root.
DR. KOPELL: Okay, fair enough.
DR. KATZ: Howard?
DR. FIELDS: I'm going to make a strong prediction, and maybe it's even helpful, that if you were able to have a way to assess small fiber function, let's just say thermal stimulation, if you find that there's an area where the initial

Page 156

stimulation is for a given stimulus intensity reported as less painful by the patient, and then with repeated stimulation, it becomes very, very painful, even more than the normal person, I've only seen that in patients who have nerve injury.
So my guess would be, based on going back to Bishop and Landau [ph] the gate-control hypothesis, the fact that large diameter fibers have lower threshold, those are the patients that are going to get the biggest and most dramatic effect from spinal cord stimulation. So in addition to having Rick's criteria, I would say I'd like to have some objective evidence for a sensory abnormality in the distribution of their pain.
I think if you had that, that's your ideal patient, and I predict that those patients as a group will do very well. And I would also say that if you had that evidence, you could go to the payers and say, "Look, we have something. This is the ideal scientifically based treatment for this condition. Nothing else does anything like this." The problem for the manufacturers is that's going

Page 157

to be a relatively small population of patients, which is why the payers will be more willing to cover it. So it's best for the patient.

DR. NORTH: We've looked at that or tried to look at that, and for it, not as thoroughly as we might have done or as you might like. It was difficult to show any association with response to simulation.

DR. KATZ: How was that done?

DR. NORTH: But I'm with you. An objective basis for complaint of pain and an objectively demonstrable nerve abnormality would be nice to have.

DR. KATZ: Any other thoughts about what painful disorder these idealized studies should be performed in? Andrea?

DR. TRESCOT: Andrea Trescot. We were just talking about peripheral neuropathy as a treatment in that though 70 percent of them are idiopathic, you've at least got something that can be documented, and is debilitating, and potentially reversible. If what you think of the peripheral

Page 158

neuropathy as being an ischemic model, then this is a good way of potentially restoring blood flow, and therefore improving function. And for that, also looking at the peripheral vascular disease.

DR. KATZ: Anyone have any thoughts on painful peripheral polyneuropathy, painful peripheral neuropathy, ischemic peripheral vascular disease as potential patient populations for such a study?

DR. ELDABE: Sam Eldabe. I'm not sure about peripheral neuropathy. I can't really comment on that. But critical limb ischemia and ischemic diseases, there are more RCTs in that domain for SCS than any other domain. And the conclusion that you can draw from that is they're negative. You do not improve limb survival with SCS in critical limb ischemia, unfortunately, as much as we would like to believe otherwise.

DR. TRESCOT: End stage. I'm sorry, but again, we're looking at end stage. If what you look at is not limb salvage, but improvement in perfusion, if you already have a patient who's got

Page 159

gangrene, that's one thing. But if you have somebody with intermittent claudication, then you can look at walking tolerance, transcutaneous PO2 levels, healing of small ulcers. But again, once you have gangrene, if you're looking at it as an end-stage limb salvage, that would be like asking the diabetologists to now get their diabetes under control and see if that now fixes their gangrene.

DR. KATZ: Sam or Simon or someone?

DR. ELDABE: Sam Eldabe again. I think you're absolutely right, and therein you get into the problem that we have with refractory angina. You are not going to get vascular surgeons to refer you these patients at the appropriate timing. You will get them at a point where their transcutaneous PO2 is less than 10 millimeters mercury, and therein it becomes useless. This is why targeting a population that does not naturally turn up in your pain clinic poses a great difficulty in recruitment.

DR. TRESCOT: Unless you go directly to the primary care.

Page 160

DR. ELDABE: Yes, that is correct. But I suppose for refractory angina and critical limb ischemia, that's not really possible.

DR. KATZ: Sam, what does the literature on spinal cord stimulation say about efficacy in patients with ischemic claudication and not critical limb ischemia?

DR. ELDABE: I think there's a whole literature about which stage you capture these patients at, and Angela's quite right. When you do transcutaneous PO2's, you get better results, but albeit, that study is non-randomized. So if that's the only positive study, but it's the only one that's non-randomized, that may tell you something in itself. But it's the only one that's captured patients at a stage that you can look at and say, yes, these are earlier stage patients.

DR. NORTH: Importantly, I think that literature is based on conventional stimulation. I'm not aware of anything really with the new parasthesia-free waveforms, and that's what we need to use for our blinded trial.

Page 161

DR. KATZ: Eric and then John?
DR. BUCHSER: Eric Buchser. Studies on peripheral vascular disease actually showed, under certain circumstances, benefit at one year, not beyond. So that might be a problem in terms of cost effectiveness.
The other thing is that if you look objectively at change in blood flow, nobody has shown any change in blood flow except for microcirculation. So actually, the objective markers and criteria for spinal cord stimulation and PVD actually are not there.
DR. KATZ: John?
DR. MARKMAN: John Markman, Rochester. I think neuropathy is very intriguing. I don't think a mixed basket of different neuropathies is very promising at all. I think the assay sensitivity is going to be vanished [indiscernible]. If you put HIV neuropathy there, we know that there's never been a positive trial ever for the drug conducted in that group. We know that there's differential treatment response across these different

Page 162

neuropathies for different anti-neuropathic pain agents.
So there's a lot of evidence to suggest that we can't mix and match and do all neuropathy, all comers. Within specific neuropathies, I think what it obligates us to do is use a parasthesia-free system because I think historically, at least in my hands and I think in others probably as well, covering the distal feet, where oftentimes people have the most severe spontaneous neuropathic pain, is incredibly hard with a parasthesia-based system, and to do it with great reliability.
So I do think it's a unique opportunity in patients who have spontaneous pain syndrome, not evoked, but who have a lot of spontaneous pain with a parasthesia-free system in a single neuropathy, potentially, with some sort of quantitative tests; or even just to show sensory deficit as part of the inclusion, I think that would be uniquely promising.
DR. KATZ: John, you know what I'm going to ask you.

Page 163

DR. NORTH: I have no plans for dinner.
(Laughter.)
DR. KATZ: Which neuropathy?
DR. MARKMAN: I think that -- well, that's a great suggestion. Andrew just mentioned alcohol-related neuropathy. It's one that's not commonly studied in the drug world, but it is ubiquitous. And it is one which I think is a very attractive one. But the only thing that gives me pause about that is that no drug company has ever really sought to explore it, and I'm sure they've thought about it.
So the options would be -- and as you know, there are multiple neuropathy trials going on now in phase 2 in the drug world for small fiber neuropathy, which are largely based on punch biopsy, and then there are several diabetic peripheral neuropathy trials also ongoing.
My inclination is to avoid diabetic peripheral neuropathy for two reasons. Number one is many of these patients do not have a stable underlying pain pattern in my experience; they have

Page 164

a variable pain pattern. And many of them really have numbness more than pain, or by the time we see them, they have numbness. So there's a lack of stability to the underlying spontaneous pain phenotype in the diabetic neuropathy population, which I think adds an extra potential for measurement error.
Second, obviously, is that for most of us who implant on a regular basis, one of the only times you really see infections in a way that you can't minimize the risk is in patients with diabetes, especially at the pocket. I think that's the one group -- especially their HBA1c's are not that well controlled. For me, that's been a challenge in that population.
So I think that the infection risk as well as the instability of the phenotype make diabetes less attractive and push me more towards chemotherapy-induced neuropathy or a very, very esoteric -- maybe neuropathy in the setting of monoclonal gammopathy, no immunocompromise, relatively common, very easy to characterize in

Page 165

terms of laboratory studies, and not a lot of treatments, so you kind of have a niche unto yourself. So I would go for monoclonal gammopathy.

DR. KATZ: Rick?

DR. LOESER: I think that's going to be a dead end; the numbers involved and the fact that we are sitting here in the country with this immense number of failed back surgery patients who are very difficult to deal with. I think that it is possible to characterize those who have had surgery, who have a clear-cut neuropathic pain within their pain syndrome, I think that's the proper target for a study studies such as this. I think going to arcane, rare neurologic disorders is just the wrong way to go.

DR. KATZ: Brian, from a payer perspective, it seems like the goal is to advance the availability of this therapy for the most common indication, which is failed back surgery syndrome. Scientifically it may be super interesting to do peripheral neuropathy, but from a payer perspective, if that's our goal, how important is

Page 166

it actually to stick with the a syndrome that we're trying to advance the therapy for?

DR. KOPELL: I think to stick with failed back syndrome.

DR. THOMSON: Certainly in the UK, NICE recommended spinal cord stim for FBSS CRPS, but importantly indicated ischemic pain as being an experimental indication, and therefore needed to see more evidence. So I think if one of the recommendations we want to make as part of this get-together is where would one want to place your bets on the roulette table to do a placebo-controlled trial, I think what I'm hearing is that that would be FBSS. Then the issue, as Sam was saying, how do we best characterize those patients within our heterogeneous space to maximize the likelihood of a treatment effect?

DR. KOPELL: And I would go even further is to kind of recapitulate a little bit about what you did in '07, which is intervening before they get to the fourth, fifth, and sixth surgery. So maybe one single intervention, and then those patients and

Page 167

those patients only, because not only would you basically show a cost benefit, but the implication is if you can stop them from getting another fusion, boy, now you're talking real cost savings

DR. TAYLOR: Yeah, and then that's much better than medication.

DR. KATZ: Salim?

DR. HAYEK: I agree with John Markman. Small fiber neuropathy is a great target population. There's a high response rate. The problem is it's hard for the average practitioner to diagnose the condition because they have to do the biopsy, and not all centers have biopsies. I agree with him also on the diabetic neuropathy. A lot of those patients go from having a mixed neuropathy with small fiber, to becoming pure large fiber, which is a non-painful neuropathy and they have a high risk of infection and complications.

Therefore, I think failed back surgery syndrome with predominant maybe leg pain component and basically being neuropathic is the obvious target. However, I cannot emphasize how it is

Page 168

important for our study going forward to make sure that we exclude fibromyalgia, not based on history, but based on actual investigators, screening the patients with validated screens for fibromyalgia. Because if you include those patients in your study -- and a lot of the pain studies do not put that as an exclusion criteria, especially with failed back surgery patients -- we're going to see a lot of failures, and that's because these patients will complain every other day about new pain syndrome.

DR. KATZ: Tell us your rule of thumb, Salim. Tell us your rule of thumb for how you select patients for treatment. Who do you rule out?

DR. HAYEK: Beside fibromyalgia? Smokers, worker compensation, fibromyalgia, active litigation, pain in multiple areas besides fibromyalgia.

DR. KATZ: One more.

DR. HAYEK: Opioids and unemployed on disability? The list could keep going.

Page 169

DR. KOPELL: But patients that present to you with back and leg pain without any sort of surgery in the past or any sort of significant injury to them.
DR. HAYEK: For our study purpose, I would stay away from it.
DR. KOPELL: You do not.
DR. HAYEK: I don't. I only have two patients in 20 years of practice that I implanted for radicular pain with not having had previous surgery.
DR. THOMSON: Simon Thomson here. I think we're also missing out on some other things which would help us predict. I put quite a lot of store -- I mean, I hear what you're saying, everybody's signing. But also, those patients who respond well to a root block, they really do get good pain relief, and it last for a few weeks, few months, and they actually improve. But they keep getting recurrence and keep coming back, and you can't carry on forever doing it. They're, I think, always good candidates.

Page 170

DR. NORTH: All of these I think are very good potential indications. But again, we're talking about a blinded trial. That means parasthesia-free stimulation. And the evidence for the efficacy, that is basically in failed back and CRPS.
DR. ELDABE: Can I ask a question?
DR. NORTH: And I'm glad nobody's mentioned CRPS.
DR. KATZ: You mentioned it just now.
(Laughter.)
DR. NORTH: Okay, fair enough. I think that's problematic. I always found it so clinically. They're basically two forms of it, the surgical standpoint. One, I think understand that I understand; the other, I have no clue.
DR. THOMSON: Like diabetes, there are 5 types of diabetes now, disease trajectories. And I think the CRPS story will be something like that.
DR. NORTH: We need a marker for that.
DR. THOMSON: And it won't be type 1 and type 2, a named nerve injury or not named injury.

