IMMPACT XVIII - Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials

---

June 5, 2015

A Matter of Record
(301) 890-4188

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

IMMPACT XVIII

Ensuring Data Quality in Clinical Trials of Pain Treatments
Considerations for Study and Conduct

Friday, June 5, 2015

Willard Hotel
Washington, DC

C O N T E N T S

P R O C E E D I N G S

MALE SPEAKER: Good morning to everyone. Just a couple reminders to you before we start formally, the housekeeping things to keep in mind. Please remember to speak into the microphone when you want to be asking a question for those in the audience. Make sure your cell phones are, in fact, silenced.

Checkout time from your room, just so you know it, is 12:00 o'clock. If you haven't checked out, you can do it at the coffee break or at lunchtime. We will try and end the meeting within a reasonable time.

Remember what your mission is for the rest of this session, rest of today, which is at the end of the day before we let you out the door, we are going to begin talking about hopefully putting some information together that Bob is going to summarize for you, and then start working towards having this recommendation paper, considerations for improving data quality based on the conversations that we've had here.

MALE SPEAKER: And we submit it to the journal.

MALE SPEAKER: To be submitted to a reputable journal unspecified at this particular moment.

(Laughter.)

MALE SPEAKER: Although she's not in the room, I want to also remind you about taxis because it's Friday afternoon and they want to make sure there's enough taxis, to check at the -- if you haven't already done so, what time you're going to be needing a taxi so that, in fact, can be taken care of by -- and Valorie is doing that.

She's not in here, and I want to just thank Valorie Thompson and Andrea Speckin, who were the two people who coordinated this meeting, did all the correspondence with you, got you the information, took care of all the logistics

I think from our experience, they've been extremely helpful, very effective in doing that. Hopefully, you've all had a reasonable experience in getting here.

Page 5

(Applause.)
MALE SPEAKER: If you have any questions or comments regarding your trip back or checking out or what have you, definitely check with them. They'll be able to help you.
So let me turn it over to Mike, who's going to finish off the session that we had begun, and we were slightly off target.
I want to thank Paul for being willing to be flexible on the timing.
DR. McDERMOTT: Okay. It's my pleasure to introduce Paul Schuette, who is a mathematical statistician and the scientific computing coordinator at the FDA Center for Drug Evaluation and Research. He's going to give the FDA perspective on, as you can see, contents data, quality issues in the design and analysis of trials.
Presentation – Paul Schuette
DR. SCHUETTE: The standard disclaimer, if you don't like what I say, blame me, not the people I work for. A little bit of an outline. We'll

Page 6

talk about data quality, and analysis quality to some extent, reviewer experiences, some monitoring and some conclusions.
So data quality. I think it's pretty much a given, it's been accepted, that we cannot inspect our way to quality. According to Deming, "Eliminate the need for inspection on a mass basis by building quality into the product in the first place." So Deming is perhaps the quality guru from the '50s, '60s, '70s and '80s.
This is one of his 14 points. This is actually reflected in one of our guidance documents. "Monitoring or oversight alone cannot ensure quality. Rather, quality is an overarching objective that must be built into the clinical trial enterprise. FDA recommends a quality risk management approach to clinical trials."
Let me tell you where that came from because that's an important document, which is "Oversight of Clinical Investigations: A Risk-Based Approach to Monitoring."
I would say that FDA has embraced the

Page 7

quality by design paradigm. It was mentioned by Nat that this occurs typically in the manufacturing side of the house, and there's actually an FDA guidance that spells that out. I would also argue that what we're attempting to do with the good XP guidances, the GXP acronyms that we have all over the place, are partially an attempt to correspond to this as well.
GCP most people know, good clinical practice. GMP is good manufacturing practice. How about GPVP? Good pharmacovigilance practice. And does anyone want to take a stab at GLP?
MALE SPEAKER: [Inaudible].
DR. SCHUETTE: Yeah. Government tends to specialize in TLAs, three-letter acronyms.
(Laughter.)
DR. SCHUETTE: So there will be a test. CDISC. Let's see. CDISC stands for Clinical Data Interchange Standards Consortium. So FDA, some of our other regulatory agencies, and representatives from both academe and sponsors have worked to try to develop data standards.

Page 8

The data standards are not perfect. They don't answer every potential issue, but they're at least a step in the right direction. And I think we can say, okay, you're missing this data. It's a lot more obvious when we all have to report the data the same way. So those are out there.
CDISC has published a therapeutic area standard for -- I'll say it's version 1. I'm not saying it's perfect again, but it's at least an attempt in the right direction. And I would also argue that the very fact that we have a prespecified statistical analysis plan is in some sense related to data quality and analysis quality.
Statistical quality concerns. Missing data. This is perhaps one of the big ones nowadays. There's a National Academy of Sciences report that is certainly something to look at. There is an EMA report. There is an FDA guidance in development, I am told, and all dealing with that.
The basic approach that I think we're saying to missing data is don't. Avoid missing data, and part of that is looking at how things are designed.

Page 9

Do the study design and study conduct minimize missing data? How do the protocol and the statistical analysis plan propose dealing with analysis of missing data? So we would like to have that more specified as we go along.

By the way, last observation carried forward is not considered a good way to handle missing data. So there's more there.

Another thing that has come out, and this is a little bit older now, are the patient-reported outcomes guidance. One of the things that we do see as an issue, and it's sometimes a problem, is specifying the choice of instrument. There should be some background as to why a particular instrument has been chosen. And along with that, there needs to be sort of a complete document trail, audit trail that is available, that specifies the version number, scoring algorithm, and so forth.

Sometimes it can be very difficult to replicate results or know what's going on. And unfortunately, the choice of instrument even if you

Page 10

say it's there, there is some concern as to whether or not the version is changing over the course of a clinical trial.

So I've used a term from the modeling and simulation world, at the very last bullet is, for lack of a better term, verification, validation, and uncertainty quantification. Basically, this is called content validation and other things in the guidance, but the overall idea is does the instrument -- does it do what it says it's supposed to do, and are the results reliable? And there's been some allusion to some of those types of issues.

So data quality and the FDA submission process, let's go through some of this. So suppose a sponsor finishes their study. This is kind of what actually happens on our end. They submit an application to the electronic documents staff. It goes into our systems, and then we start looking at it.

The review teams must determine whether or not the submission is actually what we call

Page 11

fileable. This comes fairly early on. But unfortunately, this is a fairly rudimentary process. It's basically are the appropriate domains populated, is there a demographics domain, is there this other domain?

It doesn't say anything about how great the data is once it's in there. It just says is it there. So that's why I call it rudimentary checks. We're trying to put in place incrementally some better methods.

Within the CDER -- I'm located in the Office of Translational Sciences. A companion organization within the Office of Translational Sciences is the Office of Computational Sciences. And they actually have worked with CDISC to provide something they're calling a jump start service. And this actually provides some rudimentary checks of SDTM data.

So as I use these acronyms, I assume people kind of know what I'm talking about here, but this is basically, the "raw data" that come out of the CDISC model. And what this does is it checks for

Page 12

things like is the adverse event start date before the adverse event end date, those types of things, very basic types of checks that are necessary but can be missing.

We have Office of Scientific Investigation inspections, OSI. But this is usually a very small proportion of sites. We're talking about 1 to 2 percent in many cases, so not a huge amount. And data quality issues can emerge throughout the review process. So we'll follow this with some anecdotes.

This should look vaguely familiar because I think this is the exact same instance that Sharon alluded to yesterday as her first example. Reviewer reported an incident in which several members of the same family were all enrolled in a pain medication trial on a Friday evening. My understanding is the dog did it.

This raised some red flags. There were found to be some other questionable practices at this site. Turns out this was also the largest site in the trial. OSI is concerned with the

Page 13

validity of data from the site. The entire trial was excluded, and if the sponsor wants to pursue this, they have to submit new studies. So this is a fairly serious problem for the actual sponsor.

Another experience, misclassification. Rescue medications were misclassified as concomitant medications, affecting some domains, and really changed the efficacy evaluation of the product because we're looking at a combination product, in essence, rather than the actual product itself.

So just because someone says that they're employing standards doesn't mean that they are actually doing so correctly, and we need to have the standards employed in the right manner in order to be effective.

PRO. One of my division directors says that we should always have a cartoon in a presentation. But there is a serious aspect there. I would say that the circumstances in who is administering the test can matter, and the context is also important.

I've realized I've gone into government

Page 14

speak. PRO is patient-reported outcomes, of course. One of the challenges is, of course, instrument validation. Perfect validation, of course, doesn't really exist, but we want some literature and background as to why a particular instrument was used, is it fit for a purpose. And some of that is also spelled out in the PRO guidance.

One of the problems that has sometimes been seen is the use of pediatric and adult scales without making a distinction. They're both on a zero to 10 scale, gamish them together, and we're good to go, right? No. We need to look at pediatric being very different from an adult PRO potentially, or if they're going to say they're the same, there needs to be some justification.

Observer-reported outcomes. It may not be the case that a patient can actually report their pain. How do we handle this and how should that be incorporated with the other information is something to consider.

This was alluded to in some of the other

Page 15

speakers yesterday, variability of individual outcomes. Some of this is perhaps related to subject training. From our perspective, we don't know if it's a subject training, an instrument reliability issue. It's all sort of conflated. We just know that we don't think things are changing, and the responses are changing quite a bit. So we saw some challenges being addressed along those lines.

Missing values. One of the standard problems, of course, in all this area is what happens if we have a missing value. Some of that could be related to the choice of the instrument. Sharon was saying, for example, that if we don't have the option to respond in the right way, what do most of us do, is we just stop the survey at that point. So there may be multiple reasons why there's missing values, and we think that needs to be explored further.

Rescue medication. We're looking at analyzing the efficacy and safety of medications, and, if you will, it's necessary to prevent more

Page 16

missing data, but it can also be a complicating factor. So the use of rescue medications for breakthrough pain in both acute and chronic pain trial poses a challenge for efficacy analyses.

There is a draft guidance that has come out. It's listed in the references, but it just came out last year. And they do allude to some of these types of issues, that it is a little bit more complicated when we're actually looking at this type of analysis.

The other thing that we've already seen a problem with are opioids misclassified as concomitant meds rather than rescue medications. And one of the problems is that when we're looking at these rescue medications and looking at different rates and different trials, and we're looking at an overall submission, how do we deal with integrating these results across all the trials in a submission is one of the issues that our reviewers are mentioning.

Some other issues, incorrectly coded AEs. That's actually something that we've seen some

Page 17

problems with. In some sense, that's also a quality control issue.

Correctly ascertaining the recorded reason for withdrawal. For most of us who do work with reviews, this is one of our pet peeves. Lost to follow-up is not a good reason. We need to have better follow-up as to why someone withdrew from a trial. Did they move away? Did they die? Did they experience an adverse event? Was it for lack of efficacy?

Those type of things need to be included as part of the protocol and actually more effort to ascertain what's going on for those purposes. Follow-up with phone calls, reaching out more.

Lab values. This is also a quality control issue. In some cases, we were calling this investigator error. In some cases, we don't know if it's incompetence, ineptitude, not the proper training with the instrumentation. But it does create some problems and issues. And again, missing values, something to harp on is the missing value issue, but pain is one of the areas that

Page 18

there's a higher proportion than some of the others. And one thing that one of our reviewers mentioned was the need for better tools to discover misconduct in errors.

This gets us to monitoring, and we have two basic types, the onsite monitoring, source data validation -- was something that Amy was talking about yesterday -- and centralized monitoring where we're doing a remote evaluation.

We do have a FDA guidance on the topic, and there is a recognition that onsite monitoring is time consuming, expensive, and not always necessary. And we can even add another point is that it doesn't always catch the problem.

So centralized monitoring. Let me quote straight from the guidance. "FDA encourages greater use of centralized monitoring practices where appropriate than has been the case historically with corresponding less emphasis on onsite monitoring." And this might even get into some of the issues that Nat was pointing out yesterday with we see some things being flagged

Page 19

that may not be necessarily quite as relevant for the scientific question under concern.

Centralized monitoring can be an important component of a risk-based monitoring plan, so we're focusing on sort of this risk-based idea. The guidance has some details, but let me outline, the key steps are to identify the critical data and processes; do a risk assessment of those, keeping in mind who will be actually entering the data and those processes; considering risk factors; and also, developing a plan.

Even with the best centralized monitoring, remote evaluation, we still think there will be need for onsite monitoring, at least in some cases. So it's sort of an entire approach, but we think onsite monitoring can be reduced in some cases and targeted more specifically.

Statistics and central monitoring. This will look vaguely familiar from Amy's talk. Distribution of data is one of the things we're looking at. Too much variation, too little variation, outlier, inlier detection.

Page 20

The general trend that we want to look for in some sense are results that are too good to be true, or conversely, they're way, way off scale from everyone else. One of the issues people who fudge data seem to not know how to look at a calendar.

(Laughter.)

DR. SCHUETTE: Maybe that's part of the numeracy training that we were alluding to.

But we want to examine the differences between and within sites, and we're also looking at some ideas from data anomaly detection. The word "fraud" has certain legal connotations, so we'll refer to things like misconduct or data anomaly. And we also need to make the results coherent to non-statisticians or data scientists. So those are some of the issues that are involved.

Here are some of the initiatives that we're starting, and we're not there yet. We're working with companies to bring commercial software into FDA for evaluation, research, and development, particularly for data anomaly detection.

Page 21

Amy alluded to the fact that some of these programs require hundreds if not thousands of individual tests. Many of these are actually going to be simulated. This is a very high -- for lack of a better term, high performance computing environment is needed to actually carry this out for the requisite level that we would like. So we're looking at using our FDA high performance computing environment to actually be able to carry out some of that.

We're also looking to improve the statistical methods to determine some ways we can filter out some of the false positives and false negatives.

We're also looking at improving our existing office of scientific investigation site selection tool. We just brought in -- I'm actually the person that's working on that. We just brought in a graduate student who will be working with us this summer to do a little bit of data mining in terms of looking at the data.

There is a potential for our Janus clinical

Page 22

trials repository. So one of the long-term goals is basically we're trying to develop, essentially what it boils down to, a long-term data warehouse for at least the SDTM data. And depending on funding -- that's kind of gone on a herky-jerky type of fashion. If you're familiar with sometimes riding the Metro, kind of how it lurches and stops and moves forward, that's kind of how Janus has actually proceeded. Sequester did a number on our ability to put that in place.

So we're still trying to get all that involved, and one of the ideas that's been mooted about is to actually include some of these data checks as part of putting the trial data into the repository. So there's potential. We can't say we're there yet. So this is more aspirational than operational.

Do have some conclusions. I think we're making progress, but there is definitely room for improvement. On a lot of things, we're at version 1. And as you know, for any software release, version 1 is not necessarily where you

Page 23

really want it to be.

One of the items that we've talked about perhaps overall, and I've labeled this, is can we better articulate what we mean by good clinical trial practices, good data practices?

What I mean by that is that if we look at sort of the areas by themselves, clinical practice, manufacturing, other aspects, they're all sort of individual discrete domains. But the clinical trial itself starts with a plan, a design, coming up with endpoints, how are we going to measure it, recruitment, setting up the sites. That entire process is something that I think we can improve.

I will say onsite and centralized monitoring are complementary and not mutually exclusive approaches. We're looking at a blended approach for future. And we do need to develop and implement some better tools for what we're calling data anomaly detection.

Let me phrase it this way. Here are the four guidances I referenced. Basically, if you enter these titles into Google, they'll pop up.