Page 171

It's not going to be that.
DR. KATZ: Eric?
DR. BUCHSER: One question I have to the panel, what's your experience with pain in part of the leg or arm that has sensory deficit? Because my experience -- and that relates to polyneuropathy -- if you do have a significant sensory deficit, it simply doesn't work. So those should be exclusion criteria as well.
DR. HAYEK: If it's complete deafferentation, I agree with you, but if it's partial deafferentation, I don't know. Perhaps our neurosurgery colleagues like Brian can comment on that. Maybe it's effective.
MALE VOICE: Say this again.
DR. HAYEK: I was referring to the question by Eric on deafferentation. He's saying if there's sensory loss, then perhaps we should not put a stimulator.
MALE VOICE: I'm not sure. I guess the idea of a surgical question there. That's a --
DR. HAYEK: I was just referring to the

Page 172

deafferentation, how much deafferentation there is in the extremity. If there's full deafferentation or complete deafferentation as in phantom limb pain, it's probably not effective. Maybe Sam published on this.
MALE VOICE: It's sort of an ideology.
DR. ELDABE: Sam Eldabe. I suppose if you look at the result of spinal cord stimulation in CRPS type 1, and you compare it to CRPS type 2, you end up with a completely different set of results because of what you say --
DR. FIELDS: Which one does it work for?
DR. ELDABE: It works well for type 1, but it doesn't work so well for type 2 where you have complete loss of nerve function.
DR. BUCHSER: In all the studies that have been done in post-herpetic neuralgia, for instance, if you look at the results, there's no control size of course, but the case series show relatively miserable results, really.
DR. KATZ: Eric, you're referring to -- your question is about complete the deafferentation in

Page 173

some named nerve segment --
DR. BUCHSER: That's right.
DR. KATZ: -- as opposed to just a minor sensory loss that you have work to detect.
DR. BUCHSER: The only studies I know where minor sensory loss has been treated and has responded is actually in diabetic polyneuropathy, where the sensory loss is incomplete. But in my experience, at least, when you do -- I don't know exactly what that means, but a significant sensory loss where patients really have a loss of sensation, I have been very unsuccessful with those patients.
DR. KATZ: Rick, do you have any comments on that?
DR. NORTH: In some of our older papers that were large samples, not RCTs, we look specifically at sensory loss evident on clinical exam as a prognostic factor and saw no effect.
DR. THOMSON: I concur. I know we're looking for the best model in which to do the perfect study. Maybe this isn't the perfect model,

Page 174

but I think we should be careful that -- I think we've all seen patients where spinal cord stimulations work nicely in people with sensory loss. Profound deafferentation, yes, it's less likely, but these things can happen.
DR. KATZ: Howard?
DR. FIELDS: Howard fields, San Francisco. I was very specific about my prediction. The prediction is that when you do the sensory testing, the initial stimuli are perceived as less intense. As you do repeated stimulation, it becomes more intense than normal. So what that suggests to me is that there is nerve injury in that part of the brain. There may be even hyperexcitability among the remaining fibers.
The main change that's occurring is summation in the central nervous system. And if that's the case, and some aspects of the gate control hypothesis are correct, and what you think you're stimulating is what you're actually stimulating, the prediction would be those would be the ideal patients.

Page 175

So you have a thermal stimulus that only activates C fibers, nociceptors, and you show that they're still intact and contributing to the pain problem, those should be the patients in which it works. So a blanket statement about sensory loss is probably inadequate for the assessment of the patient.
DR. NORTH: I think that's a great idea for optimizing patient selection, and that's part of what we're here to talk about. Then there is the generic study methodology to be applied for all conditions to avoid some of the problems with the research to date.
DR. ELDABE: I have a question for you, Rick, based on the generic study methodology. You mentioned that if we run sham-controlled trial, we would have to exclude conventional stimulation.
DR. NORTH: No. What I have in mind is the trial design coming out of the study would include conventional as one of the tunes that's in the jukebox, that each patient is going to successively try different forms of stimulation. My reference

Page 176

to conventional had to do with the fact that the claims for high-frequency burst and so forth, which are parasthesia free, are limited to a subset of common diagnoses at this point. So in planning the first ever study to show benefit over placebo, we probably should look at those.
DR. ELDABE: I'm just trying to through if you were to randomize a population of whatever been condition to a number of stimulation parameters, including conventional stimulation, and high frequency, and something else, and you put a sham arm, there is no requirement for the people randomized in the sham arm to feel anything because they will be part of three groups who may or may not feel stimulation.
So therefore, the idea of having a sham, we don't have to get into the complexities of the sham of conventional stimulation.
DR. NORTH: But you're saying they might be randomized to that first. We haven't gotten into the details of implantation, but I would assume that we would put in the electrodes using

Page 177

parasthesia mapping. So when it came time to use conventional stimulation -- or for that matter burst for which the data have been gathered with electrodes placed using parasthesia mapping and a trial using parasthesia mapping -- that that would have been done. So everybody will get that on the way into the study, and then they'll be randomized.

DR. ELDABE: That's a good point, but I suppose it forces you down the road of trialing, and it forces you down the road to parasthesia trialing, which may or may not be desirable. If you didn't want to do that, and you implanted everybody in the sham arm with a high-frequency stimulator, would that work?

DR. NORTH: Is the question now are you going to screen patients for implantation using any one or all of the available tunes?

DR. ELDABE: Assuming that you don't screen them because that adds a level of complexity into the story that is never ending, particularly when you get to a sham group.

DR. THOMSON: I think there's an awful lot

Page 178

of plague on has the patient ever experienced a parasthesia or not. Well, that completely ruins the result for subthreshold stimulation or subperception. I think that's another of these things that's been made up, but I think it's true. Especially in this jukebox idea, I think it will be brilliant to have multiple opportunities where the patients don't feel anything, and then, oh, up pops one that you do feel. I think that would be fine.

DR. NORTH: You might make it part of the routine follow-up visit for reprogramming, that there be a brief test of conventional stimulation. Verify the patient feels parasthesia. That reassures everybody that the system is working, whatever waveform you're delivering. And then you go on to do whatever you're going to do.

DR. KATZ: Rod, did you have a comment?

DR. TAYLOR: Just to help you bring us back to the task at hand, I think we're not here to design the perfect SCS study. We're here to hopefully present maybe a checklist of recommendations going forward for what a study

Page 179

would be. One of my observations is that we've got some really talented people in this room thinking about this question, so maybe one of the spin-offs of this meeting would be to bring together such a group to deliver such a trial.

I think one of the observations from the UK is it's very hard to get NIHR to put their hands in the research pockets to fund neuromodulation research. But it sounds as though the NIH, particularly with this biomarkers call, there could be an opportunity there, and we might be able to grasp that.

So here's a trial that would be done by -- investigator led. We would have all of the tunes and Rick's jukebox. It could be called the jukebox trial --

(Laughter.)

DR. TAYLOR: -- and we randomize people to it. And we do that definitive study. I know that's not the purpose of being here. I'm really enjoying the discussion, but I really wonder if that might be a useful ACTTION post of this

Page 180

meeting, is to follow that up. If anybody's interested, I would certainly be very interested in being part of that group.

DR. KATZ: Bob, Dennis, what do you think about -- just to emphasize what you said, Rod, this is not a protocol design meeting. That would be a different kind of a meeting. It wouldn't have so many people, and it would be very focused on a specific hypothesis-driven protocol development approach; although it is part of this meeting to identify what the key scientific questions are and what general research consideration should be, and addressing any of these scientific questions.

Bob, Dennis, what do you think about this spin-off where such a protocol would be advanced?

DR. DWORKIN: I think potentially it's an interesting idea. My sense from having been at multiple NIH meetings over the past 18 months is they see the development of biomarkers as a kind of sequential process where you begin by identifying potential biomarkers, validating the biomarkers, going through the FDA hoops, and then at the end of

Page 181

this lengthy process doing a clinical trial to test whether the biomarker is either a pharmacodynamic biomarker or predictive biomarker.

So I'm not sure -- though I completely agree with what you're suggesting, Rod -- that this can be done at one time rather than waiting for several steps. But I think their approach is this kind of very tortured, sequential approach rather than designing a clinical trial where we do it all at the same time. But we can look into this, and I think it would be a great question to send them.

DR. KATZ: At this meeting on Wednesday, the way that they presented their approach to funding the biomarker initiative is that they presented two different pathways. The first pathway they called the discovery pathway, which was if you really don't have any preliminary data, well then, you can do a small little study and get some preliminary data. But if you already have what they consider to be adequate preliminary data, then they can jump you over that first step into the next step, which is a validation study.

Page 182

DR. DWORKIN: But of course, we're in the discovery phase. Howard had hypotheses, but Howard doesn't have hard data.

DR. KATZ: I don't know to what extent there are data sitting in the literature or that people may have generated on the validity and reliability of various QST measures in patients with lumbosacral radiculopathy. That would be step one. My suspicion is that there's probably a lot of that kind of data sitting in the literature, and then maybe some pilot data or sub-studies that have looked at the relationship between that and outcome.

So I wouldn't dismiss it offhand that there might be enough about pilot data sitting there.

Dennis?

DR. TURK: Dennis Turk. Let me agree with Rod about what the purpose of this meeting is and a potential detailed protocol design as another purpose. I think for the purpose of this particular meeting and this particular paper, what we're really looking for is essentially addressing

Page 183

some of the issues that we heard from Sam and that we heard from some other speakers, is what was Nate pulling together.

So I think the nuances of which specific population is the best one or the idea one, or what exact experimental design you want to use to confirm something, that's fine. But I think for this purpose, we're trying to give -- imagine that the audience for this paper or from this meeting is somebody who's going to go back to their company, or clinic, or research center on Monday, and they're going to start thinking about how can we plan a study that investigates spinal cord stimulators as an effective treatment. What can we give them for guidance and not just wait for NIH to take as long as NIH will take for doing anything that they're going to do.

So I think it's not an unimportant issue. I think it's just more in the weeds of detail that you have some people here who not easily, but who could sit together and craft a protocol, but I don't think that's the exact -- from my

Page 184

perspective, the purpose of this meeting.

DR. KATZ: Yeah, I agree with that. So in the 8 minutes and 36 seconds that we have left in the session, I want to see if we can get at two targeted questions. The first is in relation to Sam's recommendations for how to deal with these programming issues in spinal cord stimulator studies, do people feel comfortable with those recommendations or does anybody have any additional -- would anybody propose any changes to what Sam has recommended?

Eric?

DR. BUCHSER: I totally agree with what Sam said. I'm just wondering how you should actually report those programming changing, and issues, and variability because it's so wide. It's so different from patient to patient, that I imagine the way it could look in a paper would be totally undigestible.

DR. KATZ: Sam?

DR. ELDABE: It's a very good point, and that's one of the reasons why people fail to report

Page 185

this, it's because of the limited space in a publication, and it's not a priority. But some people have done a very good job of doing that, particularly the critical limb ischemia paper, albeit what they did is they published a technical paper.

But as a minimum, if you and I knew what was the mean amplitude, what was the standard deviation of the amplitude and what was the range, we could have a stab at guessing where were these patients? If we don't have any information, it's not particularly useful. So I think as a minimum data set, that's all we want in the main publication.

DR. BUCHSER: Would you stratify that by lead positioning
or cathode positioning?

DR. THOMSON: The spatial target is very important. You call it the cathode. Some others might call it the central point of stimulation. That's very important. Then I think the mode of stimulation, so how you switch, whether you're using, if you like, an anatomical target and a

Page 186

high-frequency thing, or whether you're using best parasthesia position, and then a subperception program.

I think I remember trying to do this for the trial stim study. There are only a limited number of, at the moment, categories, and we can define those and make those our criteria for reporting.

DR. KATZ: Yes, Turo?

DR. NURMIKKO: Just briefly, about sham stimulation, just thinking of what you were saying, Simon, about the Alkaisy study, where 37 percent or so, actually who were supposed to have a proper sham didn't turn out to have that. And that was then revealed after the study. So I just wonder whether there should be something about indication or some kind of an attempt by the investigators prior to the study to ensure that the sham really is sham.

We are taken it as a given, that high frequency obviously has all these features that would seem to make sham totally appropriate, but it might be something to look into prior to so that

Page 187

you're absolutely sure, even if it was looking at some patient samples to check that everything works as planned.

DR. KATZ: Anyone have any thoughts on that? Simon? Anyone?

DR. ELDABE: I think the 37 percent report was in one of the stimulation arms of the study. I suppose when they -- well, patients in the sham group do report parasthesias. That's quite normal. It's not unusual. I'm not sure it's not desirable. It does happen, and I don't see the problem with that at all.

DR. KATZ: Well, since it's good to end every session with a little bit of statistics just to work up our appetite for our lunch, Jen, first of all, I want to say this is probably the only presentation on causal estimands I have heard that I actually feel like I understood.

(Laughter.)

DR. KATZ: And I've heard many of them, so that's an accomplishment in and of itself.

Now that you've presented those options,

Page 188

what are your thoughts on what sort of estimand approach will be appropriate and the sorts of studies we're talking about here?

DR. GEWANDTER: I don't really like to be prescriptive about what to choose. I like to let you choose based on the explanation. For this intervention, I think estimand 3 is good because it's not like a drug, where if you have an AE, you can just stop taking it. We really want to know --

I feel like sometimes for drugs, I think, well, as long as we know how well it works for the people who can actually take it, that's a good thing because people can kind of come and go as they please; whereas this is more of a permanent thing, so I would prefer I think to know the estimate for the people that's actually attributable to the treatment, and therefore estimand 3, which would be more of a jump to reference if you have to stop using the treatment because of an AE.

DR. KATZ: Can you remind everybody what the

Page 189

implications would be for the design and conduct of the study if one were to use that approach?

DR. GEWANDTER: The implication would be that -- so I made it simple by making the intercurrent event I was talking about discontinuation due to an AE, which is pretty serious. There are other more simple intercurrent events, like you took a rescue medication, or you took a medication that wasn't allowed in the study.

So it's more nuance. You have to decide what to do with all those different types of intercurrent events. But the implication would be that if someone withdraws for an AE, you don't necessarily have to follow them or try to get their data. You can kind of just forget about them from a resource perspective. But the caveat to that is -- of Mike McDermott, he would say, well, you want to do sensitivity analyses, so you might want those data anyway. But from a primary analysis perspective, if that's your estimand, you don't necessarily have to work super hard to get those data.

Page 190

But with that said, if someone does drop out, you want to work really hard to figure out why and to know why in a good way. And what we found looking at the FDA database is that a lot of times it's protocol violation and patient withdrew. It's very well characterized, and that makes it really hard to do a targeted imputation that I've explained; so making sure in your trials that you really have a good exit interview, if you will, with the patients and a good understanding of why they decided they don't want to be in the trial anymore.

DR. KATZ: I can say that we've sort of revised the dropout forms, and they're being used more and more in the pharmaceutical industry, where it's no longer sufficient to just check a box that says patient withdrew consent and see you later. So we're forcing investigators to have a more detailed set of checkboxes and then also to write a narrative. When a patient drops out, write a little story about why that patient dropped out so that a human being could look at that later and see

Page 191

if they agreed with where you put your checkbox. And FDA's pushing people, and we've provided some of those forms to different studies. So that seems like it would be important here.

Rod, do you have any additional comments on that issue?

DR. TAYLOR: No.

DR. THOMSON: I've got a comment. One of the things that happens is that these are patients, particularly if you're doing the back pain dominant over leg pain, they end up with a lot of comorbidity and new onset pains, new onset diseases, and it's very common. So if you're trying to do -- you get to your data collection, the question I want to know is like they've got this new onset of pain or whatever, do you try and take a pain score of what their original pain was all about? And then you've got your quality of life measures, and yet they're ill because they've got pancreatitis or something.

I find these things -- I'm just amazed at how ill these people get.

Page 192

DR. GEWANDTER: I think that's a very different question than the estimand question. That's a question of, if you enroll people in the study who have fibromyalgia, for example, can they distinguish between the pain that the spinal cord stimulator was for and the fibromyalgia pain. And I think that's depending on -- unless they're completely discrete, like John's patient in the video, they have completely discrete locations, I think that's very challenging.

So I think that's more of maybe an inclusion/exclusion criteria. And I know that, like was said, the failed back surgery syndrome, people are the ones who are the sickest, so that's hard. But I don't think that's an estimand issue. The only thing I would say is if someone drops out for an AE that's not related to the treatment, that's related to something else, you might treat them differently in your imputation than you would for just like they got pancreatitis and they're not in the study anymore. That's not an AE related to the treatment. That would be more of a -- I would

Page 193

treat that person as a missing at random situation.
So whatever their pain experience, hopefully they can distinguish in the leg or wherever you're hoping the spinal cord stimulator is treating them, at the time of their dropout, you impute their data using the other people in that group, and you look at their trajectory, and you don't consider that an AE.
DR. KATZ: Let me change the subject. I need to take a vote, so get your arm ready to vote. There is an hour and a half lunch break right now, and the schedule says that we're supposed to finish at 4:00. It's Friday afternoon, and I'm not sure if it takes people 90 minutes to eat. But I'm going to take a vote, and I'm going to ask you to raise your hand if you want to shorten the lunch break to an hour instead of an hour and a half, and then stop at 3:30.
(Hands raised.)
DR. KATZ: Well, I guess I don't really have to take the answer for the other --
(Laughter.)

Page 194

DR. KATZ: -- so we'll break for lunch now, and if people could return here at 1:00, we'll aim to finish at 3:30.
(Whereupon, at 12:02 p.m., a lunch recess was taken.)

Page 195

A F T E R N O O N S E S S I O N
(1:05 p.m.)
Group Discussion
DR. KATZ: Okay, everyone; homestretch. So now is the time of the meeting where we start to actually think about producing something out of this discussion. So as all of you know, the intent and the normal way that an IMMPACT meeting works is that there's a summary paper that comes out afterwards, in this particular case, that will summarize the discussions that we've had over the last 24 hours or so.
Jen was suggesting a separate paper on reporting of clinical trials and spinal cord stimulation. As you guys know, there are lots of papers on how to report different kinds of trials. And then, Ewan, your review what seem to logically be, if we can get a third, a paper that would come out of this meeting.
So that seems to be the current discussion. Why don't we maybe begin by talking about that, what the output could be, and then we'll move on

Page 196

from there to content.
Yes, Dennis?
DR. TURK: The way that IMMPACT meetings run is that although Nate's going to draft the first version of this manuscript, it will be circulated to all of you and probably several times with the comments. You should all realize, and we realize, that you're going to think of things tonight, oh, I wish I had said or I forgot to mention. So you'll have a chance to see this. So you don't have to assume that everything is in stone at the end of today, but basically it's enough to give Nate a first shot at this.
All of you will be invited to be authors, and, obviously, it's your decision if you do want to or don't want to be an author. And that includes the industry people. Everybody's invited. If for some reason you don't want to, that's fine, as well. If you choose not to or your company doesn't want you to, we will ask can we acknowledge that you were at the meeting, just for that reality.

Page 197

But just don't think that in the next two and a half hours, whatever it's going to be, that everything is going to be resolved; it's the final version; you're not going to see this again. And let me add to that a plea for Nate, which he hasn't made yet, which is when he drafts this up and circulates it, if you look at the number of people, if everybody sits on it for two or three months, this will take interminable to do. So the idea, typically, is to try to give you some reasonable time frame, like two weeks, or a month, or whatever Nate chooses to want to do.

Even if you want to say it looks good, fine with me, at least acknowledge that because the worst thing is we don't know. We send you an email with a draft, and we don't hear, and there's got to be multiple reminders and requests. So please try to be responsive either to just simply say you like it or you have some important points, or whatever you want to do, but don't just let it sit there in your 800th email, and poor Nate has to keep annoying you and nudging you.

Page 198

This will happen all the way from drafts to the submissions because a number of the journals require all the authors to acknowledge and give approval. Some IMMPACT papers takes forever just to get people to be willing to go to the website and do the two clicks that you have to do to say that, yes, you're an author and you agree to serve on this. So please be responsive.

DR. KATZ: Thank you, Dennis.

DR. NORTH: What is the rough timetable for first draft and ultimately submission?

DR. KATZ: I'm not sure. I think what I'll do is after this meeting is over and I can gather my wits and see what else is going on, I'll send an email out to everyone and let them know what the rough timelines are. I'd like to get everyone a draft within a month. That may feel ambitious, but give me a few days to get back to you on that.

DR. McNICOL: Can you just clarify what you think the two or the three papers will be? I think it will be between our paper and what Jennifer suggests, or would those be two distinct papers?

Page 199

DR. KATZ: Well, that's what I want to talk about now, actually. So it's open for discussion. It seems to me like the content of this meeting would be one paper, which would be something like research considerations in clinical trials, spinal cord stimulation for chronic pain, something like that, which is a very kind of typical IMMPACT style paper.

We could play a little bit with the language of that. Bob and I were having a little bit of a sidebar earlier; do we want it to be research considerations; do we want it to be research recommendations? But let's just put that in the world of detail for the moment.

Then, Ewan, it seems to me like your systematic review stands on its own two feet as a review of methodology in this area. And in some sense, as it did yesterday, would serve as a foundation for what we're doing here. That seems to be a coherent second paper. And I'm seeing a lot of heads go up and down with respect to that.

Then Jen and I just had a quick sidebar

Page 200

during lunch, which is by no means a formal discussion, about whether it would also make sense to do a separate paper on reporting recommendations. And I guess that would be based on your paper as well. You find the deficits, and then she leads the paper on reporting recommendations. But that's just something for discussion.

Do you think it's better to combine them or what do you two think?

DR. McNICOL: Jen's giving me a look that suggests no.

DR. GEWANDTER: No, no, no.

(Laughter.)

DR. GEWANDTER: I was taking a look back [inaudible -off mic], and I was leading it.

DR. KATZ: Yes, that was skillfully done. I give myself credit for that.

(Laughter.)

DR. GEWANDTER: I think it would be easiest because you guys are done [inaudible - off mic].

DR. KATZ: Can you speak into your mic, Jen,

Page 201

please?
DR. GEWANDTER: Sorry. My name is Jen Gewandter.
I think it just depends on how long it gets. I think if you go back to Sam's talk, like all of those different details that he talked about, I don't know that they would necessarily naturally come up in systematic review of results.
DR. TAYLOR: It wouldn't.
DR. GEWANDTER: So it might --
DR. McNICOL: I would certainly encourage the separation.
Ewan McNicol, Tufts. If we're going to do two separate papers or three separate papers, would it make sense to submit all three papers to the same journal, so there's a sort of continuum on here, or do they not fit well together? Is one journal going to be more about -- is one going to be for neuromodulation, for example, and is one more like what Jennifer's done in the past, where she's done this sort of systematic review and I think made some recommendations based on that

Page 202

review as well?
DR. GEWANDTER: Yeah. I usually do the systematic review and make some very brief recommendations in one paper. We did this with the CIPN. I wrote a systematic review with a few recommendations, and then we did the follow-up paper on the design recommendations. They happen to go to the same journal, but we didn't put them in at the same time. It was just that the same journal wanted them.
So it wasn't like they wanted a full picture, if you will, altogether. I think your manuscript would stand on its own. I think just some of the more nuanced intricacies of all the things that are important to report might not have even been things that you looked for in your systematic review. I don't remember what your manual was like.
DR. McNICOL: We definitely missed some stuff.
DR. GEWANDTER: I think, at least I've found, it's kind of hard to comment on things you

Page 203

didn't actually look for in your systematic review. So what I would advise is you write yours first and just write whatever fits well and whatever you think is good. And then if we still think after that, that a more expanded thing is necessary, we could do that. That's what I recommend.
DR. McNICOL: That's good. I have a question since we're getting into the specifics of publication here. During my talk, I mentioned that we didn't incorporate extension studies and we didn't incorporate angina studies. Do people have a feeling for if we should incorporate them and how we should incorporate them?
DR. KATZ: Anyone have any thoughts about incorporation of angina studies?
What was the other thing, Ewan? I couldn't quite hear you.
DR. McNICOL: Sorry. Extension studies.
DR. KATZ: Extension studies and angina studies. Anyone have any feeling about whether Ewan should incorporate that in his review? Sam?
DR. ELDABE: I think angina studies, it

Page 204

would be easy for you to incorporate those. There are a number, I think 6 or 7 of them. I don't see a reason why they should be treated differently.
DR. McNICOL: This may be a question more for Jennifer.
Do you think we can use the same coding manual for angina studies as we do for -- Rod, you're nodding.
DR. TAYLOR: You can because dare I say, we've had the pleasure of systematically reviewing that literature. So there is a citation to Eldabe. I think Sam was the first author or I was the first author, doing the review you've done on angina. But it's quite a date, so I think you bring everything all together, including angina, would be sensible. To leave angina out would seem a bit perverse because actually, as we'll find, some of the qualities of the trials and actually the quality of the results is probably better than we have in some of the rest of the literature. And there are some lessons that we might learn from that and that we would benefit from. At worse, I

Page 205

can understand you haven't included them, Ewan, [inaudible - mic fades].

DR. McNICOL: We're happy to do that with the caveat that that's some extra work that we need to do that may add a few weeks on to the submission process. But I think if you feel it completes the process, we'd be happy to do it.

Rod, a follow-up to that, do you think we should be out in the extension studies, then?

DR. TAYLOR: I wasn't quite sure what you meant by extension studies. What do you mean by that?

DR. McNICOL: Maybe that's not the correct term. I can't remember what --

DR. ELDABE: I may be able to help you. I think you're talking about publications of a longer follow-up.

DR. McNICOL: Exactly.

DR. ELDABE: Yeah. So studies are published 12 months, then they come back and publish it 24 months.

DR. McNICOL: Exactly.

Page 206

DR. GEWANDTER: In the past.

MALE VOICE: Sorry. Can you please state your name?

DR. GEWANDTER: Sorry. Jen Gewandter.

In the past, I've only done the primary analysis, so that's how I handled that. And I don't know if maybe for this specific indication, it might be more interesting to keep them. But we were interested in things like did they identify the primary and how was it designed in the first place. So therefore, I don't know that putting another publication of the same trial would add a lot to it, and it might overly weight that trial in terms of the overall results.

DR. McNICOL: Ewan McNicol, Tufts. I think Bob's suggestion fairly early on in this process was that we mentioned the extension studies as, hey, this is out here, but we don't actually make it part of the review or the systematic part of the review in of itself. Would that make sense?

DR. KATZ: Nate Katz. The only question I would have about that, and this is just a question,

Page 207

is to what extent -- if somebody does a short-term randomized-controlled trial, and then does multi-year open-label extension, they may relegate their more detailed methods for evaluating safety, and they may push that into the open-label extension.

So I don't know enough about this literature to know whether that happens here. But if part of your findings is going to be how well safety was reported. Then I wonder whether you'd need some of those open-label extension to see where the safety methods actually were implemented.

That's just a question. Maybe some of you know whether that's an issue or not.

DR. ELDABE: Sam Eldabe. I think most of the studies that Ewan is alluding to are following the same protocol, but for 24 months instead of 12 months or so. It's not an open-label extension as you would find in drug trials. But you are absolutely spot-on in that there are some aspects of the study that are reported in 24 months extension that are not reported in the 12 months.

Page 208

And sometimes the safety is mainly reported in 24 months.

DR. KATZ: I guess what I would suggest, Ewan -- and this is just a suggestion -- is to maybe take a look at a couple of those studies and see if they do indeed seem to have safety methodology in what you're not calling the extension studies. And if there is really more meat on the bone in terms of how safety was assessed in these extension studies, that it may be worth pulling them in simply to be able to comment on strengths and weaknesses about how safety is reported.

DR. McNICOL: I think Sam has quoted it correctly. I think there are there safety data quoted in these studies that aren't really looked at or maybe are not even valid for shorter term, 3 months or 6 months.