Page 24

This is the National Academy of Sciences' report. And I'd like to thank my colleagues in the DB II, analgesics review team, Freda Cooner, Feng Lee, Kate Meaker, James Travis, Yan Zhou, and also to my colleague Scott Como for his input on PRO issues. And I finished on time.

(Applause.)

Q&A and Panel Discussion

DR. McDERMOTT: I'd like to invite the speakers and panelists to please come up. So we had three terrific presentations, and I think to start things, I'll just maybe summarize quickly some of the questions that either came up directly in the presentations or in my own mind.

In terms of central statistical monitoring, a number of issues related to -- I guess one could put it as cost effectiveness of traditional monitoring versus central statistical monitoring. I think that can use some investigation.

The issues of what does one check in an individual study, when, how often do we have to check things, who does the checking, what are the

Page 25

triggers. There are certain actions that might be taken, but what should precipitate these actions and how strange do these anomalies have to be before we take action? What actions should be taken in any particular case?

Paul raised some issues at the very end about potential standards for clinical trial practice, which I thought was sort of interesting. And Rick in his presentation raised a lot of questions actually about site monitoring, things about selecting -- dealing with other countries, for example, that one has to worry about the feasibility of recruitment versus quality issues, of course; having investigator meetings face to face versus having webinars, the sort of training issues that are associated with that; delegation of responsibilities from investigators to coordinators, who's overseeing, is there adequate supervision of the people to whom a lot of the trial tasks are being delegated; issues concerning informed consent training; and a bunch of other things that were raised.

Page 26

So I want to open it up to first the floor for any questions.

John, you're always first.

JOHN: I guess that's what I get for sitting up front. There have been some great talks, and I think the move towards trying to make things more efficient with central monitoring and not worrying. And being willing to say that site monitoring actually might not always serve the right purpose I think is a real step forward because, obviously, it's a lot of effort involved and so on.

The thing that I haven't heard as much is that in trying to implement all of these, there are a couple of considerations that I think we ought to take into account. And that is, is there a way to do it more efficiently and more effectively? And when we come to a fork in the road, could we perhaps, if they're equal in terms of the benefits to monitoring, could we choose the one that's more efficient or is likely to work more effectively?

I wondered -- to the panelists in general, but specifically with regards to Paul's

Page 27

talk -- whether there might be some guidances or evidence that could be put forward, or studies that could be conducted to say, all right, you need to comply with these, but the best way to do that is use a pocket data entry system and to make sure the data gets entered and checked, et cetera, et cetera.

I'm wondering whether there's a way to carry it that next step, something that Nat's been working on, which is to try and make all of this work in a better way, both from the perspective of keeping track of it obviously, but also from the perspective of actually getting it done.

DR. McDERMOTT: Paul, do you want to?

DR. SCHUETTE: I am not aware of anything that says where the dividing lines are for onsite versus these others. I think it's still fairly early days to actually determine which method is best. And I think over time, we'll see that methods evolve as to how we approach the best way of collecting the data and inputting it.

We saw, for example, that bring your own

Page 28

device type of things to clinical trials has upsides and downsides. Provision of various things, devices to the subjects has its own issues.

So as far as I know, we don't have any real guidance for that, and I think that's probably an area where the folks in the field can really help out. And if they can -- I'll lapse into FDA-speak. If they can work collaboratively together to develop best trial practices, that would certainly be something I think the agency would tend to support.

DR. WASAN: It's Ajay Wasan. So quick question. A lot of us who do investigator-initiated trials use REDcap, and REDcap on many levels kind of addressed a lot of the concerns that all of you have raised. And it's being used even more for bigger trials the NIH or PCORI are funding.

So I just want to get what's your sense to what extent, when REDcap is used well, that it actually is a pretty good data platform for capturing a lot of high quality type of data, as

Page 29

you-all outlined.
DR. SCHUETTE: We do not endorse any commercial product.
DR. WASAN: That's not a commercial product. This is funded by Vanderbilt. It's an NIH effort. I just want to get a sense of in general --
DR. SCHUETTE: We do not -- let me phrase it this way. We do not support any commercial or specific product by itself. If it's fit for use and for other things, we do not stand in the way, but we don't necessarily, for example, support SAS. We don't necessarily say you have to use R. So we try to stay away from particular platforms' endorsements.
DR. WASAN: And I'm sorry to be difficult. Let me just rephrase. I just want to get a general sense of -- the process that REDcap uses, that's all throughout NIH. In general, what's the sense of good and bad of that platform? That's all.
DR. SCHUETTE: That's again, one of those areas where, unfortunately -- and I'm not trying to be smart-alecky or anything else, but I can't

Page 30

comment.
DR. WASAN: Okay. Thanks.
DR. McDERMOTT: Have others here used REDcap on the panel?
So the one thing -- I haven't used it either, but some of my colleagues have. But one thing about REDcap or any other platform, I guess, is there are things you can build into the system to prevent certain kinds of errors, of course, like range checks and so forth.
But I think a lot of what -- and in particular Amy talked about yesterday was the very many logic checks, for example, or other kinds of checks that wouldn't be automatically produced by something like REDcap. And so a lot of people have moved toward electronic data capture with REDcap and other systems, but there's, I think, a fair amount of effort that needs to go into programming. And I don't know how specific that needs to be for each trial, how portable some of these methods are. I suspect that to a large extent, they are. But it's really a different level, I think, than what

Page 31

we're typically used to.
One thing that sort of strikes me as interesting is that in a lot of ways, these ideas aren't new. I mean, Nat talked about quality control way back when for industry. But a lot of the -- I think the first publication of this that was really noticed by people was in the late 1990s, and then it was about 10 years of not a whole lot. And all of a sudden, there were a lot of papers coming out now about this.
I suspect this has to do with cutting costs and so forth and trying to move away from traditional monitoring, but I'm sort of curious as to why the sudden interest and why there was this sort of long dead period. I don't know if anyone here has a comment about that.
If not, go ahead.
MALE SPEAKER: Yes, I have a separate question. It's kind of just a practical question. I'm a clinician and have been doing clinical trials forever. I don't know if the FDA -- one of the things we get all the time in the new electronic

Page 32

medical age is how to let monitors have access or not to EMRs.
So I have a clinical practice, and I have a research practice. We keep a firewall, obviously, between the two companies for HIPAA reasons. But the monitors increasingly beat us up when they come and they want to have access to our EMRs. So we take what we think are pertinent records from the practice and send those over to the research folks, print those out.
But obviously, it would be an enormous effort on our part if we let a monitor sit in front of the EMR, not to even get into whether I think it's ethical or not for them to have complete access to a patient's file, and then access to every other file in that Epic system and so forth.
So I don't know. Has the FDA dealt with that at all, or do you guys have any position on monitoring coming in? I mean, we want them to have access and quality data. That really has been a practical problem, has been an issue for us.
DR. SCHUETTE: I'm not directly involved

Page 33

with that level of approach since I'm in the Office of Biostatistics as opposed to Office of Compliance. But generally speaking, I would encourage you to reach out to your contacts in the Office of Compliance and actually say here are our concerns and to actually say, okay, how can we address these needs.

In some case, one can do extraction from a database and then make that available. That's just one possible approach, but the short answer is I don't know. But I would certainly encourage you to reach out to your appropriate contacts.

MALE SPEAKER: Great. Thanks. That is what we do. We do extraction out, and then let them have access to that, but that never seems to satisfy them. I don't know.

(Laughter.)

DR. SCHUETTE: Well, no, but that's good. I'm glad they're not satisfied because -- I mean, I was trying to raise the question as to whether we should try to collect more, not less, from the source documents to be distinguished from source

Page 34

data. But I agree, compliance, HIPAA, privacy, we get it.

DR. McDERMOTT: Bob?

DR. DWORKIN: Yes. Rick, I had a question about what you were describing as Teva's new policy. If I understood you correctly, you were saying that you would go to a site and you want to see source documents on the patients, and it sounded like you really mean medical records, for example, from their primary care clinician.

Now, I've been at meetings with other sponsors and CROs, where what we hear is that when patients are identified through, for example, advertising on TV, that it's kind of impossibly difficult to get the physicians or nurse practitioners, what have you, in the community to fax medical records, and that if that's required by the protocol, the CRO, whatever, aren't going to be able to get any patients through advertising.

So when I was listening to you talk, there was a real disconnect between what I've heard about the impossibility of getting medical records and

Page 35

what you said is the Teva policy, is you want to see the medical records.

DR. MALAMUT: No, I mean, I didn't say it was always successful because you're right. The study subjects who come through advertising, it is a challenge. But I think it's more of an increase in effort on the part of us to get those records, find out from the patient who their primary care physician is, get them to sign a release and get the records.

DR. DWORKIN: So to me, it seems almost essential if you're recruiting a patient for a low back pain study and you get the patient through advertising, and you then succeed in getting the patient's medical records, and it seems that they've never mentioned to their clinician in the past three years having back pain, that seems like a red flag.

DR. MALAMUT: Well, not only is a red flag, it's an exclusion. So again, I would argue that those patients are precisely the patients we need to get the records on. And I agree with you. We

Page 36

strive for 100 percent. Do we achieve 100 percent? Of course not. But I think if we don't try for 100 percent, we won't even get close. And for the reasons we've talked about, you won't have the right patient in your study.

DR. McDERMOTT: Okay. In the back?

MALE SPEAKER: I just want to follow up on this because it's really an interesting subject. If you're a clinician in private practice or academia and a patient is transferred to you, their medical records come with them, almost always, invariably. And if they don't come with them, there's a problem, and somebody's going to have to fix it.

I mean, you just don't take on a patient -- if you have a pain patient that I've inherited -- I don't do that anymore, but when I was at Emory, you always get the medical record. So I think this idea of not being able to get the medical record is a little bit suspect. It sounds to the non-clinician like, wow, that's a big deal, but it happened in the late '90s.

Page 37

But one point I do want to make is this, is that in terms of the amount of data we're collecting, it's huge. And when everything is important, to my mind, nothing is important. And so I do think you need to have focus in this world, and I think one of the things that remote monitoring and some of the things you're recommending to try to get actually kind of addresses that issue.

But one of the things that comes up a lot is 100 percent SDV. I might have missed that talk. But that's part of what I'm very curious about what everybody's opinion is on that, where you go to the site. The primary endpoint is what it is. Obviously, you need to -- for the data that you're really focusing on in terms of your submission, obviously, that's important. There are a lot of ancillary endpoints people collect, and measures they collect, how well that needs to be done. Obviously, safety needs to be done well.

But I wonder if people can talk about the SDV issue because I find like a lot of people are

Page 38

really focused on getting everything.

MS. KIRKWOOD: I agree. I think for our center, we don't do 100 percent SDV. We're an academic center. We can't afford to do that.

I agree that not all data is as important as some of the data and that -- I mean, we are sort of looking at lots of data in the kind of methods that I described, but it was more to look at patterns rather than actually to try and correct everything, and that it's much more important to make sure that your primary outcome measures are all there and all correct than it is all of these blood measurements that you take at every cycle of chemotherapy that no one's ever going to use, and especially if we, on top of that, also collected safety data about those sorts of things.

I think some studies that have been done, there doesn't seem to be any proof that 100 percent SDV is necessary and adds anything.

DR. SCHUETTE: And I would actually go with the guidance on the topic for risk-based. So that's basically -- it sounds like it's pretty much

Page 39

an outline for the type of difference that you might be looking for.

DR. SINGLA: So just to get back on this source -- I mean medical records issue, as an investigator who has done this for a long time and recruits a lot of patients through advertising and different types of patients, I just wanted to provide some insight, which is that first of all, it is very difficult to get medical records from recruited patients. It's not like clinic patients because clinic patients are inside the healthcare system, and as such, there's an expectation that they will come with records.

So it is hard. It's also hard because the screening period is typically like 28 days. You got to get the patient randomized, and by the time you get their medical records, you're right at the end of that, and then you can't rescreen them. So those are the difficulties.

It is possible, however, like you said, Rick. We've had studies where it was mandated, and then we found a way to do it sometimes. But then

Page 40

we lost patients as well. It was a lot of effort, and then it's a question of is it worth it?

So I think it requires a decision on the sponsor's part to decide a priori, what is the disease being studied and do we need medical records.

For example, low back pain, yes, you probably do, and patients who are on opioids, it's a serious issue. Surgical third molar extraction, 19-year-old patient probably doesn't have any medical records. You're going to bang your head against the wall trying to find the medical records from the time they were like 11 and went to see their pediatrician. And you know that they have third molars. You can see them on X-ray, and you basically know they're healthy. Do you really need them in that situation? Probably not. It's a six-hour study. They're going to have the indicated surgery.

So I think that when the sponsor says -- like, puts a half approach towards the situation and says, well, we'd like them, then it

Page 41

just causes a lot of confusion. Nobody does it. And by the time the trial is 50 percent over, you just stop trying because the sponsor has given up as well.

So that's just a practicality of what I've seen happen over the years. So I think that you have to consider the disease and then make a choice as a sponsor. Yes, you must have them, and then increase the screening period if you're going to force that.

DR. MALAMUT: So it's a pragmatic approach is what you're suggesting --

DR. SINGLA: Yeah.

DR. MALAMUT: -- not a mandated thou shalt provide source records, but a little more pragmatic. You're right. Bunionectomy study with younger people or maybe not as important. But I'm really going to argue strongly about certain -- I'm still leaning towards the idea of this 100 percent idea even though I know we can be pragmatic. But we have to study the right patients.

DR. SINGLA: And I'm just saying, for low

Page 42

back pain, maybe you do want 100 percent.

DR. MALAMUT: Yes, I think so.

DR. SINGLA: They're on opioids. It's a serious condition, multiple medications. You can't verify they have the condition whereas in a bunionectomy, you can verify it, those kind of things.

DR. DWORKIN: So I would push and say that clearly, there are exceptions, and I think, Neil, you did a great job of giving us an example of an exception.

But if we're talking about chronic pain in adults, whether it's low back pain or osteoarthritis or diabetic peripheral neuropathy, and assuming it is correct that there are study kingpins not only in Boston but elsewhere who are going to clinicaltrials.gov and finding out the inclusion/exclusion criteria of trials, and feeding patients into those trials, then don't we really want to see the clinician's record on 100 percent of these adults with symptomatic chronic pain? Otherwise, I mean, we're just making it real easy

Page 43

for the study kingpins.

DR. MALAMUT: You could make the argument that the patients you can't get records on and disappear, maybe they weren't the ones you wanted. I mean, maybe you're losing out on some very good study patients, but the tradeoff.

DR. McDERMOTT: Michael, did you have a --

DR. ROWBOTHAM: I just wanted to go a little further with some of the discussion this morning and then yesterday about data that's too good to be true. So one of the possibilities also is that you could have a positive trial and it comes out negative because of fraud or fabrication on study sites just trying to increase their numbers.

So could I hear a little bit more from both the FDA and the industry perspective as to what kind of data checking is likely to be done routinely to make sure that the data that was sent in a file to the FDA actually is legitimate data?

DR. SCHUETTE: Right now, it's more on a trial-by-trial basis. There really isn't an entire across the submission look at data quality on each

Page 44

and every aspect. So what we have right now is fairly rudimentary checks, sort of things like calendar dates and some of the other types of items.

The example that Sharon gave was noteworthy in the sense that there was actually a comparison between one site -- one trial and another trial, trials that were conducted in the U.S. and outside. And they said an unbelievable response rate over here and a middling response rate over here. So that was a case where we could just say, just by looking at it, this is way too good to be true.

That's actually where we are right now. What we'd like to do with some of these other aspects is to look at it in a more coherent fashion, and that's still a matter of development from our perspective.