DR. NORTH: Rick North. Just as a practical matter -- you already spoke to this in another way -- I think it's important that you cite all these studies, include them in your bibliography,

Page 209

so that readers don't think you're unaware of them.
DR. McNICOL: That's good. I was saying to Jane yesterday that in manuscripts, often it's stuck in an appendix somewhere that is really difficult for someone to access. When we do a Cochrane review, we list everything because Cochrane reviews are 200 pages long. So I want to balance practicality with completeness.
DR. KATZ: Yes. Ewan, since we're on the topic of your paper, is there any other feedback you think would be helpful from this group today in terms of proceeding with that project?
DR. McNICOL: I think those were my main -- and as you say, this isn't finalizing things. This is just giving us something to work with that we can send to you guys, and then we'll comment from there. But I just didn't want there to be some fundamental aspect of it, that we submit something to you and you're like, you should've thought about this. You're going to have to go back and start over. But I think everything else is details that we can iron out once we submit our

Page 210

first draft to you.
DR. KATZ: In terms of who is involved with that paper, I don't know if Bob, Dennis, Ewan, if you had any -- were you thinking that everybody in this room would be an author?
DR. DWORKIN: The commitment was already made. It was the steering committee for this meeting. So I think it's basically the speakers of this meeting. We had an email trail going back a year and a half, so that's been resolved.
DR. KATZ: That will make it much easier. But I think, Ewan, if you feel like you need specific input from somebody in this room who's not on that list, I'm sure they'd be happy to correspond with you.
DR. McNICOL: Yes. Sam's already offered. Thanks.
DR. KATZ: And you can throw them any acknowledgements or whatever.
DR. THOMSON: Simon Thomson here. Can I just say two things? One is this SSED, do you include that in a systematic review? Is it allowed

Page 211

to be included in there?
DR. McNICOL: Just to clarify, the SSED is the safety and efficacy paper.
MALE VOICE: Is this a British thing as opposed to an American thing?
DR. THOMSON: It's American. It's an FDA thing.
MS. LEITNER: It's a requirement [inaudible - off mic]. Sorry.
DR. KATZ: And say your name.
MS. LEITNER: Angela Leitner. SSED is a FDA requirement. When you submit a PMA, they go through it and say what you can and cannot say, and it's a basis for your claims that you can make for your product. So it usually provides a lot more complete information based on how you prespecified you would analyze your data. And you can look up the example of the Nevro one and see how complete it is, and make sure it is included.
DR. McNICOL: Jennifer, have you done that for yours?
DR. GEWANDTER: No, I have not. I think

Page 212

your review was of how these trials were reported in the peer-reviewed literature. I totally get your point that you get like all this extra interesting information from the SSEDs, but that was not necessarily the objective of what you were trying to do.
So I think it would be a very interesting but different review of the literature. I think that would be interesting to kind of shed light on some of these issues that we're talking about in terms of the weaknesses of these studies, and actually maybe contrasting them to what was actually reported in the papers would be super interesting. But it's just I think a little bit outside of the scope of what Ewan was asked to do.
DR. KATZ: Rod, did you have comment on that?
DR. TAYLOR: I would just entirely agree with Jennifer. The equivalent in Europe is called an EPAR, European public assessment report. But again, in Cochrane reviews, we normally don't look at those. We use the published evidence.

Page 213

There is an interesting second review question, which is what is the disparity between EPARs and the American equivalent and the published paper? But I would encourage others to follow that up as a separate publication.

DR. McNICOL: I will say that for the studies we looked at, if they were registered and there was information that we couldn't find in the preliminary or the published paper, we would look on clinicaltrials.gov, or whatever it was, for that extra information. So we did go into a little bit more depth, but I wonder if that's the limit of the depth that we should go into.

DR. KATZ: Yes, I would say so.

All great. So that's project number one, is the systematic review. I want to just take a step back out of this rabbit hole a little bit and just talk about the overall framework of what we're trying to accomplish. So I think we've settled that.

Bob, I saw your hand up earlier when we were talking about just the overall set of papers we

Page 214

were thinking about publishing. Did you have something to say about that?

DR. DWORKIN: Bob Dworkin. I think the primary paper from this meeting, having to do with recommendations, checklists for the design of randomized clinical trials of SCS, to me is also recommendations for reporting because we're going to have a bunch of recommendations about the best gold standard practices for these trials. And implicit in that is if you don't follow what we're suggesting, you really should kind of provide the rationale for not attending to our recommendations in your publication.

So I see the possibility of a final table in this primary paper maybe being reporting recommendations, but I don't clearly see -- and maybe we should just defer it until the end of our discussion. I don't clearly see a third publication on reporting because it seems implicit in our design recommendations.

DR. KATZ: Jen? Can you use your mic?

DR. GEWANDTER: That's fair.

Page 215

DR. KATZ: Fair. That was Jen who said fair.

So you are in agreement with folding that into this publication?

DR. GEWANDTER: I think putting a final table that explicitly outlines what you'd want to see is a good idea, and just see how long it is.

DR. KATZ: All right. I'm comfortable with that plan of trying to fold it all into one. If it turns out that that part of the paper becomes so bulky that it needs to give birth to a separate paper, we can make that decision as we go.

Great.

DR. THOMSON: Simon Thomson here. I put it in a symposia suggestion to the INS in Sydney, which is late May 2019, to really sort of -- I had a 2-hour symposium suggestion with various people about how we got to this place. And then Robert Dworkin would then present the findings of our group.

They cherry-picked, and they've taken you for something, and they've taken Rod for something

Page 216

else, and they've invited me to present the findings of our group in a half-hour plenary. And I'm asking you is that okay and can you help me?

DR. KATZ: Yes and yes. I think it's great, and certainly I'll be happy to help. Does anyone else have any other feelings about that symposium?

DR. THOMSON: I was thinking I could do a double act with somebody. I think the trouble is they want to fund two speakers to fill one slot. But if you're already covered, that would be okay. I could reply back and say, yes, I accept, but we want to do a double act.

DR. KATZ: We can talk more about that offline, but certainly in principle, I'll be happy to help with that process in whatever way makes sense.

DR. TAYLOR: Nate, just on the publication specifically, in the spirit of transparency, again, because I've twisted their arm -- so I'm Rod Taylor, by the way, if you hadn't already guessed. Thanks. Sorry, folks.

We will be taking the cost effectiveness

Page 217

review to publication as well, we being really myself, Sam's graciously signed up and so has Brian. But in the spirit of openness again, we will be writing that up for publication. If anybody in this room has a burning desire to be part of that, they'd be more than welcome to email me if you're comfortable with that, and we would add them into the authorship, obviously, with the normal expectations of what authorship would be here.

If I can just throw that out there, so people have got my email and want to follow up outside the meeting. But I would see this as being another paper that's kind of resulted from this get-together, if that's okay.

DR. KATZ: Thank you, Rod, for that. So if anybody's interested in being a co-author on the cost effectiveness publication, then shoot Rod an email.

All right. I think then that actually brings us to the topic at hand, which is what's going to go in this paper, research considerations,

Page 218

recommendations for RCTs of spinal cord stimulation for chronic pain.

Yes, Rick? Say your name.

DR. NORTH: Well, this is a subtopic.

DR. KATZ: That's Rick North.

DR. NORTH: We had a lot of discussions about specifics of study designs that might be adopted, and I wonder whether this group would be anything but complemented if the companies that are here, perhaps in conjunction with subgroups like IoN, were to pick up some of those ideas and run with them and get a trial going.

Were that to happen, what would be the proper way to acknowledge the genesis of the ideas? Waiting for the publication to come out so we could cite is not the American way.

(Laughter.)

MALE VOICE: Shoot first, yeah.

DR. KATZ: Go ahead, Bob.

DR. DWORKIN: Bob Dworkin. Typically within, I don't know, 6 to 8 weeks of the meeting, we on the IMMPACT website have all the slide

Page 219

presentations, the agenda, the list of participants, and then a full transcript, high-quality transcript of the meeting. So one possibility would be to cite that website that has essentially all of the meeting. And in fact, what I'm saying right now will be in the transcript.

(Laughter.)

DR. KATZ: Rick, does that meet that meet your needs?

DR. NORTH: Yes, good answer.

DR. KATZ: Great. So back to the content of this paper. As you heard from Bob, the current thinking is to summarize the presentations. Obviously, there are some redundancies, so reshuffle that into an order that makes sense and is cohesive. It's not going to be this talk was about this and that talk was about that. It's going to be obviously organized topically so that it flows in a way that makes sense.

Then presumably, the heart of that paper will be some sort of checklist of -- let's just call them recommendations for these sorts of

Page 220

clinical trials. So you can imagine in your mind's eye a table that has sections in it, and the sections have little check boxes in them, or things like that.

So let's put that all in on our mind's eye right now and open it up for discussion. What do people think should be in that checklist?

Now, we've been talking about this for 36 hours now, so I think we all have a general sense of what those things are, but now in the last hour and a half or so is an opportunity for people to raise their hand and say, well, in case it wasn't clear, I really think this thing should be in the checklist, or in case it wasn't clear, I think this thing should not be in the checklist.

So I'll open up the floor for those sorts of comments. Rick?

DR. NORTH: Rick North, again. WikiStim, which Jane and I organized, already has in place a candidate checklist, if you will. And the SCS subsection of the site has a list of close to 200 variables that potentially can be filled out for

Page 221

any research paper in the field. So I'd just offer that as a starting point.

DR. KATZ: Great. Are those recommendations for research methods or are they checklists for what should be reported in a paper? What are they? Jane?

MS. SHIPLEY: Jane Shipley. They're all of the above.

And actually, I'm in the process of redoing this and elaborating, trying to keep the number of major checklists at 200, but offering but offering answers that can be easily copied and pasted into a sheet.

For instance, on the analytical methods, for example, I'll list all the methods out and somebody can grab them. But right now, even in this format that I'm hoping to improve, it still should be useful as a starting point for you. There are 200 -- and you can download the whole data category list. It's a CSV.

If you search a paper that's not been completed -- that would be the best thing to

Page 222

do -- a completed one will have things filled out, but search one that's not completed and just hit CSV, and you'll get the citation, but you'll also get all the data categories.

DR. NORTH: Or download the entry form, which has examples as well as definitions.

MS. SHIPLEY: Yeah, yeah. There's a submission for it.

DR. NORTH: Rather than characterize this as a potential guideline, it's more a database of what people have reported to date.

DR. KATZ: I see. Okay, great. Thanks for that.

MS. SHIPLEY: There's a lot that needs to be added, and I've been informed by this meeting. And I also was thinking about how to acknowledge that, and I was happy to hear about the URL because that will work for me, too.

DR. KATZ: Okay. Wonderful. Thanks for that. I'll find you if I need you.

Okay, great. Yes, Sam?

DR. ELDABE: I appreciate we've had an

Page 223

excellent talk on outcome measures, but you guys have done quite detailed work on outcome measures. I'm not sure that there is anything specific about SCS that requires us to redo that.

DR. KATZ: Well, we can make that a question. We can certainly refer to existing published guidelines on outcome measures for clinical trials of treatments for chronic pain and use that as a starting point. Dr. Dongen mentioned to me yesterday another website, a European website, that has its own consensus recommendations for outcome measures and various disorders, including chronic pain syndromes. I'd be surprised if there were any huge discrepancies, but that will be another place to go to and cite as a source.

Yesterday, we began to have a discussion of, well, what else besides those are "peculiar" to use Rod's word, to spinal cord stimulation that we should think about including? And some did come up yesterday.

DR. NORTH: I can give you a major example that been emphasized at all at this meeting. But

Page 224

for many, many years, the technical goal of spinal cord stimulation was to elicit parasthesia that overlapped a patient's area of pain, completely, with perhaps some extraneous areas of stimulation as well.

There is a lot of literature dealing with that. You might call that a surrogate outcome measure, and it does certainly correlate with pain relief. It was felt to be, and still is to a large extent, a necessary condition for pain relief by conventional stimulation, to scratch where it itches, as it were.

DR. KATZ: So are you saying that the extent to which the degree -- for that type of stimulation, the extent to which the parasthesias overlap the area of pain should be an outcome measure, should be reported?

DR. NORTH: Over many years, it has been routinely reported as one of the outcome measures.

DR. KATZ: Great. Any others?

DR. THOMSON: Simon Thomson. We talk a lot about neuropathic pain. The systematic review is

Page 225

including ischemic pain syndromes. There are other outcomes other than pain measures in ischemic. And indeed with -- well, I'm going to call it angina. Actually, a pain score is almost a useless measure as an outcome. So we need to ask ourselves are we going to be advising in that, and the same with critical limb ischemia.

DR. KATZ: That's a great question. Should we include consideration of chronic and angina or limb ischemia in this paper?

DR. NORTH: Well, you certainly should mention them. This is Rick again. Are you going to get into them in detail? Because there are other applications ongoing and potential for SCS, and I think it would require a lot of effort to get into them all, and even then, to cover them adequately would be practically impossible.

DR. KATZ: This is a chronic pain paper --

DR. NORTH: Pain, the title.

DR. KATZ: -- so your recommendation would be to indicate in the paper somewhere that this technology is used for angina and limb ischemia,

Page 226

but that the review will not focus on those areas.

DR. NORTH: Each of those might have its own special outcome measures.

DR. KATZ: I agree with that. Does everybody else agree with that? I'm seeing the heads go up and down.

Ewan?

DR. McNICOL: I just wanted to comment on that. When we were talking about the reporting paper, we agreed that it was important to include angina studies. But now we're saying for this we shouldn't include it. Do you not think we should be more consistent across the two publications? It's either important or it's not.

DR. NORTH: It did include ischemia, if I remember correctly.

DR. KATZ: Rod, I think you had advocated for including the angina, the vascular syndromes. What's your opinion on Ewan's question? Is it worth it? Should we pitch it to be consistent or keep it for completeness?

DR. TAYLOR: Rod Taylor. If you're asking

Page 227

me, I think for comprehensiveness, I would suggest that we include the ischemic indications in the systematic review. But to keep this paper light touch, I think it's basically a statement, consideration of outcomes, and need to consider the specific indication, some of which will have their own core outcomes.

I think part of this document is sign posting, isn't it, Nate? So it would be sign posting people to maybe whether those guidelines in those specific therapy areas already exists, assuming that they do. And there are certainly an area of angina that there's many recommendations of outcome measures. So we can just point people there, but we don't need to spend a lot of time in that space. I think that's what you're suggesting, Rick.

DR. NORTH: I think you're making the point that many of the angina and limb ischemia papers were of pretty good quality. And to the extent they allow us to paint a better picture of the quality of this literature, it would make a lot of

Page 228

sense to include them as pain papers.

DR. KATZ: In the systematic review, but not to deal with it specifically in the paper on research considerations.

DR. NORTH: Right.

DR. KATZ: So there will be a little bit of inconsistency, Ewan, but it sounds like there's a rationale for that inconsistency.

Are you comfortable with that?

DR. McNICOL: As long as we discussed it and there was a rationale for not being consistent, then I'm okay with it.

DR. THOMSON: Simon Thomson. The counter, really, is the Europeans, we have quite a lot of experience with ischemic pain syndromes, and some of us -- I was involved in the early EPAR study on CCLI, and Sam and I on an incomplete feasibility study in angina, and Rod has done a systematic review, which has included those cases. I think we do between us have knowledge about these sort of outcome measures with a thought we could provide recommendations.

Page 229

DR. KATZ: Turo?
(Dr. North speaks instead.)
DR. NORTH: But they are peripheral, aren't they, to the central theme here, which is research design consideration design considerations, pain trials. When the draft comes around, if somebody wants to add a paragraph that says, by the way, in limb ischemia, other outcomes salvage.
DR. KATZ: Yes, Turo and then Dennis.
DR. NURMIKKO: Turo Nurmikko, UK. You did mention in the morning, and I think everybody agrees, that as far as chronic pain is concerned, pain measurement has to be the primary outcome somehow. But at the same time, there all sorts of issues related to that, and a paper of this kind can be emphatic enough to actually start changing minds.
It just occurred to me, to suggest that the paper would indeed discuss this goes a little beyond what the IMMPACT paper was saying and underlining the importance, especially when you deal with invasive treatment, the importance of

Page 230

such issues as quality of life and functionality, et cetera, even perhaps going as far as suggesting that in certain circumstances, you could consider either one of them to be a co-primary outcome.
DR. KATZ: Yes. Dennis?
DR. TURK: In the first two IMMPACT papers, the first one was looking at outcome domains and the second was on specific measures. In the domains, we were saying across the board in chronic pain -- this was for chronic pain trials -- these are the domains that should cover pain: physical function, emotional function, et cetera.
In the second paper where we talked about specific measures, we said there are many measures that have been developed specifically for a particular disease entity. When you want to be looking at, for example, physical function, you should be considering using those well-established -- because if you have low back pain versus upper neck and cervical pain, it may be a very different physical function measure. When they do not exist as any disease-specific measure,

Page 231

then we recommend some generic
So I think we got away from having to discuss every one of the possibilities by either referring by reference -- I think we did. We gave a few examples of references. For back pain, specifically we could use -- not the OMERACT, the WOMAC, but we gave some examples.
The idea was that physical function is important to consider. If in fact there are well established measures for angina, if there are such measures out there, then you would use those. If in fact there were none, then you would default to one of the more generic measures like the SF36 or the BPI, or something else of that kind.
So I think you can get around having -- because or else, my fear is -- at least it was our fear that there are so many specific disorders that have their already well established measures, that for us to try to cover all of them would make no sense.
DR. KATZ: Yeah. That feels like a big project to me. And then do we have to talk about

Page 232

pancreatitis and interstitial cystitis, and chronic testicular pain, and all sorts of other syndromes. So I think I'm going to err on the side of, as you said, follow the example of those papers and keep it generic, and mention that there are specific other syndromes which have their own measures and their own considerations. And if that feels inadequate to Simon or anybody else, then we can always revisit that down the road.
Yes?
MS. LEITNER: Angela Leitner. One thing I think is important is to address a nonbeliever like Dr. Fields and open up the black box of SCS, and what can we do to elevate the therapy?
(Laughter.)
MS. LEITNER: Is it FMRI, laser evoked potentials?
Who's laughing? Anyway, I think what's the best thinking on that and what can we recommend controlled trials run for objective measures that we can look at.
DR. KATZ: Do you mean objective measures of

Page 233

pain?
MS. LEITNER: Right. What's the next best thing. What are other ways we can [inaudible - mic fades] that information.
DR. KATZ: To be honest, the short answer is that there aren't any that are useful in clinical trials. Multi-day meetings just on this issue, and every day, an email crosses my desk with yet another paper/company/what have you. I don't know. That feels like it would be a big job to try to put that in this paper. I don't know.
Does anyone feel differently about it? Bob?
DR. DWORKIN: Bob Dworkin. It sounds to me from this morning's discussion, there really was a tacit consensus that when it's possible to do so, that some kind of sensory phenotyping, sensory profiling is done with a combination of patient-reported symptom measures, and they exist for neuropathic and non-neuropathic pain and quantitative sensory testing.
So we could have a soft recommendation, that in many circumstances, sensory phenotyping should

Page 234

be considered using a combination of patient-reported outcomes and quantitative sensory testing.
DR. KATZ: Howard, what do you think?
DR. FIELDS: Howard Fields, and I want to stress that I'm not a nonbeliever, but I am a practitioner of equipoise. That was so articulately described earlier.
I think just a general push toward getting more, what should I say, uniformity in a patient group, prior to entry to the study. So you wouldn't necessarily want to mix, let's say, angina patients with diabetic neuropathy because that introduces its own variability.
What I was pushing for was criteria for a neuropathic component to the pain just on the off chance that those patients might do better, so that some sort of recommendation about grouping patients or stratifying patients, if you will, selecting patients, prior to the start of this study.
The other thing that I was pushing for was we had a big discussion, and it turned out, I

Page 235

thought, fairly general agreement, that there was this idea of reprogramming that was sort of not unplanned and not scheduled, was a real hazard in terms of the interpretation of the data. So if you could have a way of doing the programming in advance of study entry, that would be a major advance in terms of the quality of the data that you would get in your ability to interpret it.
DR. KATZ: We've got a few different -- people have been advancing what I think about as sections of this checklist. There was a section on outcome measurement that we began to discuss, and there's a section on patient selection that we discussed in great detail earlier today and that Howard has brought up now again for further refinement, and one could imagine other sections as well.
So maybe what I'll do is try to march through these sections, or maybe I should just ask explicitly, and start at that high level, and then we can dig deeper in each one. So aside from a section in our checklist on patient selection, a

Page 236

section on measurement of outcomes, we'll need to have a section on this whole reprogramming issue, which seems to have one foot in outcome measurements, and one foot in patient selection, and one foot in the trial. I'm not sure exactly where to parse it out yet, but we don't have to worry about that now.
What are the other high-level sections that should be on this checklist? Something on study objectives; I would imagine that's important. Jane?
MS. SHIPLEY: [Inaudible - off mic] -- biological complications, device complication -- Jane Shipley; sorry -- stimulation side effects; cost effectiveness; implantation; description of the implantation procedure; description of the screening trial; description of stimulation parameters in mode and everything else that's available now; pain location; pain characteristics; demographic factors and study population.
DR. KATZ: Something tells me if I look on

Page 237

your checklist, I'll find the answer to my question.
MS. SHIPLEY. Yes. Thank you.
DR. KATZ: You've got it.
DR. TAYLOR: Rod Taylor. So going back to high level, I agree with Jane, many of these are what I would call reporting issues. But I think we had a cracking presentation from Sam today on considerations of how one might do placebo trials in this setting, and I think you gave us some recommendations based on the literature of how we might go about that. So I would definitely see that as being an important section.
Then the other kind of peculiarity for me is how we bring in the learning curve issue. We didn't talk a lot about it, but basically for me, that's what might be an implanter and center selection issue. Again, we don't need to be definitive, but I think part of this document is just raising it as an issue and what the considerations might be for trialists to think about in this setting.

Page 238

Then the last one I've got, and I think it's peculiarity again, I don't know if it comes under programming, but it's what I would call in the complex intervention world of intervention fidelity; how do we know that the intervention was delivered as protocolized? And I genuinely don't really know the answer to that, but I think is tremendously important in this area. And part of that is the programming because the programming is all part of it. But I think there's more than just programming and fidelity.
DR. KATZ: Right. Very good. Ro?
MS. JAIN: Roshini Jain, Boston Scientific. This section on how to reduce bias perhaps, especially given these are all patient-reported outcomes, so reducing expectation bias; how do you frame these, say, programming paradigms; et cetera, et cetera.
DR. KATZ: Thank you. I won't forget my own talk. I probably will, but thanks for the reminder.
Greg?

Page 239

DR. FIORE: Greg Fiore. I just wanted to add onto that any measures taken to minimize likelihood of placebo response, so along the lines of sham response.
DR. KATZ: Thanks. Sam?
DR. ELDABE: Just to bring up Rod's point on intervention and fidelity, I think you have a number of components to this, and that's the surgery, and that's very well described in every publication, the programming and we've got a detail on how this is done. Then there is the instructions that are given to the patients as they leave the clinic, and finally the adherence to the therapy, and most devices will record when the device has been switched off inadvertently.
DR. KATZ: Great. Any other high-level sections to this checklist?
DR. THOMSON: Simon Thomson here. I think, Jen, the management of missing data because it's not just going to be pain scores missing; it will be all sorts of things, visits. Then I think the big thing is adverse event reporting and whether we

Page 240

can give advice as to what are classified as adverse events.
DR. KATZ: Yes, Salim?
DR. HAYEK: Salim Hayek. As long as we've had spinal cord stimulation, we've always felt that, well, it could be helpful in neuropathic more than nociceptive pain, but really, you can't find concrete evidence for either assumption. And the latest studies all say, well, for back pain or leg pain, which we know, at least for back pain, could be a component of nociceptive pain in addition to neuropathic pain.
So I'm not sure if this is something that we should put in there, but perhaps for pain studies on spinal cord stimulation, we should make a recommendation that there may be some effort to try to identify the percent or the contribution of neuropathic pain to the overall pain of the patient.
DR. KATZ: How would you do it?
DR. HAYEK: There are neuropathic scales or validated questionnaires. For example, the one --

Page 241

MALE VOICE: DM4 [inaudible - off mic]?
DR. HAYEK: Yes, and then the one by Ralf Baron.
MALE VOICE: Pain Detect.
DR. HAYEK: Pain Detect.
DR. KATZ: All right. And we have Howard's suggestions from earlier. Actually, I had Andrea in the back, and then I'll go to Robert.
DR. TRESCOT: Andrea Trescot. I'm not sure I heard the functional outcomes evaluations, Fitbits, iWatch, walking tolerance, the recognition that we have to look at pain scores and we have to look at global improved -- or global patient perceived outcomes. We still have to look at something that's objective. And if we're not going to look at medicines per se, then we need to look at something that is functional.
DR. KATZ: So just to respond to that, we have to look at medicines.
DR. TRESCOT: And medicines as well, but medicines, unfortunately, in this day and age, the opioids are being taken down whether patients are

Page 242

hurting or not. They are being taken off their opioids whether they're hurting or not. So to look at the opioid level is not going to give you any indication, today, of what their level of pain is.
DR. KATZ: They still need to be quantified in any case, of course.
DR. TRESCOT: I agree.
DR. KATZ: Yes. I think we all agree that there in this list of recommendations, we will recommend that there needs to be, at minimum, a patient-reported outcome measure related to physical function; as Dennis just said, a generic one if there's no disease-specific one available; a disease-specific one if there is one available. I think you're talking about performance-based outcome measures where you actually have patients do things, and you measure what it is that they do.
Maybe it's worth having a minute or two discussion about that. What do people feel about the role of performance-based outcome measures in clinical trials with spinal cord stimulation?
DR. THOMSON: I thought we'd agreed

Page 243

that -- IMMPACT has done a lot on outcomes. We're literally going to -- I mean, it's really, are there any outcomes that are different because we're treating people with spinal cord stimulation.
DR. KATZ: Are there?
DR. TRESCOT: Generalized activity. Sorry. This is Andrea again. Generalized activity, improvement in sleep, all of which can be measured by an activity watch. So if they're getting up 4 and 5 times a night or they're sleeping through the night; if they've increased their general level of activity --
DR. THOMSON: But these are not particularly new outcomes because of spinal cord stimulation.
DR. TRESCOT: Well, I would argue that they would because you're now getting people -- improvement in sleep is something we've been able to show over and over again, is something that the --
DR. THOMSON: Sorry, Andrea. These are already outcomes that are measured in any pain trial. They're recommended by the IMMPACT ACTTION

Page 244

group in any pain trial.
DR. TRESCOT: I see. I'm sorry.
DR. THOMSON: Do you see what I mean?
DR. TRESCOT: Okay.
DR. THOMSON: So it's like what are the outcomes that are specific to our therapies, if any, when treating pain?
DR. TRESCOT: Put them on their treadmill.
DR. NORTH: [Inaudible - off mic].
DR. KATZ: Can you speak in your mic, please? Rick North.
DR. NORTH: Rick North. The watch than Andrea refers to is a wonderful tool for measuring an outcome, but it's common to stimulators and other pain trials.
DR. THOMSON: It needs to be validated.
DR. NORTH: Yes, good point.
DR. KATZ: Yes, Robert?
DR. VAN DONGEN: Can I make one comment? Robert van Dongen. Did we cover the psychosocial aspects enough? I know it's common for the other pain syndromes, but psychosocial existential

Page 245

problems can be a problem in these patients in treatment.
DR. KATZ: Yes. Bob?
DR. DWORKIN: I think Simon raises an important point, which is, shouldn't the bulk of this article be what is specific to spinal cord stimulation trials? We don't really need to talk about that it's important to capture adverse events using state-of-the-art methods, and that we need to report the patient's age and the patient's sex, et cetera.
I think that so much of that already exists in the literature, that our contribution really can be to emphasize what's specific to these trials and not these generic questions. I mean, of course we want to get anxiety and depression as outcomes, maybe as moderators of treatment response, but that's true in every chronic pain trial, and we can cite all of that literature. But it's what doesn't exist in the literature or best practice recommendations for these clinical trials.
DR. NORTH: Bob, I agree with you. Rick

Page 246

North again.
This narrows things right down to a couple of SCS specific technology examples. There's an implant with accelerometers that can be used to measure activity. And one of the manufacturers is now using the aforementioned watch as part of the programming system. So SCS gives us opportunities to measure outcome that are specific to the modality.
DR. KATZ: Yes, Eric?
DR. BUCHSER: Eric Buchser. I think we should be very careful with the physical activity because the data that's out there does not actually support the fact that the level of physical activity is correlated with the intensity of pain. If you look at fibromyalgia, for instance, where it's been done extensively, it's how physical activity is distributed over the day that is different, but the total amount of walking distance and in those studies that have looked at the speed, the stride lines, and all that, actually do not correlate with the physical activity on the whole.