DR. MALAMUT: I think, as I said, we do our best -- is that the right word? -- to verify everything that the study sites tell us, what they write down, what's documented. A lot of the efforts go towards making sure data is entered.

Page 45

We're a little limited in being able at the time of monitoring, at the time after last patient out, to be able to verify that a pain score is actually an accurate pain score. So a lot of the too good to be true may have to come later during the all too short time to look at the data and then in analyses later. And of course, we'd like it to be good, so there's a bias.

When we see something that's successful, we say aha, we were right, but then we do try to be critical and look for those patterns, and do take it to individual sites, individual regions, and try to see not only why a study may not have succeeded but why did a study succeed, where did the positive data come from.

DR. McDERMOTT: I think --

DR. DEVINE: Eric.

DR. McDERMOTT: Eric, sorry.

DR. DEVINE: I've come to this meeting to beat a single drum. Everyone knows what I'm going to say. But we want people to exclude professional subjects from their study before they get in.

Page 46

But assuming the potential that not all sites will make the efforts to get the medical records or design the study so it isn't really vulnerable to being gamed by the professional subject, and maybe there's collusion on the part of the investigator who really wants to enroll quickly, do you think that it's possible, from a monitoring perspective, to pick up on patterns of data that are indicative not of a fraudulent site but of a fraudulent subject?

Over a lot of observations of what a particular disease looks like and the way the subject answers questions in the trial, would you see a pattern that you could pick up on that professional subject because their answers are similar to someone that's malingering with a learning disorder because they need services, or someone who is malingering with pain because they want the drug?

MS. KIRKWOOD: To be honest, it's not something I ever considered until this meeting because it's not something we have in the disease

Page 47

area I work in. You don't get professional subjects in cancer trials. But it's definitely something that I think would be interesting to look at.

I think it might be hard to do without examples of it. So what you would really need is a trial where you've got a data set with patients that you have identified as professional subjects, and if you could get that, then it would definitely be something that would be very interesting to look at, whether they could be picked out. But it's not something I've looked at because it's not something I'd imagined before this meeting really.

DR. SCHUETTE: And I will say that we don't have -- we'd be interested in the exact same thing with actual examples, and we're working with commercial developers in some ways to try to do research and development.

I've already mentioned after reading your paper, Eric, this idea to one person -- one developer suggested that this was a product niche that if they were willing to, they could actually

Page 48

pursue.

DR. McDERMOTT: Way in the back.

DR. SIMON: So slight note of caution for those of us who actually serve on executive committees or DSMBs or DMCs or any of these things. In the last 10 years, it's actually been more common than not that we've seen inappropriate or inadequate or mismanaged oversight by hired CROs that require buy-in from the physician or caregiver supervisory group.

People 20 years ago, when I was a trialist, 25 years ago, were hesitant to become proactive with the CRO or the company doing the study because they felt that that was their purview. But the events that I'm referring to are so egregious and so extraordinary that, in fact, if we do not exert some responsibility, we cannot allow or expect only the regulatory group to pick this up because some of this stuff is extraordinarily buried in the database and it may not become evident. And yet theoretically, a DMC chair and his team should be cognizant of what's going on in the trial.

Page 49

So the caution that I'm trying to suggest is that we keep asking questions that are asking the FDA to have an answer for how these things get controlled, but it's going to have to require cooperation between the external groups that are supposedly overseeing the studies and the internal groups.

Just for example, most recently I am serving as a kind of monitor of a trial, and I was asked to review the SOPs of the CRO regarding the history of the patient. And the concerns were for a chronic pain trial that the history was only going to be taken for the previous year.

Well, the problem was, is that a lot of the potential history for allergies and other issues would either have to be extracted from the chart or, God forbid, the PI on the trial at the site would actually have to take a history from the patient that is actually being recruited into his site, and he's getting paid for this.

So the reality is we have to take some responsibility for this, too, and be mature enough

Page 50

to actually do the work that we're being paid to do. And I'm not actually hearing anybody acknowledging that particular aspect of this.

Of course, nobody in this room has any belief to be malfeasant or not to do what he's supposed to do, but in fact, it's become de rigueur that the superficial nature of the supervision of the trials on our side is actually quite bad, so note of caution.

DR. McDERMOTT: Yes, I think, Rick, you said something about vendor oversight yesterday.

DR. MALAMUT: Yes. I mean, again, I was trying to be veiled yesterday because I didn't want to actually give strong opinions, although somebody last night told me they could tell everything I thought. So maybe that's good. Thanks, John.

But I really do want to hear from others about this. I think I may have made the case yesterday that exactly what Lee is saying, that we do need to take responsibility as sponsors, as CROs, as investigators to take a proper history.

If you're not at that 100 percent level on

Page 51

the patient record, we just assume that you're taking the history and finding out the allergies and the con meds, and of course, the presumed study patient may have incentives to not tell you the whole truth, but we have to at least make sure we're doing that.

The same on the vendors. I tried to make the point yesterday that we all have to monitor each other. Otherwise, the data we get will maybe not reflect the true nature of the compound we're testing or be misleading. And I fully agree, Lee.

DR. SCHUETTE: If I can jump in, I'll echo what Richard has indicated and say we completely agree with Lee, particularly when we're talking about multi-regional trials and things that are done outside of this country where, as was pointed out by Sharon yesterday, we can't always, as regulatory agencies from the U.S., get access to the types of data in other countries. And it becomes particularly important that the CRO or the sponsor ensure that the data is good, that there is not misconduct at the site because some cases,

Page 52

we're blocked from going very much further than just an overall look and inspection.

DR. McDERMOTT: Okay.

MALE SPEAKER: Just an obvious follow-up observation, again, the way we did things 15, 20 years ago, just following up on some of Lee's comments, was much more personal. So there are so many more layers now in study conduct.

To your point, Rick, getting to the site is incredibly important. So all these centralized and statistical monitoring approaches are very important, but you really do need to get to the sites.

As we go up the food chain in industry, we get further and further away from the sites. In this meeting here today, I've heard us talk about some of the most important people are the CRAs. They're the most inexperienced people in the system, and we overtax them. We expect them to do monitoring. If we really want to have onsite eyes, we either have to do a better job with the CRAs or get out there ourselves.

Page 53

So in small companies -- and I don't know if others have even been to sites. I don't know, Rick, if you've been out to visit a site recently --

DR. MALAMUT: Not in a few years.

MALE SPEAKER: -- but, yeah, it's really eye-opening. And there are a lot of intangibles that you pick up. And obviously, you can put in statistical and other central monitoring schemes, but until you get to meet the investigator, the coordinator -- I mean, there isn't probably even a coordinator here at the meeting. I'm not sure if that's true or not.

But they're the quarterbacks for these studies, and maybe just as we think about this, Bob and Dennis, going forward, probably we need to get some feedback from coordinators and monitors about this guidance if we're going to be creating this paper that people will be reading.

But again, very impersonal, lots of layers. I think CROs are probably -- I mean, I've worked for CROs. I've worked with CROs. It's a problem.

Page 54

There's just too many layers from an issue or problem getting to the source. It has to go from the coordinator to the CRA, up the chain through the CRO, through the project manager, back to the sponsor. And as a CMO in a small company, I'm eight layers removed from an issue, and it gets filtered. Basically, we have to do a much better jobs.

Your thoughts, Rick?

DR. MALAMUT: Again, obviously, I agree because you made the points I made yesterday. But no, I'm envisioning a diagram with circles of all the people involved in the study with everyone connecting to everyone else because we're all monitoring each other, and we all have responsibilities, so yes.

DR. McDERMOTT: I guess way in the back and then --

MALE SPEAKER: All these comments are really good, and as somebody who as even like a year ago when I was at Merck was doing site visits, I totally endorse that. But one of the greatest

Page 55

technical devices that's been invented recently is the phone. And picking it up and just calling people, you'd be surprised how much valuable information you get just by talking to an investigator because they have phones, too. It's pretty cool.

(Laughter.)

MALE SPEAKER: Now, but the point I do want to make is a serious one, picking up from Lee's comment, is I was amazed. The way we do things in the United States, we think the whole world does it that way. And I've been used to a system where we have highly trained study coordinators who really know what they're doing better than, at one point, I did when I first got into it. These study coordinators are sometimes amazing.

Then we go to these sites, and we open them up for the first time. And in Europe what really surprised me is that study coordinators are not these really anal retentive nurses who've been, we're on the floors for years and now we're working clinical trials. They're young physicians, who are

Page 56

really both sub-PI and study coordinator.

I've expressed concerns about that on multiple occasions during multiple site visits in Europe. People have allayed my fears somewhat, but it is something that everybody should be aware of. I think doctors are pretty good, but not all of us are trained to be that focused on the minutiae, which I think a really good study coordinator needs to be.

DR. McDERMOTT: David?

DAVID: Yes, sort of in all this discussion, I think we're looking a lot at downstream events. And when you talk about drug or the kingpin, the study kingpin, that also could be a study coordinator. And I'm wondering, has anybody ever looked at sort of the study equivalent of a secret shopper where you put a sham patient in and sort of see what happens at that interface? Because I think that could be very interesting. I've never done it myself, but I'm very curious if anybody has.

DR. McDERMOTT: No takers.

Page 57

(Laughter.)
DR. MALAMUT: I don't think I want to reply to that. Secret shopper. Wow.
(Laughter.)
MALE SPEAKER: No, there have been. There have been secret research subjects, and they'll publish it. And it's a big splash article in a local paper about how they pretended to be a patient and how easy it was to get into a study and get drugs. It happens every once in a while.
DR. MALAMUT: So what was the outcome?
MALE SPEAKER: Oh, no, it always makes some great expose article. I don't know when the last one was published, but I don't think it's been done on a systemic basis.
MALE SPEAKER: Yes. I was thinking more of the quality aspect, not the journalistic aspect.
MALE SPEAKER: I was thinking about actually -- Rob and I were talking last night about the undercover diner. It goes into to see who's stealing from the cash register.
(Crosstalk.)

Page 58

MALE SPEAKER: Just to chime in real quick -- sorry. Actually, the secret shopper, people may know, in most healthcare systems is actually used as a validated way of quality checks and auditing. So there actually is a nice track record in clinical care for using it.
So it's certainly a reasonable thing to maybe consider as a bullet point in the paper. Do you consider such a -- because it has been used and validated throughout healthcare, and we use it. And I'm always interested in getting the secret shopper reports on my service. It's fascinating.
DR. McDERMOTT: Over here, and then Lee afterwards.
MALE SPEAKER: So I apologize again if this has been said. I think we talk out of both sides of our mouths. Maybe in the clinical trials area, this is a major concern, and I understand that and I appreciate that we're spending the time talking about that.
In other research enterprises in the same wing of our hospital, they keep lists of

Page 59

professional patients in laboratory analog trials that are clean of any drugs, they're free of health problems and so forth, to keep them coming back for more and more trials.
I think it's important that we, I guess as a group, appreciate that maybe in our same community, this is being thought about in quite different ways That's one point.
I guess I heard really yesterday -- it's maybe shifting the gears a little bit here. But at the point of making decisions about a person's meeting inclusion and exclusion criteria for participation in a study, which I think is really central to the integrity of the study, I heard both in presentations and, then importantly for me, in sidebar conversations with a number of people over the last day and a half a sense of, yeah, there are those. But if I really have a bad feeling about a person, I don't include them in the study.
I have a very serious concern about whether we really have additional inclusion and exclusion criteria that we're talking about here, which is we

Page 60

include people if they maybe are close to meeting the criteria in terms of, for example, age or some other criteria. Then we should be clear that there's some wiggle room there. I don't agree with that, but I imagine that that's the case.
Exclusion criteria, which is we have all these things, if the person meets those criteria, but we still don't feel -- we think maybe they're a fraudulent or professional patient, we have a hunch or gut, we exclude them. I'm very concerned about that slippery slope.
DR. MALAMUT: I won't speak for all my fellow sponsors, but in most studies, there is somewhere down at the bottom of the exclusion list a exclusion for any other reason the investigator feels the patient is not appropriate. And there's no rule that says just because a patient meets all the criteria, they have to be enrolled. I mean, it really is discretionary.
But I think there's a risk, and I think everyone in this room either is, or has been, or will be an investigator. We make that check. I

Page 61

think we're doing -- our internal system is checking to see is this really what they say they are.

My concern is more with the less expert enroller, some of the research sites who don't have that inner -- and they're checking the box. They're saying, well, they met all these criteria, therefore, they must qualify. They're not really thinking beyond that. And I don't mean to generalize.

MALE SPEAKER: I've encouraged journal editors to ask that that disclaimer, if you will, be added to every --

DR. McDERMOTT: Well, sometimes other pressures enter into this, too. I mean, pressure to recruit or to be kicked out of the trial as a site. All sorts of things come into play.

Lee, you've been waiting.

DR. SIMON: It's really interesting. This raises the problem that we've assumed that the people that are out there serving as principal investigators at the individual sites are actually

Page 62

capable of doing that because they've actually been able to recruit before or for any other criteria that are there.

We are all experienced people, and we all think we know what we're doing. And we've taught ourselves since there is no academic process of learning how to become a clinical trialist, it is catch as catch can. And basically, in the end, we've abrogated our responsibilities for these clinical trials unless we create a methodology to allow people to become clinical trialists and make it be a actual learned endeavor.

To complain that people use their own intuitive nature of deciding if somebody will get into a trial or not, which is absolutely what happens all the time, and why that last comment exists in the exclusion criteria, we have an opportunity here to identify within this manuscript what we believe should be the right way to do things.

It even goes down to how we recompense people for work. We typically create contracts

Page 63

where they get paid for recruitment, not paid for completion. And therefore, that leads to missingness, which is like a nightmare. And furthermore, it also leads to inadequate patients being recruited because the pressure, as you just mentioned, is to recruit so that you can actually remain within the trial, because otherwise, you'll be dropped out if you only have two or three people compared to somebody else that already has 45.

So I think that we in this community who believe in this process should, in fact, create an infrastructure to allow people to learn how to do these things and become certified or at least knowledgeable.

To think that this all happens at an investigator meeting that may take eight hours, and everybody is asleep and on their computer anyway during the time, is ridiculous. Let's be honest. This is a complicated process that requires real knowledge, and we should be teaching it.

DR. McDERMOTT: Laurie?

MS. BURKE: I completely agree from my

Page 64

limited knowledge post ex-FDA, from what I've seen. And I think that this isn't just one curriculum to qualify people to be clinical trialists, but it's multiple. It starts with -- there's the multiple disciplines' worth of qualification of degree programs, or whatever you want to call them, that need to be thought about. And of course, my thing is the measurement area. There's really no place that people can go to get a degree in clinical trial measurement, and that, I think, is one in and of itself.

MALE SPEAKER: Just one comment on the compensation, which I think to piggyback on Lee's comment, which is really important. Clinical trial sites don't get paid to screen in general. There are screening fees. Yes, of course, there is, but you get paid to randomize subjects. And it's completely skewed, the amount of money you get when a subject randomizes.

I think it's because sponsors want randomized subjects, which makes sense. They also find it difficult to pay for subjects who are

Page 65

screened because it's easy to inflate the number of patients you screen, so it's kind of just a necessary evil.

But I think this economic lopsidedness about randomization leads to a pressure to randomize and to not get -- because you're not getting compensated essentially for screening. It's almost like you're paying for screening, and then you get paid when you randomize subjects. That's how it is as an investigative site, which is the genesis of a lot of these problems.