Page 247

So what's implanted in the stimulators, the way I see that is more of -- it's a marketing tool. You have this fantastic thing that shows you how much or less the patient's worked, how much they were sitting or lying, and I don't think there's any scientific support to that, so I would be very careful.
DR. KATZ: Ewan?
DR. McNICOL: Ewan McNicol, Tufts. I agree with Bob that if we just list off a bunch of outcomes that we need to look at, we're really just repeating the IMMPACT recommendations from the very first paper. I think we should focus more on what is peculiar to spinal cord stimulation and what of the recommendations that were made in that 2001 paper are conversely what are not relevant to spinal cord stimulation. What things do we need to measure for drug trials that we don't need to measure for spinal cord stimulation?
That could be nothing, but I'm just trying to think is there something that's not relevant to this literature.

Page 248

DR. THOMSON: Simon Thomson here. I think this is the right way to be thinking. So device usage, for example, it might be recharging intervals. It might be -- what else can we think of? Help me. Oh, device longevity, time to explant, because explant isn't always with non-rechargeables; it's a normal thing. You would expect to take it out and put a new one in.
DR. KATZ: It seems to me that we're being somewhat idealistic about how faithful people who report trials and spinal cord stimulators are 15-year old guidelines for what the basics are of reporting. It seems to be like --
MALE VOICE: Maybe I can quote them and say refer to -- [inaudible - off mic].
DR. KATZ: It seems to me a little refresher could come in handy for this research community, so I'm tempted to suggest that we -- how much real estate in a paper does it take to mention the 6 core outcome domains? That's 6 words if I'm counting correctly. That's going to be okay. So I think we can very briefly highlight where there are

Page 249

standards, but then focus the majority of attention -- I think we can have our cake and eat it too a little bit in that way, and then focus the bulk of it on what's actually particular to spinal cord stimulation. I think I'll give that a shot and see how it comes out.

Yes, Roshini?

MS. JAIN: Roshini Jain, Boston Scientific. What about loss of therapy over time?

DR. KATZ: Loss of efficacy over time?

MS. JAIN: Correct, yes.

MS. LEITNER: Angela Leitner. I think that we really need to then define loss of effect as a group for SCS because many people define it differently, so it'd be good to have a recommendation coming out of here.

DR. KATZ: And this, of course is -- any treatment for chronic pain, the same issue, drug treatment, opioids, and loss of efficacy of opioids, as everyone knows, that's an enormous issue. As far as I know, there are no standard definitions for loss of efficacy of any type of

Page 250

chronic pain treatment, of any kind that I've ever seen.

Anyone know different? It's kind of rules of thumb. Data's presented in every imaginable way. I think we could certainly talk about the importance of attempting to define it. For a one-week study, it may not matter, but for a one-year or two-year study, researchers should figure out some way of measuring that.

Do we know enough now to be prescriptive about how that should be done?

DR. NORTH: Rick North again. From the long perspective, going back to the '70s, some of our papers have reported like 20-year maximum follow-up. We have referred to a minimalist outcome measure, which is just patients still using stimulator as some reflection that it still is helpful.

DR. FIELDS: Howard Fields. I want to pick up on what Rick just said. How many of the devices out there can subject turn on and off when they want to? Do we have that data?

Page 251

MALE VOICE: All of them.

DR. FIELDS: All of them?

DR. KATZ: All of them, I think.

DR. FIELDS: So then the question becomes, in general, what's your impression about how patients actually use the stimulators? Do they use them all the time? Do they turn them off at night? If it turns out that they use them intermittently, you could ask them what is it that the stimulator is most helpful for, how often do you use it, and how effective is it for that?

I don't know that you'd get a standard outcome measure, but you could begin to get a better idea of how the subjects actually use it now, and that way it's a lot like an analgesic drug, where somebody takes an ibuprofen when they have a headache and they don't take. And it could be very effective, but maybe they don't have a headache for 6 months. That would enter into their overall assessment of how helpful it is.

So it might be interesting if there were some data on intermittent use and how helpful that

Page 252

was when they used it.

DR. KATZ: That's interesting.

DR. LOESER: John Loeser. Howard, you have to recognize that people are warned, "You don't want to use this too much because you'll run the battery down." So you end up then with a very complex pattern of how much the patient uses it versus how much their fear is the battery's going to decrease. It's not just their pain level, in other words, that determines how long they use it, how intensely they use it, and so forth. It's a very complex issue.

DR. KATZ: Yes. I have to capture that.

DR. THOMSON: Simon Thomson. Obviously, the battery issue was an issue in the past. Most of us use rechargeable devices. It's just become a non-issue. They toggle now between these different subperception and parasthesia-based programs nowadays.

Howard, all of the above is how they use them. There will be some who will have it on all the time on subperception mode, and if they go out

Page 253

for a walk, they might put it on to parasthesia. So it is fascinating how people use them, and then, yes, there are people who basically their background pain syndrome improves over time, then they get stressed, or they do a lot of exercise, and then they turn it on. So all of the above.

DR. FIELDS: Howard Fields. That's absolutely fascinating, in a way, but it's not like there's a Twitter post. There's not like there's a way where somebody like me could look it over and get a fuller picture.

Just like John Markman's videos, that would never come out in any outcome measure, yet you hear this person say it, and you're like, "Oh, my god. This is phenomenal." All I'm asking is are we losing something by giving general measures over time when you've got something that used intermittently and you'd like to know when it's actually being used, how effective it is.

DR. KATZ: I'd like to follow up on that point, and then get back to Angela's point about measuring loss of therapeutic benefit because I

Page 254

think there's more to say about that.

Something that I've been doing a lot in the last couple of years is convincing sponsors to bolt a qualitative research component on the back of a randomized-controlled trial. So you do your randomized-controlled trial, everyone's filling out forms, and obviously what the patients fill out is limited by your imagination about what forms to give them. But then at their termination visit, have a semi-structured interview, where we ask the patients questions like, "How was that treatment for you?" and with a crossover study, "Why did you like treatment A better then treatment B?" or "What did you think about the way we designed this clinical trial, and could we design this trial in a way that might be more meaningful you?"

People say all sorts of interesting things when you've give them an opportunity to talk and you capture it. So it's qualitative data, but qualitative data is the underpinning of most of what we do. So I wonder if while we're all here, do people think that there could be any value? I

Page 255

don't know that we could recommend it as a must-do for every trial, but do people think there's any value in bringing that research method to light in the context of this paper?

DR. THOMSON: Simon Thomson here. Yes, I think one of the failings of this meeting is we haven't involved the patient voice. I think, one, it should be in our recommendations that any trial should involve some kind of patient-voice consultation, anyway. And then Sam's group, which I'm involved with, just had a qualitative study of a bunch of randomized patients having qualitative interviews. And I think it very much enriches what we do. So as you say, not for every study maybe, but I think it should be one of the recommendations, yeah, I think.

DR. KATZ: Okay Thanks. Any other further thoughts on that qualitative research angle before we get back to Angela's point? Roshini?

MS. JAIN: Roshini, Boston Scientific. I fully agree with the both of you, but just being on the other side of also doing the analysis,

Page 256

sometimes when you keep a lot of open fields, we get stories and stories, and then we're not able to glean out. So I think yes and no questions were fantastic, like would you want to do this again? Would you recommend it to a friend? But when we start going down the why, we do get lots of lengthy stories that we find it hard to discern.

DR. KATZ: I'll be happy to help you with that.

MS. JAIN: Yes, thank you.

DR. KATZ: There are all sorts of qualitative research methods that can be used to digest it, and then at least bring to light what the major themes are. And that's not the whole transcript, but it's the major themes.

Bob?

DR. DWORKIN: This goes back to Angela's question. Bob Dworkin. Nat, I don't remember, but I bet you do, how was time to loss of therapeutic response defined in the randomized withdrawal trials that had that kind of time-to-event endpoint? Because that's in the literature, both

Page 257

for pregabalin and opioids.
DR. KATZ: For those studies, it's been done in different ways, but the most common way is -- just to boil down to one thing would be time to 30 percent worsening. In pain intensity, there are composite approaches, too, so you could treat it as a time to event where if you lose 30 -- you get randomized when you have very little pain because you've responded to treatment, so maybe your pain score's a 3 or something like that.
So if your pain score goes up by 30 percent, or you drop out due to lack of efficacy, or you take the forbidden rescue medication or something like that; so you can imagine composite approaches to that. That's been done also, which brings us back to Angela's point. I wonder whether there is actually more that we can say to people along those lines about loss of therapeutic efficacy.
So obviously, if you drop out of the study because of lack of efficacy, well that's a no-brainer. If your stimulator is explanted because of lack of efficacy, which I guess is more

Page 258

or less the same thing, that's a no-brainer.
In terms of peculiarities of the spinal cord stimulation world, if the patient is not using their stimulator anymore and the reason is lack of efficacy, I guess that would be another spinal cord stimulator specific type of loss of efficacy.
Is there anything else that comes to mind that's specific to this area? Turo? Turo Nurmikko --
DR. NURMIKKO: Turo Nurmikko, yeah.
DR. KATZ: -- from the UK.
DR. NURMIKKO: I'm not sure if this is still valid, but I used to have patients who complained of suspicion of the SCS losing its effect, and I put them on an SCS holiday for a couple of weeks. So you could actually measure that and see if indeed there's an impact or at least that could be one of those measures where you try and define loss of effect.
DR. KATZ: Discontinuation. Great. Anything else on the issue of loss of efficacy? Sam?

Page 259

DR. ELDABE: Not particularly on the loss of efficacy, but on Dr. Fields; point, I think he makes a very good question about dosage. And the reality is that rather depends, as Dr. Loeser says, on what instructions you give the patient before they leave your clinic. However, all of this can be measured because all devices will give you a printout of when it was switched on and when it was switched off. This is not reported anywhere.
DR. KATZ: What would be the best way to present that as an outcome measure?
DR. ELDABE: Good point.
DR. TAYLOR: It's effectively a process outcome here, isn't it? I would see it -- again, it's sort of going back to the fidelity of the intervention, really, isn't it? So I'd see it as just being a subheading within that.
DR. KATZ: And what number would you report?
DR. ELDABE: I think you can report the percentage of hours used. So if the person has had the device implanted 100 hours, they've used it 90 hours, you can say 90 percent usage. Again, it

Page 260

ties into what instructions you give them because you might turn around and say, well, I want you to use this 20 percent of the time.
DR. NORTH: Rick North. The implanted pulse generators routinely will display through the program total therapy hours, and that will tell you what percentage of the time, since the last interrogation, the device has been up and running.
DR. KATZ: So it's essentially a percentage of your 24-hour day that you are using the stimulator.
DR. NORTH: Yes. To Howard's point, back in '77, in that big, long report, I did an APL. There were several parameters like that, that we looked at, pertinent to designing a device. What controls did the patient need to have? How often would they be making adjustments? So there is literature out there, but it's not much reported.
DR. THOMSON: Simon Thomson. Some of this business about the eye in the sky, as it were, or the spy in the can, some of the companies are much more sophisticated than others in that regard, so

Page 261

there are some that not only give you times of usage, now they'll be imputing pain scores and looking at the different modes of stimulation and what the patients put in as their own pain score. So this is very much one of those things. I think there might be a minimal dataset that we should be talking about maybe as an outcome, but reference the fact that this is going to get more and more sophisticated.

DR. NORTH: Simon, as you pointed out on a sidebar moments ago, these things need to be validated, too.

DR. KATZ: Great. Let's maybe channel the conversation on a different issue that I think, Simon, you and others have brought up already. There are rechargeable and fully implanted spinal cord stimulators. Does that have any impact on our research design recommendations in any way, which type of system is being utilized?

DR. THOMSON: Simon Thomson. I think the duration of a randomized study is never going to go more than two years. So in other words, I don't

Page 262

think it makes much difference, except for the fact that there is the recharging burden. And that can't be trivialized because for some it's a daily thing. That can be a cause of giving up on the treatment for some people. I suppose I'm saying, yes, it does make a difference, but not because of device longevity.

DR. KATZ: And what are we measuring there exactly in terms of this recharging burden?

DR. THOMSON: Well, I think I said earlier, it's like a recharging interval. What else would you say?

DR. ELDABE: Sam Eldabe. You measure the duration of time it takes a person to recharge the device, and the recharge interval means the number of days between one recharge and another, and the strength of the coupling sometimes is displayed.