To talk just one more point about what Bob Dworkin said yesterday regarding blinding clinical trial sites to when the subjects can be screened and randomized, I think that's a good idea.

When that does not occur, in other words, the decision has not been made to blind. When there's subjective criteria or like a baseline entry criteria, let's say, for OIC, patients have to have less than this many bowel movements, or for a OA study, they have to have a flare of X, Y or Z, you can see in the data -- someone's talking

Page 66

about -- we're all talking about central monitoring -- that different sites have differential rates of patients that will make it through that baseline period and sometimes widely differential. And when it is widely differential, that's a problem because it's all based on competition.

So I think when you look at central statistical monitoring, this is the key aspect because it's financially driven and it affects the randomization.

DR. McDERMOTT: Matt?

MALE SPEAKER: I think it's -- I want to pick up on comments made earlier about contracting. Now that we're contracting with sites to do clinical trials, it's kind of shocking how much contracting influences quality in the sense that investigators are never, ever, ever contracted for quality. They're only contracted for procedures, whether those are visits or EKGs or visits or histories and physicals or what have you.

Then when you try to go back to the site

Page 67

later and say, hey, we want you to really think harder about patient recruitment or think harder about how you're coaching your patients -- training your patients -- to measure pain, or get your queries resolved more quickly or whatever the quality metric is, you can't get anywhere because you don't have any financial leverage over those sites unless you are in a network where you own the sites.

I wonder whether it's almost worth a paragraph in this paper or at least some discussion of how we fall short of trying to influence quality because we fail to account for it in our contracting processes.

DR. MALAMUT: It's almost for the next study. You're right, in the middle of a study, it's very difficult, without seeing the data, to know what the quality of the data actually is. But I think the act should be on the next study, so that if I look at site X and they've recruited 30 patients, and did a great job recruiting but, in fact, most of the data had to be thrown out and

Page 68

other patients -- then that site is not going to be selected for the next study. Now, maybe that site doesn't care. But we would hope they do.

MALE SPEAKER: The facts, I think, show that if that's the current system, it's not working, because otherwise, we wouldn't be having this meetings.

I also think that as much -- ask any CRO or any sponsor, do they monitor -- do they measure the quality of the site at the end of the study, they'll all say yes. And then when you ask them to show you exactly how they do that, no one can ever come up with anything. They can't find it. That was the other CRO. Well, what we really know is whether there were any major audit findings or how many patients they recruited.

So there's a huge disconnect between what people claim they evaluate in terms of study quality and how they utilize that information for the next study and what's actually being done. So I think there are opportunities within the contracting process to pay for quality and create

Page 69

that financial leverage.
DR. McDERMOTT: Way in the back and then --
DR. HEWITT: So what I'd say is I do believe in relationships with sites, and there are sites that I've used, we've all used probably, in clinical trials that go back to the late '90s. So we all know a lot of people who do a lot of clinical trials and are good sites.
And I think when you -- just to pick on Nat's point, a high-quality site, I think for those of us who have really been in the trenches, has to do with their queries. If the data is really dirty, no study coordinator -- I mean, no CRA is going to want to pick a site again that's just given them hell for three months as they're trying to clean up the queries.
So although you could use a lot of metrics, I'll tell you, the number of queries and getting them rectified in a timely fashion is what will either get you on to the next study at work at inVentive or not because for those of us who follow these things, those are important metrics.

Page 70

DR. McDERMOTT: Roy?
DR. FREEMAN: So I want to pick up on a point made by I think it was Lee behind me and the invention of the cell phone guy, which I think was David Hewitt, but I'm not sure.
Bob Dworkin and I found ourselves at an investigator meeting a week or so ago, and we had a chat -- they called it a fireside chat -- in front of a group of study coordinators and investigators, and there were no academic investigators. These were all pay-to-play type sites.
The aim of the chat was to discuss concepts related to the placebo response, and topics included things like how to balance your desire to recruit more and more subjects versus selling the study drug as the new wonder drug and raising expectations, how to balance study retention versus being warm and nurturing and fuzzy, and again, enhancing placebo response. And I could go on about the nature of the discussion, which was kind of entertaining.
But what was eye-opening was at the end of

Page 71

it, the number of study coordinators that came up to me -- I don't know about Bob -- and said, "I've never really thought of that before." And it was so interesting to hear that.
So picking up on Lee Simon university for clinical trialists, university for study coordinators kind of concept, obviously, really difficult to operationalize. But what I do think is a really good idea, and perhaps this could go into the document as well, is at each investigator meeting, there should be a clinical trial 101 type meeting in which the protocol is not just discussed and how to do an EKG, but actually the nature of a clinical trial and what we are doing in it. This is an experiment in clinical equipoise, and all of the things that we kind of take for granted, but they actually don't understand, and they really don't.
DR. McDERMOTT: Scott?
DR. EVANS: So in statistics, we have a saying that there are lies, damn lies, and neurologists.

Page 72

(Laughter.)
DR. EVANS: You guys may have heard a different version.
I'd like to pick up on that point because much of the discussion is sort of focused on detection of fraud and malfeasance and outliers in a sense. But when I think about data quality from a broad perspective and where we can make the most impact -- and I'm someone who's been teaching clinical trials for 10 years, so hopefully, there is some academic process to this.
The first thing that comes to mind from a statistical perspective, where I do think we could make enormous impact, and it's probably old news, but it's a point that Paul made on missing data. And the National Academy of Sciences put out a report a couple of years ago, and there was a New England Journal of Medicine summary of that report a couple of years ago as well.
Basically, the message in that report is that missing data is not a data analysis problem. It's a design and conduct problem, with the message

Page 73

at prevention and dealing with this upfront.
Now, I'm a part of a new clinical trials network, and I've essentially said that efforts to minimize missing data is a standard section in the protocol, and we have to figure out ways to prevent it and deal with it because if you get it at the end, as you know, prevention is the best medicine. But fancy statistical methods are not going to rescue design and conduct flaws. So I think this whole process is really sort of a prevention issue.
Picking up on the education piece, I think that that's really an educational message, that if you can train people to understand fundamentals about clinical trials, your quality is going to go up.
Just getting people to realize the important distinction between needing to go off study because of toxicity -- or needing to go off treatment because of toxicity doesn't mean you have to go off study and that I'm going to lose your data.
So there are a number of things, and there's a checklist in the New England Journal of Medicine

Page 74

article or the Academy's report, data management practices about clear CRFs and not overburdening patients and doing things that enables them to be able to stay on study; the intention to treat principle, getting people to understand the intent to treat principle, that follow patients regardless of adherence; and their example language about if a patient wants to withdraw, whether we could actually withdraw you from treatment but still collect your data and follow you, and that has important implications.
Things in design, clearly thinking carefully about the population, whether you want to do run-in periods, which may reduce missing data later on; flexible treatment regimens, how would you handle the rescue medication issue? And this actually gets at the effectiveness versus efficacy piece. And in academic medicine, I try to get people to think more about effectiveness.
If a patient goes off treatment because of an adverse event, a toxicity associated with the medication, and therefore, maybe they go off study

Page 75

even, and you don't get their measurements that you expected to get at the end of the day, and we go to analyze pain, oh, all of a sudden, we've got a missing data problem.
Well, it's a missing data problem when you're trying to evaluate causal pathways and mechanisms of action and understand biology. It's not a missing data problem in clinical medicine. It failed the patient.
So either having to go off therapy or having to rescue them is actually part of the outcome. It's not missing in a sense and getting to think about whether you need to characterize outcomes that bring in this information.
So I've been pushing in other areas that in clinical trials these days, our tradition is we collect data on patients and then analyze the endpoints. Well, I want to reverse the order. Collect data on the endpoints and analyze the patients. That's who we're treating. That's what's going to apply in practice. And that will help eliminate some of the missing data issues.

Page 76

So there are certainly things we can do -- I did have a couple of comments about the central monitoring issue, statistical monitoring. We already do in clinical trials, as Mike mentioned, range checks and logical checks, and we can certainly turn up the temperature in that. And it will take effort and thought.
You need to decide what you're looking for, and this is not entirely clear what we're looking for. And that's going to require more thought than perhaps Amy's had to get in the tall weeds about it. But I think when you apply it, it's going to take more thought than sort of this 40,000-foot approach.
Then when you say, well, what is an outlier? And how do we decide what's true and false? This is really a diagnostic problem. You're trying to figure out whether I can diagnose fraud or outlier or something that's a problem. And so this is a balance of sensitivity and specificity in a sense.
We can detect more outliers, but you're also going to detect more false positives, and there's

Page 77

going to be a consequence to that. And frankly, I don't even know what the gold standard is. We don't even have a gold standard. It's an imperfect gold standard in many cases. So there are some real challenges here in thinking that through.

Then I think deciding how you handle a particular issue, if you've identified it -- there's been talked about intent to treat and whether you exclude them or whether you don't -- we're really going to need to think -- we're going to need a more detailed evaluation on the nature of the issue and the consequences of different actions.

If you're on a case where you're running a blinded trial and the blinding actually works, then if there's malfeasance going on or people are just enrolling patients that are just nonsense data, well, that's going to hurt assay sensitivity. But if the blinding is really -- and so it's going to hurt the ability to detect differences and so forth. But it's not necessarily differential between arms. Although if you're doing a

Page 78

noninferiority trial, it actually would bias toward non-inferiority, and you get a different problem.

So I think the blinding issue is a real important one, and I often encourage people to evaluate the success of the blind. We often say we're running blinded trials, but whether the blinding worked is a whole different issue. And people often just refrain or refuse to evaluate if it worked or not through questionnaires, and I think that may help us understand what the potential consequences of this are.

So I'll end there. Thanks. Very quiet after that.

DR. McDERMOTT: You just quieted the room.

DR. EVANS: So you guys have been thinking your lives about how to reduce pain, and statisticians think about how to inflict it.

(Laughter.)

DR. DWORKIN: This is only partly related to your comments, Scott, and follows up on what Lee was saying. I don't know whether this exists in other therapeutic areas because it's certainly

Page 79

clear it doesn't exist for pain.

In other therapeutic areas, are there kind of training certificate programs for study coordinators and principal investigators? And it goes back to what Roy said, that maybe the beginning of an investigators meeting should be some kind of general introduction to analgesic clinical trials 101.

But doing it in that kind of decentralized, leading up to the sponsor ad hoc way is clearly not as good as if some organization, for example, ACTTION, put together a two- or three-day boot camp for junior investigators, senior investigators, study coordinators, and it was kind of introduction to analgesic clinical trials 101 with people like Mike and Scott and everybody on the panel, and many of us in the room, instructing the people who come to the meeting. And they all walk home with a little certificate. It could be set up as CME, that they spent three days learning all these challenging issues about clinical trials.

So this just occurred actually completely

Page 80

independently to both Dennis and me while we were sitting here. Does that exist? Does anyone know whether anything like that exists anywhere else?

DR. McDERMOTT: You just caused eight hands to go up.

DAVID: There is an accredited organization called the American Association for Pharmaceutical Scientists, I believe. They've been around for a number of years, and they do grant some sort of a certification process. However, it's costly, it's time consuming, and a bit onerous.

So I think the solution that you propose, to have some sort of a training during an investigator meeting that is iterative, that can be accessed on a corporate or a CRO website, I think is part of the solution.

DR. DWORKIN: David, I know there are existing programs, but the ones I'm familiar with are all generic. I'm talking about something that's pain specific. I don't know whether it's just chronic pain or chronic and acute pain.

So is there something like for

Page 81

cardiovascular clinical trialists or people who do type 2 diabetes trials? Does anyone know of anything that's -- any kind of training program for clinical trialists, both investigators and coordinators, that's therapeutic area specific?

DR. EVANS: So recently, because of problems associated with performing rheumatoid arthritis clinical trials, which require you to actually do hands-on outcomes -- and it turns out that the most experienced rheumatologists can't do a physical exam appropriately, therefore, this has been studied -- people like at Keystone and others in Canada actually put together training programs for investigators at investigators meetings, where they go in and get tested whether they can actually feel tender and swollen joints. I mean, it's like wait a minute, this is what I do for a living, and yet, in fact, passing such a test is ridiculous.

Furthermore, there's another training system for injectable drugs, intra-articular drugs. As it turns out, even the most experienced orthopedist and rheumatologist miss 33 to 40 percent of the

Page 82

time getting the needle in the joint. My God, even the knee, which is like the size of Manhattan.

(Laughter.)

DR. EVANS: So they actually have created very specified training programs for this, and then they also have reminder programs during the time of the trial to bring people back up to date. And you cannot become a PI at a site for some of these trials unless, in fact, you pass this.

DR. FARRAR: Just a quick comment on that. Misha Backonja about seven years created a tape for study on exam to identify neuropathic pain, for instance. So there are very specific instances like Lee is defining. And you're right. There are other training programs for coordinators.

At Penn, we had a four-week training program, but it's not specific on a particular area. And I think that there could be real benefit to doing that.

DR. McDERMOTT: Mike?

DR. ROWBOTHAM: To Lee's comment, maybe an aside, that's why they went into research.

Page 83

(Laughter.)

(Crosstalk.)

DR. ROWBOTHAM: Their offices are so full of patients, they don't have time to deal with doing research.

But I just wanted to make one point. Sometimes the investigators meetings will have a lot of materials. I've done this like on how to examine postherpetic neuralgia patients and do sensory mapping and injections and stuff.

We create those, but one issue that comes up, especially at the organization that I'm in where we do a lot of cancer trials, is that you can have sub-investigators enroll patients as long as they've been trained by the PI on how to do everything.

So then you've moved one step away from what actually was covered at the investigators meeting, and we have to spend a lot of time making sure that when a PI trains a sub-I that we're confident as a research organization that the sub-I really does know what they're doing on the protocol. And that

Page 84

may not actually happen at all research organizations.

DR. MALAMUT: I think I had briefly hinted at that yesterday, that whoever shows up at the investigator meeting, often at the site, they may not have gotten the benefit of the training we're proposing. So if we're going to put some kind of training at the IM, we're going to then have to insist that everyone who's involved in the study will show up and get that training, which isn't always possible.

In regard to training, we've done it a few times in studies where we've wanted to stratify or assess the presence of allodynia or mechanical hypersensitivity. So we put together training tapes with -- I guess Brett Stacey did our most recent one.

So I think amongst everyone, there might be individual training videos and pieces specific to different studies, and maybe they just need all that to come together.

DR. McDERMOTT: Our coordinating center will

Page 85

do webinar training or they will have -- it's been like that when you have new personnel coming onboard. I don't know if that's standard across everyone, but that's what we typically do.

In the back?

MALE SPEAKER: There are some aspects of this that kind of exist within -- as people have said, that exist now like GCP training, everybody has to. All investigators hate the fact that they have to do their GCP training every time they do a clinical trial with a sponsor. That's a big issue.

But certainly with site initiation visit, there's a lot of training that should be going on as well as stuff that goes on at the investigator meeting. And certainly for site initiation visits, they should be able to get pretty good training, and there's training online.

But with all of that said, I think the point Bob is making is a good one, is that we're really going beyond something. And I think if you can get like a certification so that people really get it and understand it, I think that's a huge, huge

Page 86

thing because medical schools aren't going to do it. Medical schools have been -- we've been trying to teach medical schools to teach about chronic pain for 20 years. They're still not going to do it. So I endorse the idea.

MALE SPEAKER: I guess I have a quick question. How is this being handled for international folks where we're dealing with multi-regional trials and folks who are not always proficient or even fluent in the language in which the primary training materials have been created?

MALE SPEAKER: Well, that actually is an issue that people have to address. Frequently, these materials are translated. Certainly, all the patient materials are translated into other languages. But it is a good point that you need to -- and then is the translation right is a big thing as well. There are all these dedicated translation services that go back, translation forward to make sure that they've got it right. And then your CRAs speak the language, too. But it is an important issue.

Page 87

DR. McDERMOTT: Lee, did you have a comment?

DR. SIMON: I just wanted to comment that one has to wonder whether national professional societies, be it in the U.S. or in Europe, have an obligation to do that. If they're going to help their membership professionalized themselves, one of the ways to professionalize is to become a professional trialist. And what the problem has been is it's not considered an academically scholarly activity to do clinical trial work here in the United States.

So there is no real pressure on somebody who does it in an academic site to publish. I mean, if you're one of 400 investigators, you're not going to be one of the people who are going to be the author on the paper. You might get acknowledged, but then that's not recognized. And in certain institutions, even if you're a first author because it's a clinical trial, it's meaningless for academic promotion.

Until that changes, unfortunately, we're going to have to rely upon either national

Page 88

professional societies to do this, the sponsors to do this, or groups like this. And I think we can't rely because it hasn't worked so far. So something has to be changed, and perhaps this group can do that.

DR. McDERMOTT: John?

JOHN: Just one quick comment that I made to Bob that I think is probably worth saying out loud is that there's clearly a big effort now on what's called team science. It means lots of things in lots of situations. but in this particular case, one of the issues that we know about investigator meetings is that everybody gets together at the beginning, and then the coordinators go off and do their thing and the investigators go and do their thing. And that's never made sense to me because we really want the coordinators and the investigators to, like, hear the same thing so that they can hold each other accountable.

So I would argue that the best of all worlds -- obviously, you can't always do that -- you would want actually to train the team

Page 89

to work together to provide the services necessary.
DR. McDERMOTT: Nat?
DR. KATZ: I have sort of a change of the subject, which is more back to the issue of central surveillance of clinical trials. I think we need to say something about what corrective actions in response to surveillance findings are and are not appropriate. If we're going to be monitoring for quality interpreted one way or another, then the next question is, well, if you find something, what are you going to do about it? Otherwise, there's no purpose to surveillance unless it's connected to some type of corrective action.
The risk-based monitoring guidance has a lot of information about possible corrective actions, but it's a very suggestive and non-specific and certainly not focused in our therapeutic area. And I wonder if folks on our panel maybe could comment on what types of corrective action are and are not appropriate because as we're designing these systems, we need to know.
DR. SCHUETTE: I think the -- well, there's

Page 90

two types of things. At the FDA, we get the data at the end. So corrective action is basically exclusionary from our perspective, although that can help. We're thinking of using some of these tools to help determine where we can send our office of scientific investigation inspectors.
Before it happens, though, during the course, if it's done at the sponsor level, certainly, there's an opportunity to go in and make an intervention either by going through and training, confirming what's going on and making sure that the site investigator is following protocol, that their staff is following protocol, that there's a potential for corrective action there.
I think if it's done sort of in combination with sort of a DMC approach, that there's certainly a possibility for, shall we say, rescuing some of the sites.
MALE SPEAKER: I think this is an important question. I mentioned before that in some ways you want to understand the nature of it and

Page 91

whether -- the success of the blind, I think was one important issue. But also the people doing analyses are the ones making decisions about whether there's an exclusion from the database or not. In some ways, I want those people blinded. I don't want that to be potentially based on treatment assignment, either.
I think there are consequences, of course, with exclusions. There was a mention in one of the talks earlier about, well, of course, you're going to lose power because it's going to have fewer patients. So there's one issue.
But it's probably not the biggest issue. The biggest issue is whether, one, if you analyze, say, what's left after you exclude, how generalizable is it or have you hurt generalizability because now you're selecting. Is it differential between treatment and is what I have left now a distorted view of what I started with?
So if the malfeasance is actually a result of poor results, I see poor results so I make them

Page 92

more positive. Well, clearly, if I exclude those, I'm excluding what was poor results, and if I include them, then I've got trouble.
So in some ways, you've got a real tough statistical problem because the data that you're either excluding or including is biased either way, and you've got real informative censoring problem going on. So that's where I think that much of the effort from a statistical standpoint, you've got a real hard problem.
Therefore, I think that as with the message with the missing data problem, that efforts are about prevention and avoidance.
DR. McDERMOTT: I get the sense that your question was about an earlier stage of correction, though.
DR. KATZ: Yes, let me maybe follow up with a more specific version of my question. What patient level corrections of problems that are observed in the patient performance in the clinical trial would and would not be acceptable?
So for example, yesterday we heard that if a

Page 93

patient is providing nonsensical data, we can't correct that on a patient level, and we need to accept that. I think that was the message we got from Sharon yesterday. The approach would be to sort of provide as much general training as possible across the board, and then cross our fingers and hope for the best without any type of for cause response in terms of retraining the patient on how to use the instruments more effectively. That was the message I got yesterday.

So I think that's one example of what I'm trying to ask in a more general way, which is what -- and we can talk about patient level corrective action or site level corrective action, which I think was the comment, Paul, that you made earlier, that site level -- what I heard from you, Paul, is that site level corrections in general sort of anything goes, right, except for unblinding and sort of obvious violations of the rules of clinical trial conduct.

I'm getting at least the beginning of an impression from you that, from your perspective,

Page 94

virtually anything could go at a site level. If there are exceptions to that, it would be great to know about.

How about on a patient level? Let's say, for example, the patient's not being compliant with their electronic diary. Well, is it okay to call the patient and say can you be more compliant with your electronic diary?

So it seems like there probably are some things that you would consider forbidden on a patient level as a for-cause response and some things that would be considered acceptable. I'd like to know what those are, and then the same thing on a site level.

DR. SCHUETTE: I'm not necessarily the person to ask for some of those. I think everything has to be consistent with the protocol.

Now, I think that's something that's actually discussed further, and maybe there can be other meetings where what can be done could be discussed. But I don't have a great answer.

The types of things that I see where things

Page 95

are going out of sync can be an issue, and I think Amy referred to one site where they basically said, "Hmm. Your adverse event reporting rate seems to be a little off," and just that intervention seemed to be sufficient.

So in some cases, maybe just actually being a reminder or actually part of a site inspection from the sponsor saying, "Show me how you're doing this" could be sufficient. So I think that's the level that I'm talking about. But that's actually something that would have to go through an entire process that's separate and distinct.

So I don't have a great answer for each and every aspect here.

DR. DWORKIN: This is on the agenda for this afternoon, this exact issue of what can be done, midstream course corrections versus what can be done legitimately after database lock and evidence of something funky is discovered. So this is pretty high on the agenda for this afternoon, Nat.

DR. McDERMOTT: I think that we're going to have -- I know there are other questions, but I've

Page 96

been told it's time for a coffee break. So 10:30, we'll reconvene. Thank you.

(Applause.)

(Whereupon, a recess was taken.)

DR. McDERMOTT: It's terrific that everybody is so stimulated for the discussions, but please take your seats and quiet down because we want to move on to the next session.

I want to congratulate you all. By attending this meeting, you are all now qualified at phase 1 to be a clinical investigator. However, to be a fully -- non-provisional -- qualified IMMPACT trialist, it is essential that you have to respond to the manuscript, which will eventually be drafted. And responding is not sufficient because the certifying committee, consisting of Dr. Dworkin and Dr. Turk, will evaluate the quality of your responses and comments to the manuscript. So simply saying "good job" will not do it. We need to have your input.

So thank you all very much. And thank all of the speakers, both yesterday and today, for

Page 97

really creating what I think, in hearing from the people around the room, from listening in at people's discussions, has been extremely interesting, stimulating, exciting.

So thank all the speakers for accomplishing that so far. We're really moving forward now. This session will take us up to lunch and then after lunch we'll really have an opportunity to spend more time discussing some of the kinds of issues that we've been talking about.

In the last panel discussion, we were really starting to segue nicely into what we want to do now, which is -- yesterday and early this morning, there was a perspective of quality in clinical trials that was really coming from somewhat of the ideal what we'd like to see, what we really need to do, what we need to accomplish. But there is another side to that balance, which is the people who are actually in the clinical trials trenches doing the work.

We started going into those things, and what we thought we would do when we organized this is to

Page 98

allow people who are sort of at the other side of this, who are actually in the trenches trying to do the work, to try and carry out the best possible clinical trial that they can given the realities.

I think the old adage that we need to be cautious about having the perfect be the enemy of the good is that we have to face some reality to what we can feasibly and appropriately do, given that we are knowledgeable about some of the problems that could occur, but yet we still have to find ways to get these trials done.

So what we're going to do in this session is take the perspective from somewhat in the trenches, from the clinical perspective, from the CROs, from the companies. We heard a bit about company perspective from Rick Malamut, so we'll follow-up on that.

Before I introduce our next speaker, I want to say it's John Markman. And I want to say, belatedly, happy birthday, John, yesterday. I understand you are finally now eligible to have alcohol consumption in some states.

Page 99

(Laughter.)

DR. McDERMOTT: Our next speaker is going to be John Markman, who is a -- unfortunately, he was told -- neurologist, and the bad news we heard from Scott Evans about what neurologists are. So John Markman is a neurologist, and he is director of one of the pain programs at the University of Rochester. He has been involved with a large number of clinical trials, particularly related to back pain.

John is going to, I think, give us some perspective of what it's like actually being there as the clinician and being the investigator in these types of trials and maybe having some reflections on what he has heard to this point and how it influences the concerns he may see about being able to accomplish some of these things.

So, John, you're up.

Presentation – John Markman

DR. MARKMAN: The big 3-0.

(Laughter.)

DR. MARKMAN: I just want to, first of all,

Page 100

thank Bob and Dennis. I think I come to these meetings -- and I know this was said yesterday -- and this is one of the most professionally rewarding moments of the year for me. I think I come out of this room thinking that there are a cadre of incredibly talented people around the world who actually can move the field forward and have sort of the knowledge and the reach and the wherewithal and the energy to do it. And so, I always leave these meetings energized.

So it's a privilege to be here, and it's certainly a privilege to speak here. I'd also like to thank Valerie and Andrea for shepherding us to this moment.

I'm going to try and provide what I will call an academic perspective. This is in italics. These are some of my relationships in terms of research, as well as consulting. I serve on DSMBs. I have served as a special government employee. But most importantly, for this talk, I've buttonholed most of the people in this room to get their opinions on my talk before I gave it, which

Page 101

is a real insurance policy against hostile comments.

(Laughter.)

DR. MARKMAN: So phone calls, long runs, cocktails, just about everyone here. So I think I'm in good shape.

So I want to come back to where Nat started when he talked about quality as the ability of the system to detect, in our case, an analgesic signal. And he introduced these two twin notions, one of scientific quality that had to do I think with the question being asked about analgesic signal. And then he sort of parsed this into a second concept or construct, which was regulatory quality, which was a little bit more about fidelity to the rules of the trial and the execution piece.

As a sort of preamble to my talk, I think all of us need to think about, to the extent that we subscribe to these two constructs, what is their relationship? Is it hierarchical? Is the scientific somehow superior or more important or privileged relative to the regulatory? Are they

Page 102

are on par? Are they on an equal footing? And does it matter the question you are asking? And the way they are balanced might depend whether you're doing a phase 3 confirmatory trial or whether you're doing a phase 2 exploratory study to develop differential response to a certain neuropathic pain phenotype. And maybe how you weight these things will be different.

The way I was thinking about more colloquially, if you have a hierarchical relationship, and you put the scientific on top, I think you sort of run the risk of saying that foolish consistency is the hobgoblin of little minds, and you sort of take the Emersonian view that somehow the details are not quite that important; whereas if you say they are on par, in my mind, what you're suggesting is sort of God is in the details, in the sense that when Flaubert said that, he was saying that with a really important creation or a significant scientific work, it's equally inspiring or equally powerful as you get down to the most minute details.

Page 103

When you look at the whole creation versus when you look at the particular sentence, in his case, it's equally well constructed. So I think we're going to try and kind of balance between those concepts.

The second set of concepts I think we're going to work through are also, again, as I alluded to, this notion about whether scientific and regulatory quality, whether these consideration weigh equally across different types of scientific questions.

In an explanatory trial, a trial with an explanatory goal, versus a trial that has a pragmatic goal, versus one that has an exploratory goal, and how these different types of questions that we ask in clinical trials -- I'm going to give you an illustration of that in a moment -- how do those different types of trials help us think about how we would emphasize quality more or less in a particular case.

So I'm going to start with this trial, which last fall, a lot of people were talking about and I

Page 104

found particularly interesting. And this is a small trial done at a relatively small shop compared to a lot of the trials that we've been talking about today.

This is a study looking at the effect of oxcarbazepine on peripheral neuropathic pain, and it asked the question whether a response to oxcarbazepine was differential with different phenotypes of peripheral nerve injury pain syndromes. And the take-home message of the study was peripheral neuropathic pain stratified by phenotype. And I know there are different views in the room on NNTs, but this is one way of illustrating there was a difference by phenotype.

Basically, the notion of this trial was the idea that the irritable nociceptor phenotype, which is related to sensitization of the unmyelinated cutaneous nociceptors, and where there was small fiber function preserved -- again, that has a particular -- and I won't go too deeply into the weeds here, but this obviously has a very particular clinical signature with regard to exam

Page 105

findings -- that would respond differentially to oxcarbazepine than a picture, which was the sort of non-irritable nociceptor phenotype, the deafferented small fiber picture with profound pain and temperature impairment, and oftentimes associated with allodynia.

Again, the hypothesis was that the compound would respond differentially to oxcarbazepine with these two different phenotypes. And what they found was very interesting. And I'm not going to go into all the details, but they found that there was a significant treatment by phenotype interaction for the irritable nociceptor group relative to the other group.

But it was a very aggressive dosing schedule. I think it went to 2400 milligrams, and there was a very large amount of dropout in this study. And as they say, the high dropout due to adverse side effects, which led to low power ultimately led them to do an analysis, which is very hard to follow in the manuscript, at least for me, and I've read it several times.

Page 106

They said that they used last observation carried forward, and basically they went from 281 subjects screened to -- what they would hope to have would be 97 randomized, and end up with only 39 subjects. There is no irritable nociceptor phenotype placebo interaction at all when you look at the data they provide.

As you learn further when you read the manuscript, 81 percent of the patients correctly guessed their treatment allocation in period one and 84 percent in period two.

So I think the question that this raises for me is I thought this was a very important study. I think I learned a lot, and I think it's a study that needs to be replicated and sort of revisited because it raises some important questions. So for me, there is a fair amount of scientific quality in this study.

I think from the point of view of regulatory quality, obviously, we just heard about LOCF analysis, we've talked about a study being underpowered. So clearly this does not have a

Page 107

sense of regulatory quality in the sense that we're asking of these large phase 3 confirmatory trials, where a drug is going to go into millions, if not tens of millions, of patients.

But there is an enormous amount of scientific quality in this trial, and I think it's going to be studied further and further. How do we think about how we approach quality evaluation in a study like this differently than in a phase 2 trial?

This is another trial, one that we recently published, which has a slightly different set of issues, but also was underpowered. And it wasn't underpowered because of drug tolerability. It was underpowered because one of the study drugs was pulled from the market, so basically we terminated the study early.

But this was a novel design, again. Our novel design is a single-dose design for a problem called neurogenic claudication, which is the evoked pain in the low back and legs that patients have when they are standing and walking. We used a

Page 108

paradigm where we put patients on the treadmill, and we'd only enroll them if we can induce moderate pain from a baseline of mild pain.