DR. NORTH: Rick North again. One of the companies has introduced an externally powered passive device, like the radio frequency devices that I grew up with and used in preference to primary cell devices for many years. And that

Page 263

changes the nature of trials, potentially, in a big way. It in principle eliminates the need for battery replacement because they're no [indiscernible] components.

DR. KATZ: Is there anything else that we would want to alert people doing studies using that device, that they ought to be considering when they design and carry out such studies?

DR. NORTH: Well, the main one that comes to mind is that it lacks a pulse generator and can be put in and tunneled as a trial device. But because nothing emerges through the skin, when the patient comes back and says, "Doctor, this is working just fine; can't I keep it?" the answer is finally yes.

DR. HAYEK: The one downfall of this therapy is that the patient has to wear an external transmitter at all times when the patient desires the therapy, which may be not convenient if the patient is swimming or sleeping in bed. Sometimes there are certain limitations to that therapy.

DR. THOMSON: Simon Thomson here. I think there is this kind of patient satisfaction with the

Page 264

powering of the system measure because that would be one, having to wear an external cloak in order to provide energy to your implanted lead, and then the difficulties some patients can get in with centering their recharger over the implantable device. So yeah, there needs to be some kind of satisfaction measure with just the recharging component.

DR. KATZ: Sorry, Angela, just one second. Has anybody ever developed any patient satisfaction instruments that are specific to spinal cord stimulation?

(No response.)

DR. KATZ: No. Okay. Andrea?

DR. TRESCOT: Andrea Trescot, Stimwave. There are a couple of issues. One is that we have been able to show a very prolonged effect from short-term stimulation, so patients can wear this for an hour or two, and they get long-term relief.

The second is that when they're doing the recharging, some of the systems are requiring 2 and 3 hours of recharging sitting in a chair or

Page 265

connected to a wall socket, where with this external device, they're wearing it in their clothes; as long as they have their clothes on, they're getting stimulation.

So yes, having also grown up with the old ANS RF receivers, there are some issues, but if you compare -- I think what will be interesting is comparing the rechargeable IPG, not the non-rechargeable. The non-rechargeable is a whole different issue. And actually in Alaska, I put in a lot of non-rechargeables because I have patients who live in dry cabins. They don't have running water. They don't have electricity. They want to go off moose hunting for two weeks. So for those patients, it's most important for them not to have to be hooked up daily to a charging station.

DR. LOESER: There's a new outcome measure.

(Laughter.)

MALE VOICE: Moose hunting.

(Crosstalk.)

DR. TRESCOT: Are you able to moose hunt? So yes, there is a mindset that has to change, but

Page 266

I would argue that most of us don't pay attention to the time our patient has to be spending recharging their systems. We've sort of ignored that because it's not something that we're seeing. But I think being able to develop some sort of patient satisfaction would be huge because when you're comparing a rechargeable system where you have to physically sit in the chair for that period of time, or you have to physically stay hooked up to the wall unit for a period of time, it is something that we don't really tell the patients would be going on.

DR. KATZ: It's interesting that no one's tried to turn that into a measure. With medications, there's this measure, for example, called the TSQM, the Treatment Satisfaction Questionnaire for Medications, and it's like a 14-item questionnaire that gets at how convenient it is to take your medication, and how much it hurts when you get injected, and how big the pills are. It gets at these patient-centric issues that patients care a lot about, but you're not going to

Page 267

capture your 6 IMMPACT core outcome domains.

Roshini?

MS. JAIN: Roshini, Boston Scientific. For our studies, typically we've adapted the TSQM.

DR. KATZ: Oh, okay.

MS. JAIN: Of course, it's not validated because we've now modified it specifically for devices, but we have an adapted version of TSQM that we use as you're describing.

DR. KATZ: Have you thought about cleaning it up and putting it out in the literature?

MS. JAIN: Roshini, Boston Scientific. I have not, but we'll do so.

DR. KATZ: Maybe that will be an organized way of capturing the patient's perception of these issues that seem to be -- if I were hunting moose, I wouldn't want to have to plug myself into a tree --

(Laughter.)

DR. KATZ: -- and wait for a day or whatever.

Great. So that gets at patient

Page 268

satisfaction. That gets up at rechargeable versus non-rechargeable issue. Let's now rise back up to the 50,000-foot level.

Are there any other sections -- can anyone think of a section of this recommendations checklist that's been left out?

DR. THOMSON: Simon Thomson, and I'm not sure whether this is -- but patient-related outcome measures, I'm not sure whether we've covered that well enough, and I'm not quite sure whether the IMMPACT thing covers it quite well enough, you know, expectations met.

DR. KATZ: So do you mean measurement of outcome?

DR. THOMSON: Yeah. So I'm back on outcome, and I know this isn't what you wanted. But I just feel we haven't completed that yet.

DR. KATZ: So I think where my thoughts are about that right now is that we have the IMMPACT core outcome domains for clinical trials. Then there was a subsequent impact publication that recommended some measures of those domains. I

Page 269

think we can just highlight the fact that those exist. I don't think we need to relist all those measures here.

Then we have some spinal cord stimulator specific patient-reported outcome measures that I think I have a little inventory of that I've been taking notes on all along the way. And I think that's where it sits right now.

Dennis?

DR. TURK: There is an IMMPACT paper that reviewed all the physical function measures, both from patient satisfaction, to family member responses, to physical activity. Ann Taylor was the first author; that it was about two or three years ago. I can't remember which impact meeting it was. So when you get to the physical function, if you want to get a little more specific, then just assess physical function.

We did review all of those, and importantly we made a distinction between self-report measures of physical function, performance measures of physical function, and clinician or clinic-based

Page 270

physical function because we have subsequently found that those don't correlate all that well. So what patients tell you they can do, what they can actually do and what they do in the clinic don't necessarily tell us the same thing.

DR. KATZ: Dennis, who was the first author of that again?

DR. TURK: Ann Taylor.

DR. DWORKIN: And there are also IMMPACT articles on phenotyping. Rob Edwards is the first author, and on biomarkers with Shannon Smith as the first author.

DR. KATZ: Simon, was there anything else that you think we've left out here?

DR. THOMSON: Are we at 50,000 feet?

DR. KATZ: Yes.

DR. THOMSON: All right.

DR. KATZ: Categories.

DR. THOMSON: Categories.

DR. TRESCOT: I'll take [inaudible - off mic].

DR. KATZ: Can you speak into your mic, please?

Page 271

DR. TRESCOT: Just a joke. And we'll take history for 300.

(Laughter.)

DR. KATZ: Okay. So I'm not hearing anyone thinking that there are some major chunks of this research recommendations checklist that we've left out, so we can dive down to 25,000 feet now. We've got all the sections mapped out. Are there any specific elements of this? I don't want to rehash what we've been discussing over the last 36 hours.

Are there any specific -- and I will look at your 200-item checklist. Are there any specific recommendations for spinal cord stimulator trials and chronic pain that have not been mentioned yet that should be in this paper? Jane?

MS. SHIPLEY: Jane Shipley. This is a small thing maybe, but I'm changing my system. I'm not saying inclusion criteria and exclusion criteria anymore; just patient selection criteria because too often a study will say something in a positive way, and then say the same thing in a negative way,

Page 272

and it's just plain stupid. So it should just be -- that's a small thing.

DR. KATZ: That's a great point. I run into that issue all the time. Excellent.

Yes, Eric?

DR. BUCHSER: Eric Buchser. Something that's been diluted before, you can derive from the parameter that you're using: frequency, voltage, or intensity, and so on. And that would be the charge per second because that could be something that you could mention without people having to work it out from the data they have. The charge per second should actually be mentioned I think.

DR. KATZ: Charge per second

DR. BUCHSER: Charge per second.

DR. KATZ: Great. Anything else, particularly if it's spinal cord stimulator specific? Jane? Jane Shipley from Baltimore.

MS. SHIPLEY: Jane Shipley from Baltimore. And I would say this should be for everybody. Now my brain just died because you said "from Baltimore."

Page 273

I think Sam brought this up. Too often, people mix up methods and results. One of the things I'm trying to do is say follow-up duration planned, and then later on in methods, say follow-up duration achieved.

MALE VOICE: That's result.

MS. SHIPLEY: I'm sorry, and result. So method is follow-up duration plan and then achieved. So I'm trying to push people to see that there's an actual distinction, so they're not putting results right in the methods. And there are other places we can do that: programming parameters planned, programming parameters achieved. Papers are a mess right now, in general. The literature's a mess.

DR. KATZ: Great. I want to revert back a little bit to our recommendations on reporting of safety. We did have a robust discussion of that, really, throughout this meeting, but I want to see if I can bring that down to what we think we're actually going to recommend.

So capturing adverse events, we don't need

Page 274

to take up real estate in this paper about that. It's a regulatory and universal requirement. Now, whether there are any habitual gaps in how that's reported in the spinal cord stimulator literature, that might be worth attending to.

But what I really want to get at now is, is there any template, for lack of a better word, any specific list -- maybe the right way to ask this question is, is there a list of adverse events of special interest -- that's how I would think about it in a drug trial -- where we can offer a standardized set of terminology for adverse events that occur, that are particular to spinal cord stimulation, so people aren't using five different words to mean the same thing or the same words to mean different things?

Is this something that would be helpful? Is there something like this out there? What do people think about that? Salim?

DR. HAYEK: Salim Hayek, Cleveland. The devil is in the details. You can make it as extensive as you want. For example, lead

Page 275

migration, which is a very common occurrence, especially with percutaneous leads, can be clinically meaningful or clinically meaningless. If a lead moves 2 or 3 millimeters, with current stimulator devices, most of the time you're able to recapture parasthesia if you're looking for parasthesia, or pain relief if it's parasthesia free, without the need for a revision.

So I think from a technical standpoint of complications, the ones that are clinically meaningful are the ones that result in revision or loss of therapeutic efficacy. The biological complications I think should be recorded as far as infection. Again, if it's a deep infection, it almost invariably results in an explant. Related to that is adherence to previous guidelines that that have been published, including the NAC guidelines. Many of the people here were co-authors on these.

So I guess you can be as extensive as possible, but there are certain ones that we should not miss in putting out a list, and I'll be happy

Page 276

to help with that.

DR. KATZ: Are there guidelines in the literature now for how to report adverse events specific to spinal cord stimulation?

MALE VOICE: No, not that I know of.

DR. KATZ: Would that be useful? And I'm not talking about a 50-page guideline inside of a 5-page paper, but would it be useful even to provide a little table that suggests a standardized terminology for these particular sorts of adverse events?

DR. NORTH: I would say there are guidelines. Salim's paper, the Tim Deer intact paper on complications, and Tracy Cameron's paper, all use the same basic scheme: biological, technical, et cetera, although not published as guidelines. And WikiStim has a corresponding list.

DR. THOMSON: Simon Thomson here. There's relatedness to the device, relatedness to the procedure. Then there is the need for reoperation. We struggle when we're doing studies because being hospitalized in reoperation comes out as an SAE,

Page 277

whereas for us it's kind of a routine thing; okay, lead migrations. It's not a disaster but they need a reoperation. It's not really an SAE. So that would be helpful to the field. It needs to be recorded, but at least to take it down from being an SAE.

Sam?

DR. ELDABE: A couple of points. I think the SAE and AE classification is not ours to change. This is regulatory, and that remains where it is. It would be useful, as you say, to produce a list of stimulator-specific complications that we would like studies to report on specifically and in detail. The list is available in many publications. As Rick mentioned, there are a number of publications that are guidelines, but these guidelines are how to avoid complications. They're not guidelines on how to report them, because once you start issuing guidelines on how to report something, you have to define it. And none of these guidelines, to my mind, define what it is that we mean by lead migration.

Page 278

DR. KATZ: Rod?

DR. TAYLOR: So just on this theme, and I guess as a generic comment, at the end of the day, you're in the chair, Nate. But I think one of the things we talked about over a beer last night was the CONSORT guidelines for reporting are quite useful because they often give you exemplars of good practice. They actually verbatim take bits of text out of papers and say this is how to do it.

So I guess I might encourage that within the fight that it needs to be a manuscript that will fit in a journal. But I think illustrating, I find those as a trialist hugely helpful, not just saying do this, but here's an example of how you could do it well. And I think there are some examples of good practice here. So where they are there, again, I would just like to encourage we sign post, and obviously adverse events would be maybe one of those areas we've got some good practice.

DR. NORTH: Rick North. The replacement of a primary cell implant is not generally reported as a complication. It's more like an expected end of

Page 279

battery life. But I think it deserves consideration, especially when there are externally powered alternatives, including rechargeables. So I wouldn't just sweep that under the rug and call it routine maintenance.

DR. HAYEK: In some of the papers, it is reported as a complication, but you're right; it's something that is expected. However, if replacing the battery leads to an infection and an explant of the whole system, then it's also a potential complication because of that.

DR. KATZ: Salim, how about if you help me come up with a table that lists a recommended -- preferably copies and paste, but these things always require some kind of clean-up. You think it's there and you can copy and paste it, and then you realize that there's something not quite applicable. So how about if you help me come up with that table?

DR. HAYEK: Happy to do so and to share with everyone who has to edit it. But I'll circulate a draft.

Page 280

DR. KATZ: Oh, yes. Everyone will certainly have their thoughts about it, I'm sure.

DR. SINGH: Rahul from MHRA, London. I think this is a problem throughout all research fields, whether it's oncology, orthopedics, whatever you want to say. Competent authorities actually have codes for adverse event reporting, which are usually adhered to when they're doing the clinical trials. But the issue is when a device is already CE marked and when clinicians do their own prospective, or even retrospective, studies that they want to publish, they don't adhere to these, or they don't know about these codes that are present.