Basically, this is a phase 2 clinical trial platform that we're trying to validate. This is very exploratory. It's a single-dose design, trying to test a lot of compounds against this very, very common clinical pattern. And it was active placebo controlled.

We ended up terminating this study early. And the problem with that is that in terminating earlier, basically it became underreported to detect a 2-minute difference in the onset of moderate pain intensity.

But the fact that it was not sort of at the level of the threshold that we were looking for when we originally designed it didn't mean that we didn't learn anything. It basically learned that with a larger confidence interval, what we could really say was that the results suggested, in this case, oxymorphone and propoxyphene-acetaminophen combinations could not improve or did not have any

Page 109

evidence that it demonstrated basically providing any more than 5 minutes of low pain walking.

So we wanted it to be 2 minutes. It turned out it was 5 minutes. Now, whether that's clinically relevant or not and whether we should be allowed to change our prespecified endpoint is an important question to ask. But I do think for us, there is some scientific good quality to this. There is obviously not regulatory quality. So, again, how would we think about this differently, this flexibility?

So I'm going to talk now about -- with that preamble, I'm going to talk about the spectrum of academic clinical sites because I'm giving the academic talk, and then I'm going to talk about data quality, and then I'll talk about some future considerations.

So an academic medical center -- this is the academic talk. But we all know that what it means to be an academic medical center is something which is in an incredible amount of flux. There are 119 or so of them in the United States, and they're

Page 110

very diverse. They're very different. And they are rapidly becoming these regional networks where we have a set of laws how, which is turning them into either duopolies or triopolies in major cities.

That has important implications for patient fraud, of course, because that means that basically every one of us is going to be in one of these large systems' medical records very soon, if you're not already, and it's going to be -- and I know Dr. Rauck was raising this issue about let a monitor sit in front of your Epic console or your Cerner system.

But there are basically three large medical records out there. It's a very consolidated industry. The hospital industry is getting more and more consolidated. Physicians are employees, and basically there aren't going to be that many medical systems out there, and we're all going to be in these systems.

So I think some of the fraud issues are going to be harder and harder to achieve in this

Page 111

changing environment. That's one thought.

Another thought is that these are the largest employers and the economic engines in the region when you read their annual reports. I think that this has a little bit reprioritized the commercial interests in the output of a clinical trial, and that has implications when you're trying to do a small trial and there is maybe some potential intellectual property. It makes contracting harder. It makes it harder to be a small site because the level of scrutiny even any trial you do gets at the level of contracting. It just creates a whole other level of review, in my experience, and in terms of attention, and it makes a little harder to work in that environment.

There is also an increasing division of labor in these large systems where there aren't really that many clinician researcher investigators anymore. You're being asked to sort of differentiate from -- you're not going to be a triple set anymore. That really is going to go away.

Page 112

I think at certain institutions, like the University of Washington, where there is $1 billion almost of annual funding and sponsored funding and a few others, that may not be the case. But in the vast majority of those 119 medial centers, I do think that the imperatives of serving their local region and living up to the Affordable Care Act will not really allow for a system that has people who want to live this hybrid life where they practice 40 percent of the time and do research 60 percent of some variation there. At least at our institution, I think there's a lot of pressure to sort of differentiate further.

Then lastly, I think that as we've talked about a lot, there are less training opportunities in these environments, and one of the reasons there are less training opportunities is because we're moving sort of to a winner-take-all funding of the infrastructure of these places.

If your institution has a CTSA or a CTSI, you have a huge largess from the government, which supports this infrastructure not only for training

Page 113

future investigators, but also supporting a lot of pilot work where people get skills.

But the reality is that the grants that are sort of in the next tier down are far, far smaller, and the ability to support a robust clinical research infrastructure for a lot of those 120 academic medical centers is going to go away, because they're not going to be in the 25 or 30 places that get those big grants.

So I think that there is going to be an erosion in who is an academic medical center and what that means over the next 15 years, because unless there is a real change in the funding environment, I think what an academic medical center is will look very different. So that's just a simple preamble.

So what is it like in an academic medical center when you're running a research enterprise? Well, I think all of us probably have relatively similar structures at some level. There is an IRB, and there's a lot of, obviously, review that goes along with that, with the consent and a lot of

Page 114

ancillary reviews about risk and other things potentially.

Then there is the projects administration component, which has a lot to do with financial reporting and budget allocation and CMS reconciliation of care and those sorts of details. And a lot of the sort of quality checking there is just making sure that those two parts of the organization are talking and the data you submit to one is reconciled with the other.

Then there is the academic department level, scientific merit, which his very different from department to department. I work in a department with 11 faculty, with one clinical investigator basically, and there are departments with 200 faculty with 60 investigators or 60 people who are doing some clinical projects. And the review process for scientific merit is very different across those different kinds of academic departments.

Then at the site center, there is a lot of training for folks who do sponsored studies,

Page 115

because they go to study site trainings, and we send them off to organizational trainings through the ACR and others.

But the reality is there is not as much sort of moment-to-moment supervision of sites in an academic center for those that are doing sort of non-industry trials. There are some, but it certainly pales in comparison to the kind of moment-to-moment supervision you would have if you were, say, doing clinical work.

If you tried to take a patient with an OR and don't sign the day of surgery update, nobody will hang the bag of antibiotic. That's happening in real time. That's not some monitor call at the end of the day, that's not some query that you're answering. That's a query that's right now every second of every day.

So even though there are some checks, and the IRB will come and they'll audit your site and make sure you're being compliant, that's happening in a real time lag. It's not happening in real time, whereas in clinical practice at these

Page 116

institutions. It's literally happening on a minute-to-minute basis.

So again, what goes on in an academic medical center is I think very diverse. I'm the little dot, the little yellow dot down there are the bottom. I called this week to find out how many investigator-initiated and clinical research-sponsor studies there are at our institution. There are almost 300. We have $400 million of sponsored funding, and only 5 percent of that, really, $20 million, is from dedicated drug trials.

Now, there are different ways to account for that money, but it's a relatively small amount of the total research pie when you think about the organization and what their priorities are.

There's a broad range of investigators and sites. There are sites like mine, which are a single investigator doing a combination of sponsored research studies, as well as investigator-initiated trials, and then there's a group, of which Dr. Dworkin and Dr. McDermott are

Page 117

professors, where there is an entire clinical trial infrastructure, and they're leading multicenter, international, investigator-initiated trials in Parkinson's disease and Huntington's study group, which are of a whole different order of magnitude. They have their own in-house attorney and a core team of biostatisticians and a materials department.

You can imagine managing quality in these two different types of environments, just completely different efforts. The amount of quality control you need and what you can ask is sort of like being a public company versus being a tailor shop or a dry cleaners on your corner. You just can't ask whether it's in compliance for both of those structures.

So I'm just going to give you a snapshot of where my perspective comes from over the last five years. I've been a primary investigator in probably about 20 trials. These are mostly in neuropathic pain and OA. They have been in small molecules, they have been in oral drugs, IV drugs,

Page 118

biologics, device studies, and neuromodulation tools, abuse-deterrent opioid formulations, opioid-induced constipation studies, long-term open label studies, and obviously every possible design you can imagine in that area.

Then we have a lot of single-site investigator trials. I showed you a little snippet of one, a crossover trial for low back pain. And like many small investigators, I tend to do a lot of crossover trials. And again, those are largely in low back pain, but we also do trials with pain syndrome phenotpying.

We do them in outcome studies for our department and larger pain in the community, and we also do a lot of things related to service delivery in pain, for example, in urine tox screens, where for us the quality issue is really about the laboratory more than it is about other issues in that case.

But the main reason I do these sponsor trials, at least initially, was to gain the skill set and train my team to do our own trials, because

Page 119

that was really how we learned best practices, because one-half percent of the funding from the NIH is for clinical trials, some small, paltry amount.

So where was I going to get the skill set to do these trials? The only way to really earn what's done in industry and the best practices would be for me to do those trials and learn by going to investigator meetings, and sitting down with monitors, and looking at the protocols myself and trying to figure out what to do.

I know Lee has addressed this and others, but this is really the core issue. It was an on-the-job process where I learned one trial at a time.

Again, I think that the question here when we think about quality, and this is what I tried to raise in the beginning, is the attempts that we're going to take to minimize sources of error at the level of identification, at the level of prevention, at the level of management, again, may not be exactly even across these two types of

Page 120

enterprises.

You know that today is my birthday, and you know that I'm a Gemini. And we have the sense that there are sort of two Johns in the world, there is placebo John and there is assay-sensitivity John. Right? And I live these two lives, and my office and my clinical research space are the same space.

Neil Singla and I were talking about the challenges of what that might have impact on quality for and how that might affect assay sensitivity when you have patients come into that same environment. And these are the kinds of quality issues which a lot of academic sites contend with. There are a lot of sites that have their own clinical practice.

But I want to come now to the factory floor, if you will, because I see myself doing clinical trials more in the manner of someone who has a little independent bookstore and every morning goes out and hoses off the sidewalks, and people come in and look at the new books; or running a micro brewery, not some big Budweiser or Heineken

Page 121

brewery.
The nuts and bolts of our organization when we're doing clinical trials are the coordinators. They are running the trials. They are doing every one of those assessments, and their office is only eight feet from mine, and I'm talking to them all day long about all these decisions. But I think Rob pointed out this and it was very poignant to me, they are the guts of this.
So I wanted to ask them some direct questions, and I did this in the couple weeks before I came so we could hear from them.
By the way, you don't get to be a coordinator in my group unless you have gone to Catholic high school, unless you went to West Point. You have to be a very rule-oriented person. You cannot be someone who studied the hermeneutics of French modernism --
(Laughter.)
DR. MARKMAN: -- because that's not what I'm looking for. I'm looking for someone who follows rules, and they just are completely rule-bound.

Page 122

I found two incredible women who are incredibly rule-bound, and you need that. And I think that David made a great point about why you would be concerned that your physician might be your clinical trial coordinator? I think it's a total concern because clinicians are taught to use their own judgment, and that's not what you should be doing in a lot of these trials, as we learned yesterday even about training. You don't want people on the ground using their own judgment every moment. You want people following the rules every moment.
So I asked Maria, "What's the most important training experience you've had recently?" And here is what she said.
(Whereupon, a video recording was played.)
DR. MARKMAN: I got chills when I saw this video. I was looking at this, and I was like, oh, my god, we had this issue yesterday. Are you allowed to retrain people on the scale or can you only do it the first time? I was like, oh, my god, I'm going to show this video and people are going

Page 123

to be like -- but I think this has been something which has been very powerful, I think.
We talked a lot about patient engagement and making our patients our partners in research. It's incredibly important to make your coordinators your partners in research, and you can see how important it is to them to feel like they're getting more consistent reports and they're helping patients do that.
I think that rather than having them convince the patient that they're on the wonder drug or trying to guess what they're on, what they're actually do with the patients is coach them into being better subjects.
That's a fairly neutral thing, actually. I think about it, and I think that's actually a fairly positive thing. And you can see how important it is to them in their work because they are concerned about arbitrariness and randomness, and you saw Maria, the first woman who spoke, perfectly say -- I asked her does this make a difference. She says, "How do I know?" And that's

Page 124

the right answer, right?
She doesn't know if this makes a difference, and she's not doing it to make a difference on the outcome of a trial. She's not doing this to increase the analgesic signal of the trial. She's doing this to have a less sort of arbitrary interaction with these people, and that's her goal. And that's a more meaningful interaction with them.
So I think it's important to talk about what's your motivation for doing clinical trials within an academic medical center -- Neil Singla and I were talking about this yesterday -- as opposed to an independent research group that basically their business is doing clinical trials, whereas in my case, how do we choose which trial to do. I think that's important.
I know Rick gave a great talk yesterday talking about how companies and CROs -- we're going to hear from David -- choose sites. But here is how sites choose trials.
What drives my choices are what I'm interested in. I'm here today, and I'm interested

Page 125

in trying to develop better pain treatments for the problems that I see every day in practice for which I have to stare at people and don't have good answers, or I have drugs which are intolerable, unsafe, or don't help them very much.

So I tend to be in trials that are related to the indications of the target populations where I see the unmet need every day in my practice, and I feel it. And I also tend to be interested in being in trials where, as I said, my team is going to learn best practices and learn from being in those trials.

We're going to learn how to do an IV trial or follow the potential immunologic complications of being on a biologic. And I want my team to learn how to do that, and collect those samples, and send them down, and send them out, and store them in all those records, and do those follow-up exams.

We participate in a certain herpetic pain trial because I want them to learn how to use a tuning fork. But that's a lot of how I choose to

Page 126

sort of decide what we're going to do.

Now, they have a different agenda. They have an entire Excel scoring sheet that they got at their own industry conference, which is probably like this, just in a bigger ballroom. And it comes with a spreadsheet, which looks at different components of clinical trial complexity, inclusion/exclusion criteria, study design, screening steps, the study procedures themselves, the duration of the study. And you can see those bullet points down there. I just redacted that from their spreadsheet.

So they're looking at operational complexity. They're looking at -- and we both are, looking at the feasibility, can we get these patients? Can we keep these patients.

Then we're obviously also looking at the financial impact, and we've talked a little bit about financial incentives for folks in the system, obviously. My financial incentive for the system is to break even. I want to keep it going. This makes my clinical life richer.

Page 127

I think that I'm part of a larger project, which is important and helpful to people. And again, this is not sort of my core sort of compensation. This is no compensation, basically. This is just something that I think is important to do and offer something to my work every day on the larger purpose.

I think that's a different motivation to do this kind of research than at other centers, and I do think that may or may not matter. I don't know how impacts quality, but I think it matters at some level.

So we talk a little bit more about recruitment, because I think recruitment issues -- as we've touched on it a lot of ways. And I thought Dr. Kerns' comment about what we're really assessing for when we're screening patients for a trial is sort of how engaged they might be and things like that.

There's a whole covert set of screenings before the inclusion/exclusion criteria, I think, when a site is looking at patients, which are not

Page 128

as explicit perhaps as we think they are. And I thought that was a great, insightful comment.

This is from the WIRB application when you're trying to fill this out, but this is the only thing I could find about incentives for enrolling patients. Will the PI or the research team receive recruitment bonuses, yes or no? That's it in all of our work. And basically this is how WIRB defines a recruitment bonus or incentive.

But otherwise, the university doesn't really specifically ask me too much about it. There are some lines in our IRB, but it's not a direct question like this. But this is really all we have at our institution.