So if we did want to harmonize and have a common language, that would be good starting point. These codes are jargons to me, to be fair. They're numbers and letters and stuff, but if you follow the pathway of those codes, they do come down to a comment like dislocation. You can have a dislocation of a total hip replacement, or a dislocation or migration of a spinal cord

Page 281

stimulator, for example. So that's one way of looking at it as well I think.
DR. KATZ: So I'll take that as an offer to look at whatever we come up with and provide comments.
(Laughter.)
DR. KATZ: Thank you for volunteering.
DR. FIORE: It's Greg Fiore. I should probably also volunteer for that. I spent most of my career in pharmacovigilance and safety, and might as well put it to use.
DR. KATZ: You're hired. Great.
Great. Any other comments about this adverse event tracking? It sounds like we've agreed that it would be helpful to provide some kind of rubric. We've got a plan for working on it. Any other thoughts about that issue?
DR. THOMSON: Simon Thomson here. There are going to be some things. I mean, we just touched on it talking about non-rechargeable; and is it after 5 years; is that an adverse event, that you have to change it? And there's an argument that it

Page 282

could be because you could have used a rechargeable system.
But what if it didn't last 5 years, the premature exhaustion? Now that to me is an adverse event. But what is that? Is that 12 months? Is that 2 years? Probably 2 years I would have thought. Those sorts of things, we could usefully define.
DR. KATZ: Yes, that seems really important to me because if you have to go to the operating room, that's part of the burden of therapy, whether it's expected or not. It reminds me almost of drug side effects, let's say opioid-induced constipation. Well, if you have to take a $300 a month medication to reduce your constipation in half, but it's still a problem for you, yeah, it might be expected, but it's still a burden of therapy. So it seems like it ought to be tracked in some way.
DR. HAYEK: Salim Hayek from Cleveland. Even among rechargeable devices, some of them are warranted

Page 283

for one year. Some of them commit suicide at 9 years. Some of them go for 20 years. So should we say that the 9-year one is a complication compared to the one that goes for 20 years? There's a lot of gray zones in there.
DR. KATZ: All right. Well, luckily we don't have to solve every problem, but if we can at least advance the agenda one step and provide a suggested template, then maybe that'll be helpful.
Roshini?
MS. JAIN: Roshini, Boston Scientific. Just two comments. Device complications, and especially those that result in an adverse event, just to be able to cull those out because each of them are different and significant. The second thing was, is there any thought on standardizing this with the MAUDE database that's out there? Just a thought.
DR. KATZ: Yeah. That's a good question. Roshini, I think you're referring to the study we just did on the MAUDE. We just did a study on the MAUDE database, looking at spinal cord stimulator complications, and we spent a lot of time trying to

Page 284

make heads or tails out of what's in that database, which was not an easy task, and which left open a lot of questions.
But maybe once Salim comes up with his suggested template, I can cross-check that with how the MAUDE database is, what the codes are, what the dropdown fields are the MAUDE database and make sure at least we try to pay attention to relating the two. Yeah, that's a great idea.
MR. BOSLEY: Bernie Bosley from Nuvectra. I think this is also important to standardize this or at least communicate it more because the adverse events I see, some of them are defined in terms of harm; some are cause of harm. They're kind of mixed up a little bit. In Europe coming up, we're going to be asked to do active postmarket surveillance, and if the definitions in postmarket aren't the same as clinical, we're going to have trouble there, too.
DR. KATZ: It's starting to feel to me like, although I don't want to -- it's starting to feel to me like this could be a separate project,

Page 285

figuring out what the classification system would be of adverse events in this area, and making sure that it makes sense on one side of the Atlantic, the other side of the Atlantic, and regulatory and clinical. That's not going to be part of this paper, but I don't know.

Do others think that that would be a useful side project?

DR. HAYEK: Salim Hayek. Should we assign numbers or points for severity of a complication, for example, if it's clinically meaningful or significant? Surgical revision, it gets higher points or is more serious than if it's a potential nuisance like pain over the generator -- or discomfort over the generator site; one that leads to surgery and one that doesn't lead to surgery.

DR. KATZ: What we have now is you can be classified as an SAE. If it requires hospitalization, medical intervention to prevent serious harm, that's got its own list of regulatory definitions, so that obviously will need to be captured; it's a requirement. And then for adverse

Page 286

events that are not SAEs, we have regulatory definitions for what counts as mild, moderate, and severe, and every adverse events should be coded that way already.

Do we need to go beyond that for this paper?

DR. HAYEK: Is outpatient revision surgery of a generator a hospitalization? Would that be considered a hospitalization? Because the patient is not admitted in the hospital; that's an outpatient procedure.

DR. FIORE: This is Greg Fiore. Nate, the way procedures are handled in the MedDRA coding schema that was referred to earlier is it's the diagnosis. The condition that led to the procedure is what is the adverse event, and the procedures can be captured separately. So if this group would want to recommend that investigators prespecify procedures that might be performed and list those not as adverse events but as procedures for treatment or for revision, that might be worthwhile.

DR. KATZ: Greg, do you have a list of the

Page 287

MedDRA codes that are relevant to spinal cord stimulator complications? Actually, does MedDRA even apply here? Because it's supposed to be for drug adverse events. Isn't that what the D stands for?

DR. FIORE: Dictionary for Drug Regulatory -- yeah.

DR. KATZ: Drug, right?

DR. FIORE: Yeah.

MS. JAIN: Roshini, Boston Scientific. Yes, we use MedDRA coding.

DR. KATZ: You do?

MS. JAIN: Just because it's standardized as well.

DR. KATZ: So you have a list of MedDRA codes that are applicable to this situation, so that once Salim comes up with his list, we can at least see where there's any relationship? Are there more than 10?

MALE VOICE: [Inaudible - off mic].

DR. KATZ: That's what I was afraid of.

(Laughter.)

Page 288

DR. KATZ: All right. I'm willing to receive it. Let's put it that way.

Greg?

DR. FIORE: One more point on that if I can. Greg Fiore. What's very useful when analyzing safety information is that if somebody took the time up front to map the terms to categories that are useful for the indication. The dictionary has -- someone here probably knows -- hundreds of thousands of terms in it, and mapping up front. And then even letting investigators choose from a prespecified list is always helpful. Health authorities like it, and it allows us to really achieve meaning, because sometimes the terms are pretty esoteric.

DR. KATZ: That's actually a really good point. With MedDRA, you could have the same event and see 6 different MedDRA preferred terms that it's coded to, which makes life very confusing for everybody. I don't know. I don't see us doing that mapping procedure in this paper; at least I don't see myself doing it, but let's see how far we get.

Page 289

All right. Any other comments on the issue of safety for now?
(No response.)
DR. KATZ: All right. Well, are there any topics, any recommendations that anybody feels should be included in this list of recommendations that have not been mentioned yesterday or today?
DR. SINGH: Rahul Singh, MHRA. We're talking about -- I understand the scope of this forum, but is there anyone innovating or trying to amend, make changes to the spinal cord stimulators to improve outcomes and reduce adverse events? Is there anyone that you know of, apart from the manufacturers -- obviously, they've got their own R&D departments who are trying to excel in that area, possibly.
Is there anyone in that area doing that that you're aware of?
DR. KATZ: You mean in R&D not at a manufacturer that's trying to improve upon the technology?
DR. SINGH: Yes.

Page 290

DR. NORTH: Yes, anonymous.
DR. KATZ: Anonymous. Yes, anonymous.
Anyone else care to expand on that, on Rahul's question? Any other units anybody knows about that's working on this outside of manufacturing?
DR. NORTH: Don't spill your IP.
DR. HAYEK: Salim Hayek, Cleveland. What altitude are we at? You said any other questions, any other comments. At what level? How granular are we getting into? I mean, are we talking about inclusion/exclusion criteria for studies?
DR. KATZ: I think we can be very granular now, again, recognizing that we're not trying to design every imaginable study now, but if anybody feels that there is a general recommendation that would apply across the board in spinal cord stimulators studies for chronic pain, then now would be the time to articulate it.
DR. THOMSON: Simon Thomson here. One of the things that always vexes us is age.
DR. KATZ: Age.

Page 291

DR. THOMSON: We generally don't start doing studies in children. Is that 16 or 18? And then there's this top age, and we don't do -- it's becoming politically incorrect to put a top age. But obviously, the top age has comorbidities.
Is our recommendation that you have a top age or do you just say stuff around cognitive ability, and comorbidities that would make relevant outcome measures difficult? These are the sorts of things.
DR. KATZ: Of course that comes up in virtually every clinical trial. Does anybody feel like we should be advancing in this paper a recommendation for a top age for spinal cord stimulator studies?
(No audible response.)
DR. KATZ: I'm seeing heads shake in this sideways direction, so it seems like people feel like that should be up to the designer to struggle with that.
DR. THOMSON: The other issues are around pregnancy. IN other words, we wouldn't go out of

Page 292

our way -- we're not going to recruit pregnant people, but what happens if they become pregnant during the trial, and how do we manage that?
DR. KATZ: Is that a research design issue or is that an issue of how to take care of patients once they're in, if they're in clinical trials when things happen to them?
Research design issue. Okay. Would anybody be prepared to articulate a recommendation for how to deal with pregnancy occurring during a clinical trial in spinal cord stimulation? Should we be doing that?
DR. THOMSON: Potential for pregnancy?
DR. NORTH: Rick North. I'd say not at this level. I've been involved in cases where it's come up, patients have gotten pregnant, during pregnancy, and the pragmatic recommendation is, well, what are the alternatives for managing your pain, and has safety or efficacy been proven for them. It hasn't been for the stimulator.
DR. KATZ: Yeah. Andrea, you had a comment on that?

Page 293

DR. TRESCOT: Andrea Trescot. I was going to recommend that the criteria be that if women are of childbearing age or in the study, that there's a survey being done asking them to avoid getting pregnant for the period of the study because of the unknown effects of stimulation on the baby, and the potential that they would be dropped from the study if they become pregnant because I think the confounding issues of it can be a big problem.

Again, as my understanding, there is currently no registry for women who become pregnant during and have a spinal cord stimulator in place. I don't know of that registry. We had it with sumatriptan to monitor whether there was a problem.

DR. KATZ: Well, here's my sense of the room right now. It feels like we're kind of done. You sort of get the feeling after a while, all the big issues have been covered, and to try to spend time specifically to fill the time with more words, that doesn't feel like it has a huge amount of value to me.

So it's 2:45-ish, 2:48. So I feel that now

Page 294

is the time where you can spend a couple of minutes, if anybody feels that we've missed anything in terms of our big-picture plans, in terms of the content of this paper -- obviously, everybody's got plenty of time to look at revisions of this paper and come up with all of the most nuanced thoughts and recommendations that will come up. It doesn't have to be today. There will be plenty of shots on goal here.

Does anybody feel like we've missed anything big in terms of the content of this paper?

DR. THOMSON: We've spent a lot of time over the last 36 hours talking about blinding and single blind and double blind. We're presumably going to be making a recommendation position on that, aren't we?

DR. KATZ: Yes.

DR. THOMSON: And all the things we've talked about, we're going to be --

DR. KATZ: Yes. It seems like the wisdom was to not make a strong recommendation that there's only one way of dealing with that, but just

Page 295

to outline a number of different options, what their pros and cons are, and then let the protocol designer look at those and provide a rationale in terms of why they chose one of those or chose not to. And I feel like I've detailed enough notes to at least write a draft and get people's feedback on it.

Anything else that anybody thinks? Bob?

DR. DWORKIN: This is minor, but I'm going to ask the group's permission for us to depart from what's been a kind of unspoken IMMPACT-ACTTION policy. And that policy has been that everybody who attends both days of the meeting is invited to be a co-author on the manuscript that Nate is going to be drafting. And really, the model has been that people spend the bulk of their time; they participate in the bulk of the meeting.

I'd like to ask for permission for an exception, which is that we invite our two colleagues from CDRH, who were here only for an hour and a half yesterday, to be authors on this manuscript. I think, for all sorts of reasons that

Page 296

don't have to be said out loud, it would be very cool to have them involved in this process.

Does anyone disagree with that exception to IMMPACT and ACTTION's policy, that we invite our two colleagues from CDRH to be authors? They of course can decline.

(Affirmative nods of heads.)

DR. KATZ: The heads seem to be seem to be going up and down.

MALE VOICE: I think that's a good idea.

DR. DWORKIN: So let's include them in the process, and if they withdraw, that's up to them.

DR. KATZ: That's a great idea. Great.

Anything else? Bob? Dennis? Do you guys want to make any closing comments of any type?

DR. DWORKIN: My closing comment is simply thank you very, very much, Nate, for doing a masterful job of sewing this altogether.

DR. KATZ: Well, thank you guys for making it easy.

(Applause.)

DR. TURK: From ACTTION and IMMPACT, I want

Page 297

to thank our collaborators from the neuromodulation societies for their contributions. Those that don't know the background, there are lots of discussions and conversations going back for 6 months, 8 months, a year, trying to set this up. There's a tremendous amount of work, and to also thank all the speakers who took the time, and to request that you'll be asked to allow us to put your slides up on our website.

If there are any slides within your set that are proprietary you don't want, you can remove those. But to the extent that you'll let -- people out there who might be interested in this meeting who aren't here would find those very interesting. So you will get an invitation or request to allow us to use your slides.

Adjournment

DR. KATZ: I'd like to finally thank Rahul from MHRA for coming because, first of all, you came a super long way, as did many others. And second of all, it's so critical for us to have people here who can provide a regulatory

Page 298

perspective, so much appreciated.

Alright. Well, happy trails, safe travels everyone. See you next time.

(Whereupon, at 2:52 p.m., the meeting was adjourned.)

B

C

E

F

I

M

N

O

S

T

U

V

W