So the other issue around recruitment is how do we recruit? And I am a convert. I got into a sidebar conversation several years ago with Jim Campbell, who was here earlier, about recruiting for one of his trials. And he said -- I'm like, "What's the secret? How do you do it?" And he said to me, "Drive Time radio." I was like we're

Page 129

going to do Drive Time radio because that works.
I did a trial here -- this is just some graphs on -- we do an analysis on how we recruit patients to our trial to figure out how we're spending our time and our resources and what we're asking for.
This is from 60 weeks of recruitment of our first study with 260 screens, and we looked at the yield on patients who we looked to recruit from the office versus the folks we got from Drive Time. Of course, I'm interested in who is listening to Jimmy Buffett and who is listening to Rush Limbaugh and whether that's going to separate differently for placebo versus other listening preferences.
(Laughter.)
These are the deeper questions which I'm interested in. But we spent a lot of time looking at these patients. And again, we don't know if these are better or worse patients.
I have spoken with some of you about this before. There is some experience, I think, in the psychiatry literature about how patients who are

Page 130

recruited by advertising might perform differently as a clinical subject than folks who are not.
We have found this to be an incredible evaluator of recruit patients. I think as Neil made the point earlier, screening patients is extremely laborious, and screening patients from a Drive Time radio ad is much more laborious because we don't have their source records. They're not in our electronic medical record. They're out there in the world driving around the freeways of New York. So we have to go collate and get all that information, and we have to rely on them to verify what we do in a phone screen. So it is enormously labor intensive, but it's very high yield for us. When I say high yield, I mean we've got like 10 patients out of 200-plus screened.
Here is a little graph just about the different reasons why patients don't want to participate in our trials and how that shapes up whether they come from radio or they come from our office.
So we do this analysis on every mode of

Page 131

recruitment, whether it's Twitter or anything else, and we think about what are the reasons that patients turn us down, and how is it different, whether it's concomitant medications or comorbid pain conditions, and how is that different for the radio group versus our in-office group.
There is a lot of artifact there, and it's only 260 patients, and it's a small sample, so I'm not presenting this as hardcore statistical data. I'm just telling you this is kind of how we approach it as a shop to think about.
So I think academic sites in the future are going to have potentially some advantages, maybe. Again, I think that one of the questions that we face, though -- and I heard the rap yesterday in Rick's talk a little bit about I think our site has, in the past, had some low recruitment in some studies and a few others, we've been the highest recruiting site.
But the question is what is the low recruitment and how does that relate to quality? And I definitely think, in my own personal biases,

Page 132

there is definitely a relationship, especially with these large phase 3 trials.
When I'm doing a 50-person crossover trial, obviously, we're going to do the recruiting. We're going to see all those patients. We're going to know our own protocol. But I think that the first person we enroll in a protocol on the 10th, we're just handling that differently. We know the receipt, we know the drill, we're doing it again and again.
Everyone is kind of -- we have the flow diagram up on the wall. Everybody can just point to exactly where the patient is, patient number 9. It's not the same way with patient number 1. And that uncertainty affects every interaction. So I definitely think that being too low recruiting is an issue with quality, and I think that -- again, one of the reasons I got interested in the Drive Time radio is because I thought it was a way to improve our quality, because as we get more patients into the trial, we'll be better at doing that trial, and that matters to me.

Page 133

So I do think it's important for a lot of these private sites that can recruit a lot of patients. Obviously, there are dangers about the incentives, and I understand that, but I also think there is an upside to volume. And I don't know what the magic number is, and I don't know where you cross the threshold for quality, but I do think it's a compelling issue, in my opinion.

Again, we try to do -- one of the other ways we try and address the recruitment issue is we do multiple drug trials in the same population with similar indications and the same drug class. And again, that develops our expertise in dealing with that drug class.

We've done the COWS now in six different trials. We're good at doing the COWS, right? We can do it on an iPad. We've got all these little tools that everyone develops, but we can do the COWS, and we're good at it. And my clinical coordinators are good at it, and I know when they do it, it's done right.

For me, that gives me the confidence that

Page 134

we're doing a better job, because we've done six o these trials, and we use the same instruments. Again, these things are very helpful because they are not specialists in managing opioids. I'm that guy. So the more comfortable they get, though, the better they are at handling every question and query. So that's why I do a lot of overlap in what I cover.

So let me just talk to you a little bit about documentation. Let me come back to Valerie. I approached Valerie from the cardiac surgeon. She has been a cardiac CCU nurse. She worked and did cardiac surgery, different types of protocols with valves and heart replacement for many years. And then later in her career, she came to us. She is an incredible coordinator.

I think that one of the challenges -- one of the challenges of clinical medicine, but also designing clinical trials, is how much you're going to put upon people who are actually doing then trial.

When I work in the neuro ICU and a patient

Page 135

comes in, and they have an altered mental status, they might have had an intracranial hemorrhage, I'm concerned about that patient overnight. And I can tell the nursing staff to do q 10 minute neuro checks, q 10 minutes. Every 10 minutes they're going to wake the patient up and shine a light in their eyes, and disturb that patient's sleep to see if their pupil is bigger or lower.

That patient is going to have a terrible clinical outcome because they're not going to sleep because they're going be woken up every 10 minutes to look at their pupil.

Now, I'm anxious about that patient. I want to make sure that patient gets a lot of surveillance. But that's not going to help that patient, and it's probably not going to change the outcome.

I think that that's what Valerie is getting at in this thing, and she's doing it in her own way.

(Whereupon, a video recording was played.)

DR. MARKMAN: So she will stay there until

Page 136

9:00 at night until every query is answered. She is not going to go home until they are buttoned up, and I know that, and it's every day. But when you ask her to report in when the blood pressure is 139 one day and 141 the next, and she's got to deal with that, and she's got 50 of those to deal with, you can see that frustration and what that does to her work.

So again, I think Ajay mentioned earlier -- and I know this came up with Rick's comment, as well -- about the sort of data infrastructure that academic medical centers are developing and whether that's going to have important implications not only for source documents and looking at people's medical histories, but also important implications for fraud.

This is just a diagram of Epic, which is the dominant large health care system, electronic medical record in the country; i2B2, which is able to scrape that. So I can go look for -- for about 400,000 people, I can go look and see who is on

Page 137

300 milligrams of gabapentin and who has a neuropathic pain diagnosis right now with the click of a button. Now, there are some IRB issues around that and recruiting through that if I'm not touching that patient clinically, but we have the capability to do that now, and that's only to get more and more robust.

So I think that's going to have enormous implications not only for source document verification, but also for identifying fraud and also for recruitment ultimately. And obviously, the goal is that we'll have an implication for doing pragmatic clinical trials in the clinical record, as well, down the road.

So better recruitment and optimize source document verification may be something on the horizon in a consolidated health care system where your academic medical centers change.

So just to come back to the opening point, again, are the quality considerations for investigator-initiated clinical trials somehow different or related to the clinical question or

Page 138

the design of the trial or the number of subjects? We know there will be no central statistical monitoring for a trial with 20 patients in a crossover trial. That's not a feasible recommendation.

So the question is what are the types of things which could be equally prioritized in an investigator-initiated trial, and what are the things which should be less important?

In one sidebar we had yesterday, obviously, the quality issues as they relate to the primary endpoint or adverse event reporting I think are going to have to be of the highest priority and on an equal footing everywhere, in all trials. But again, quality issues related to the signatures on the CVs an whether those are up-to-date, those administrative issues, may not be the core issue that's going to improve the ability to detect the analgesic signal or test your hypothesis in an investigator-initiated trial.

So I'm going to stop there. I want to thank you all for your invitation and your attention.

Page 139

Appreciate it.

(Applause.)

DR. McDERMOTT: Thanks, John. I think that demonstrated that some neurologists are able actually to be listened to, incredible, and not liars. At least I'm going to assume so. However, we can do a reliability check because we have another neurologist who is going to come along, but now give us the perspective from -- in contrast to the academic medical center, the perspective from CROs, but also from industry about how they think about these things and what they have to do day-to-day in their actual operations.

So it is my pleasure to introduce Dr. David Hewitt, who is a card-carrying neurologist.

Presentation – David Hewitt

DR. HEWITT: Well, thank you for inviting me here today. This is something that is really important to me. In fact, I was at Merck, and some events that happened at Merck, actually probably related to the quality of doing one particular trial, led me to the interest of going into a CRO,

Page 140

at least for a while to try to understand some of the things that happened.

So in terms of my disclosures, I am currently working with inVentive, which is a contract research organization/commercial contract organization combined. As inVentive, we are involved in multiple large and small biopharmaceutical company studies and involved in a lot of commercial work, as well. Before that, I worked for Merck, and before that, Johnson & Johnson. So most of you know me.

I wanted to talk about a few things. I have a different perspective a little bit. Part of it comes from being in big pharma. I've only been in the CRO industry for about a year. So I still carry a lot of my Merck view of the world. I haven't lost that yet, and maybe a little bit of my J&J view of the world, but definitely Merck oriented.

I want to talk a little bit about quality by design, the need for -- that it's more than just a good protocol, and all the quality that I'm

Page 141

interested in, which is really not just the quality, I should say, of the clinical trial and the integrity of the data, it's really the quality of the product that you're going to bring to market that people are going to take, and it's going to potentially impact their lives.

That's important because I do think that the higher the quality is, before it even hits a patient and before it gets into a phase 3 study, the higher the quality that phase 3 study will be, as well. That would be my argument.

I'll talk a little bit about investigator sites and trial execution, but I think a lot of the points have been made already, so I won't belabor them.

I do want to talk a little bit about the differences -- and I do think this affects quality -- between big biopharmaceutical companies and small. One is that in big pharmaceutical companies -- and some of you can correct me if this has not been your experience, because I know we have got a lot of big pharma here -- is that there

Page 142

is a significant time to think, rethink, reconsider decisions, and then change things again. And I think that that churn that exists within big pharma, which I'm sure we're all familiar with, can really impact quality.

Big pharma, there is our complex decisions process with significant input from a hierarchical reporting structure, and I think there are pluses and minuses to that. It does ensure that the protocol at the end should not need to be amended too many times because it's been looked at a lot.

Small pharmaceutical companies, there are a number of stakeholders, and they can be very influential. Ultimately, there is a different intent with small pharma companies. I think it makes it different, particularly the small, almost virtual companies.

I do also want to make the point that I think money is important. Money impacts quality. There is no question in my mind that it does. We bid with a lot of companies right now, and they don't want to spend like a little extra money for

Page 143

like major things, like rater training. They say, "Well, we'll just pas by the rater training. You know, well just give us the Reader's Digest version of that."

The other thing is that small companies burn through cash verify quickly, and they can actually lose money as they're waiting for the first patient to be entered, which I think is a really stressful experience for them.

Also, their goal is often to be sold or to go public and really to make money in a relatively short period of time, which I think can influence them.

Big pharma has cash, but there are limits. They're not going to invest in everything. And the important part of cash, I think it was kind of mentioned previously, is that sites might enter patients into clinical trials based on how much a clinical trial actually pays, so they may preferentially go and enroll patients into the study that's paying more than the one that's paying less.

Page 144

Also I should say, that they are more likely to enroll patients in the study that's less complex. And we all, as bright people, like complexity.

A little bit more about contract research organizations. Obviously, there are issues with contract research organizations. Some of you work with them closely, you have positive experiences, negative experiences. My view, in general, is whether you do a study internally or externally, you're going to be complaining because it's not enrolling fast enough or there is some issue. I don't think it's germane to whether it's a CRO or a big company.

But for CRO, there is a focus on study execution, quality and speed, and the idea is that if you do one thing over and over again, whether you're at the medical monitor level or you're at the CRA level, that you're going to bet better at it.

You also have a project team, and the CROs benefit from the experience of having run multiple

Page 145

trials in the same area. On the other hand, the pharmaceutical company, the biopharm company, this may be the first time doing a study in Parkinson's disease or neuropathic pain. Maybe they haven't done a neuropathic pain study in five years, so there's a reason that a CRO may be of some value. And for small companies, there is really no other choice. They need to use a CRO.

Now, one of things I have as underlined, in italics, and shaded is the CRAs. And I think the point that John made before about the study coordinators is also made about the CRAs. The CRAs are where the rubber meets the road for clinical trials. That's the ability for the sponsor to have eyes on the ground. You know, the eyes thing? That's it.

So you really do need to have great CRAs. And I think if we can -- I could spend a lot of time talking about CRAs and study coordinators because I think they're hugely important, and they need to be experienced.

I think somebody said, "Oh, they're

Page 146

inexperienced and they're underpaid." Well, that's true, except at companies where you don't hire people who are inexperienced and underpaid.

I left the meeting yesterday, in part, to talk to a sponsor and say, yeah, the reason these people cost a little bit more is because they're experienced, and they've done this multiple times, and we get into conversations like that. And they really are the ones who are the control of ensuring the quality of the data, and I really can't stress it enough.

If you're a CRO, you know sites, you know them well. I mentioned this earlier to Nat's comment. You get to know how they respond to your queries. We also had a CRO no contracts. Contracts would be a killer, as many of you know. So if you know the sites, you know the contract issues, that could be hugely important. We have regulatory experience.

Then there is this ability to have what I think is the most exciting part of being in a CRO, for me, is I get to see all these clinical trials

Page 147

from all of these sponsors, from big and small. And I get to say to them, well, you know what? Maybe you should be using Bob Dworkin's algorithm for baseline pain.

I will throw that out there, and some people like to say, we're going to actually revise our protocol because we think that's important. And I actually had a big company do exactly that, revise their protocol just based on adding that in. But we do other things, as well.

But I think there is a huge actual joy. If you're into clinical trial and clinical trial methodology, being able to see this is great. And of course, being within and covering such a large group of people, I get to see everything from Duchenne muscular dystrophy, to Alzheimer's, Parkinson's disease, and, of course, pain.

The other thing that's kind of interesting, as well, which I like where I'm at, is that we do have this contract commercial organization. So in every protocol that we look at and that we get, we actually look at the market and the need and the

Page 148

value of it.

What's interesting there is it really will change the measures that you may want to put in by understanding who you are marketing this -- who this is actually intended for. And I say market, and that's kind of business speak. But it's really is it going to really meet the needs of the patients who might potentially be getting this drug.

But I do think that quality clinical trials begins really early on in development. When you're a big pharmaceutical company, you have a portfolio of products. And if you buy another company, like Schering-Plough, you even have more portfolio products, and you have to decide what are you going to move forward and what are you not going to move forward.

In the absence of information, really good information, it's really, really difficult. Consequently, as many of you know, there are great drugs that are withering away in the vaults at big pharma that may never get developed, were

Page 149

interesting historically. But the goal is to really make these drugs safe and effective, and that's all about quality.

Now, there used to be this idea we want to take multiple shots on goal, and that I think was a risk to quality when you're trying to push everything forward. Now, I think in most pharmaceutical companies, there is a de-risking exercise that really does ensure the quality of the molecule or the compound.

We spend a lot of time demonstrating things like target engagement, proof of pharmacology, safety, really, in order to select the optimum dose. Again, I think this all really does fit into efficacy.

In terms of the quality of the clinical trial, I think there's a tension that we need to be honest about between stopping clinical trial development early and recognizing that a lot of clinical trials, a lot of drugs that are out there, have only survived because they have had one person who was really willing to take all the shots, go

Page 150

the distance, and really be an advocate for the molecule. And those drugs, like topamine, really became very important drugs. There are other drugs. We all have examples of them.

So that's a challenge as you're trying to focus in on quality. And I think there is a focus on the clinical development plan and life cycle management, where we're kind of focused right now on I think the registration study, but really there are a lot of other studies that go on, and there's a lot that really gets built in here.

Again, in this process, there is significant time to put in to de-risking; again, a significant time put into what are the customers' needs, how can this molecule be used and how should be investigated appropriately.

There are a lot of health economic considerations. One of the things we really haven't hit on today is that these drugs that we're developing cost a lot of money, and you really have to defend the cost of these drugs.

I was just at ASCO at the cancer meeting,

Page 151

and one of the issues that came up is they've got these great drugs, checkpoint inhibitors, that cost hundreds of -- like $100,000, $200,000 a year. And there are some implications about financial toxicity, as well as the toxicity associated with the drug. And that's something that comes from -- actually, from all of Sloan Kettering. It's a big issue for them now.

I think to ensure quality, it's important to talk to KOL. I think a lot of times people ask me for KOLs. Luckily, I know a lot of you in here. Some of you may want me to stop referring people to you, but I'm often referring people who work with IMMPACT, as well as others.

I think it's also important -- and I stress this point a lot, is I think we need to put in -- to really ensure quality, we need to talk to patients more. We need to get their perspective on what they want.

The FDA is really important, there is no question about it, but it's the patient, and they're really -- at the end -- let me just say

Page 152

this. The FDA, when we're talking about quality, that's what Congress created to ensure that our products that hit the market for patients are of a high quality. And I think that needs to be put into the paper, as well, because that's important. Historically, it wasn't necessarily that way 200 years ago.

But I do think patients have a lot to offer. They have a lot to offer in terms of what are they willing to put up with in terms of pain and side effects, what does success look like. You can take a lot of the lessons from cancer and put them into pain patients, and sometimes just having a good response is a pretty good thing and how many patients have a good response.

In terms of clinical trial designs, obviously, I won't go into a lot of choice on this. But there is a decision on co-primaries versus one primary, what the secondary endpoints are.

Sometimes the studies I think get very complex as we're trying to really put science forward. There a lot of people who want to do

Page 153

quantitative sensory testing, which I'm obviously an advocate for. I think they're very important. But in large phase 3 studies in the United States, it's hard to find a lot of centers that are really good at quantitative sensory testing, and it increases the cost of the study significantly. So I do think we need to be careful about, as we add these in, there are some risks associated.

But the big thing for me is the large number of outcome measures and the complexity of the study decreases quality. There is no question about it. If you have a complex study, and you are getting the study coordinator and the investigator really annoyed, and the patient is annoyed, and if a patient has to spend four or six hours at the clinic, I think your quality is going to decrease and decrease significantly. Patient burden needs to be considered an important part of quality, as well.

Now, obviously, sometimes protocol complexity is really important. I'm working on a very complex protocol, and it's to assess a very

Page 154

important safety measure. And it has led to huge complexity. There is imaging that needs to be done. There's lab work that needs to be done. It's huge. Obviously, I can't talk about it in detail, but it's big.

There is no question, as I mentioned before, that the PI would much prefer to work on less complex studies. Part of this comes from my own experience at Merck. Merck was actually known for creating very complicated protocols. We've prided ourselves on that. And then they kind of pushed back on that, our bosses, but it was really an important issue.

I'm not going to belabor this point. I think this was discussed before. I think to have a clinical trial that's really successful, you need to blind the patient and the investigator to as many things as possible, from what the entry criteria are to really when the study drug is actually started and when it is discontinued, because if they know, there is an expectation. If they don't know when the drug is going to get

Page 155

started and they don't know when it's going to get stopped, you are actually collecting data that is going to be, I think, potentially more -- less biased, I should say.

I do want to encourage people, I think adaptive designs are really important. I think they're very informative. I think they're very efficient. I did a very nice adaptive design study, while the enriched enrollment one that many of you know is clearly and adaptive design in some ways.

I did a nice migraine study where we could add dose at the bottom end, and it was really kind of cool. But it can be very informative to clinical trials; not necessarily approved for the FDA right now for phase 3 studies, but very useful in phase 2.

Definitely it's important for efficacy and safety, and I like to think it minimizes harm. I think anything we can do to minimize harm to patients is a huge thing.

The other thing I wanted to mention, I think

Page 156

we touched on this yesterday, was the importance of the placebo, but I also an active placebo. An active placebo combined with a placebo can tell you whether you've had a failed study or not.

As some of you know, I did a very large study, international, worldwide study with thousands of patients, unfortunately, that failed. We actually were able to demonstrate rasagiline, made by Teva, very good drug, that we know works in Parkinson's disease. We were able to demonstrate that it was no better than placebo and, in fact, in certain countries, placebo was a lot better than rasagiline and preladenant, which was the drug I was studying for Parkinson's disease.

So a failed study really does suggest a problem, and it was that failed study that really said to me, you know what, as I'm looking at career options and possibly becoming -- where I could go from where I was at Merck, where should I go? And I thought, CRO would be a good place.

It is kind of like the candy store. When I left academia and went into industry, one of the

Page 157

reasons I went in -- there were two reasons. One is because I wanted power, statistical power, that is --

(Laughter.)

DR. HEWITT: -- and then the other one is that's where the drugs are. It's like -- what is it? Willy Sutton? Why do you rob banks? Because that's where the money is. I went into the pharmaceutical industry because that's where the drugs are. So it's pretty fun.

But going into a CRO is also fun because that's where the studies are, and you get to see all -- I have like 35 people reporting in to me. Each one is running -- there's a medical monitor involved in at least two or three studies. So the number of protocols I see across everything from oncology to -- it's really, I think, going to make me -- and has made me better at thinking about clinical trial issues.

I wanted to talk about rater training, which I think is huge. I think this is one of the biggest things. And really, this is where

Page 158

the -- in this article, there has got to be a lot of focus on this, as far as I'm concerned.

The placebo effect is real. It's very real. As I just mentioned, it killed my study in Parkinson's disease. There is tremendous investigator enthusiasm for the drugs that we are investigating. They're novel mechanisms often. It's really exciting. It's hard to contain your enthusiasm.

You're looking at somebody who, in general, finds it hard to contain his enthusiasm, and my patients saw that I was always concerned that I might be increasing the placebo effect.

But there is also something that to me is equally important, which so therapeutic misconception. And we've kind of talked around this issue a bit, but it's a really important issue.

One of the things I was really happy with coming to my new job at a CRO is because they had a rater training group, and I was going to be able to educate them about therapeutic misconception.

Page 159

Unfortunately, they had already had videos, and they were way ahead of me already. But this was something that was really exciting to me that I didn't even know existed when I joined.

So in this case, patients confuse participation in clinical trial with the primary care for their condition, and this is done all the time. The study subject is really a partner in clinical research and not a patient.

I know we keep referring to them as patients, but I think there are ethical issues that -- I'm into ethics -- need to be recognized, that the power relationship between a patient and doctor is sacrosanct. That is a very special relationship. And once a patient goes and makes the decision that I want to do you, Dr. Hewitt, a favor and become a study subject, that leads to a whole other set of ethical considerations.

I think we need to respect them because they're really -- and by doing them and making that study subject our partner, they understand their role better, and that is part of what rater

Page 160

training does as well.

My other concern, of course, is that in some countries, that are underinsured, clinical trials may be a way for people to get primary medical care. And of course, as many of you know, this was really an issue and a concern in India.

Obviously, the KOLs and PIs know a lot about how to assess pain in Parkinson's disease and all of these things. But on measures where the investigator matters, you really need to make sure that there is consistency, that everybody does it the same way.

I can tell you some of the world's greatest Parkinson's disease experts go into battle about how you do UPS Part 3. And it's kind of funny to watch when that hasn't been your whole thing, but it's very important to get consistency of assessments across sites.

You need dedicated teams with real expertise in rater training. You need to ensure the quality of the data collected. So they don't only just do the rater training, they do a lot of the things we

Page 161

talked about before, which is they look at consistency across measures, both qualitative and quantitatively, and that's a huge issue, too much for today. And it's already been touched on, but I can't stress that enough.

Ensure that all sites are thinking the same way. We want to provide significant materials to understand the placebo effect. So we have videos. You guys were talking about what you did live, which I think is really good. But we have videos that people can watch over and over and over again. And we say how often do you have to watch that placebo video? And we'll say, well, maybe we need to do it once every month or maybe it's every two months, but it's one of the things we talk about a lot.

There is ongoing training of patients. We've talked about this in the meeting. There is no problem. Part of being a medical monitor is to monitor the study. If things are not going well, you fix them. You still have an intent-to-treat analysis, you can't change the data, but you can

Page 162

fix and make the data that comes in following that intervention much better.

You have to train the PI and their staff. When the receptionist says, "Mrs. Dworkin, you look wonderful today and that medication must be really working for you," you know that you are pushing the placebo effect. Again, I can't mention this enough. Patients are our partners more than patients, and the question is does the patient exist?

Patients are study subjects with a specific skill set. I want to get back to this because we talked about this yesterday. We have inclusion/exclusion criteria that enrich our population and make our population different, much different than the general patient out there, and we can go into that. But there is no reason -- and we've already done this.

We select out people based on their pain intensity with the algorithm, for instance, or with their ability to -- in a run-in period, their ability to do an electronic diary. We certainly

Page 163

can exclude them based on other factors if we don't think that they're going to be able to do well, and if they can't be good observers of their own experience.

So one of the questions that I ask myself a lot is why do patients participate in clinical trials. I don't know the answer to that. I myself participated in a clinical trial when I was in medical school at the University of Rochester, and I know why I did it. I did it for money. I had somebody put a catheter -- put a tube down in my lungs, and they collected the macrophages from my lungs for a very interesting study. But I knew why I was doing it, and it was not really to better humanity. I wanted 400 bucks.

So what is the benefit for the patient? There really is none. It's really an altruistic thing, getting back to concepts of ethics.

Are they really seeking primary care? I mentioned that before.

Are they refractory pain patients? A major concern of mine.

Page 164

Are there comorbidities that will impact the results that the patients have? We've hit upon that, as well. We've hit upon professional patients.

Should subsets of patients be assessed based on QST or biomarkers, stratification based on these things?

These are all very interesting kind of endpoints that one could look at to verify the patient has what you think they might have, but also to ensure the quality of the study. John did a nice example of that with, I think, the irritated nociceptor as an example.

Now, I want to move on to sites. Sites are really important. What I always think about is that we don't do basic research and have lab rats. We basically have an extended team. And so this is the way I think about it.

I have a team that is very large, and sometimes we have 50 sites, and all those investigators and all those people are part of my team. And when I go into investigator meetings,

Page 165

and some of you may have seen me do this, I say, "You're part of our team." One of the things I want to make sure is that everybody feels like there is skin in the game. I always talk about skin in the game, so everybody is going to do the best quality work they can do.

So it is important, though, in terms of sites, to have geographic diversity, but too much diversity can be a problem. I think a lot of places, like Eastern Europe, have gotten much better at doing clinical trials, based on my experience, over the last 10 years in industry. But still you need to pick the right sites.

In Latin America, there are still issues, and we can talk about this. The placebo effect is clearly regional. If you want an example of that, I'll send you a poster on my Parkinson's disease study.

There are different practice patterns when it comes to what we're treating patients for, and that can be impactful. There are different startup times, which can impact study execution, as well.

Page 166

And then there is -- I think this point of controlled drugs has been mentioned. We talked about academic versus quality private research institutions.

I've changed my view on this. I used to think that academic institutions were the best, and I used to think that these -- what did we call them -- mills, these drug mills were horrible and that they were really bad.

But then I got to meet some of these guys, and there are sites and there are networks that are truly amazing, where they pride themselves on the quality of their work, and they will get the neurologist to do a neurology study or an -- well, you need an anesthesiologist to do a lot of anesthesia studies. But there really are Parkinson's disease experts doing Parkinson's studies. So they're very serious places, and there are more of these.

Sites are important. A large number of sites can be certainly useful to enroll studies faster. Contrary to a lot of the thoughts that

Page 167

have been doing on in this room, I don't think there's a problem with studies enrolling too fast. And one of my folks who are in industry might want to contradict me on that, but I've never seen a study enroll too fast because people are so eager to get patients in and are pushing that hard.

But I do think the longer there is an issue, it takes sometimes a long time to get sites up and running, particularly after the investigator meeting. There are issues with contracts. I mentioned some of these issues, as well.

But I did want to mention one of the points that Rob mentioned, which is I do think it is very important to visit the sites and to know the sites and talk to the sites. And I like to have a lot of conversations with sites, particularly when things aren't going well.

Again, there is huge value in face-to-face investigator meetings. I think that they know that you really are -- basically, when you're dealing with these drugs, it's like giving your baby to a stranger or to a babysitter, and you shouldn't be

Page 168

doing that lightheartedly. And I think one of the issues is that there's this thought that if you go to a CRO, you give it to the CRO, and you go away. I don't like that method. I don't like that interaction. I think you really need to have a partnership as you do these studies.

I think that part of it, whether you're a CRO or whether you're the biopharmaceutical company, face-to-face is better than on the phone. Relationships are relationships that can really last a very long time.

One of the things I do want to mention is we don't have enough sites in minority areas. This is one of the biggest failings of clinical trials right now. And I think to Bob's point, if you get your school up and running, I would like to get some hospitals that serve -- under the sites that serve minorities to really take the opportunity to put them in clinical trials, as well, because they really need to be.

I do want to mention that speed of enrollment really does negatively impact quality.

Page 169

We are always pushing, pushing, pushing, and when you do that, there is a concern that investigators may start to enroll patients who are not really truly qualified for your study.

We had one example of a study that was really positive in the first half, with a low placebo effect. In the second half, it was a huge placebo effect. It was really problematic. And we know that enrollment increases over time, and when it speeds up, we should be a little concerned.

I want to talk a little bit about diaries. My time is running out; I should be careful here. Paper diaries, I don't think they should be used anymore. Electronic diaries are very useful. They can assess compliance in the run-in period. You can avoid the hood effect that you have with paper diaries, where people fill them out on the hood of the car while they're waiting to come and see you in the clinic.

There is evidence that supports that diaries are not only filled out retrospectively, but in one study in Parkinson's disease, they were filled out

Page 170

prospectively.

(Laughter.)

DR. HEWITT: So that was a major new one to me, and I think that will be presented in the near future.

There is also the assumption that the more accurate the data is, there is an increase in quality of the study results. I know some people believe that a lot. I'm not sure if it's absolutely true, but it probably is.

Translation is an issue we talked about. Let me skip over that.

Obviously, safety monitoring is a big part of what we do, and the CRAs and the medical monitors and the MDs are a big part of that. I think this was covered pretty well previously in terms of alerts of these lab values, I think is important.

Efficacy I think is a huge issue for monitoring. It's hard to monitor efficacy in a blinded fashion, but there are methodologies that one can use to look at consistency of response on

Page 171

scales to make sure that they are aligned and correlated. And then you can look at site differences and regional differences, which I think can be helpful.

One of the big issues that we do in industry and CROs is really look at the inclusion/exclusion criteria because a lot of times they are not followed, and that leads to significant protocol, major protocol deviations. And those are high-quality problems, because then you're not really studying the population that you thought you were studying.

I do want to speak to the importance of DMCs. I think these independent data monitoring committees are huge. And if any of you want to participate in one, just let me know. People are asking all the time for people to do this. It's big. It's a big industry now, I think.

Of course, they need an independent statistician. You can assess efficacy and safety, and you can actually stop studies. But for efficacy, you can stop studies for futility, which

Page 172

I think is a very important thing, and I talked about that before in terms of interim analyses.

We've talked about risk-based monitoring. I think we've hit a lot of these things already. Let me skip over that.

So the data, I think this has also been hit on already, the use of flags in programming data checks. I do think one of the things to ensure quality is the use of soft locks. That's where you really lock the patients' data, and you really clean it up, really as you go along instead of trying to do it at the end.

Unless you're a pharmaceutical company, you may not know what I'm talking about. For those of us in the industry, that's a big issue because you spend a lot of time at the end trying to clean up data. So I'm a big advocate of soft locks. And to check, of course, the program before you finish.

We talked about this before, ensuring that people take the drug, and I was very excited by the technology that was just mentioned. I do think that the quality of a study overall begins way

Page 173

before even your selection of a molecule. It's really the intent of the pharmaceutical company. It's a decision on what mechanisms you're going to pursue and what molecules you're going to push through.

I think, obviously, you need more than a good protocol, but a good protocol is essential. We've talked about some clinical trial designs. I think there are some really interesting clinical trial designs that we're using right now, but one of the benefits of my position right now at a CRO is I get to see clinical trial designs from other areas, including oncology. And there are things like umbrella designs and basket designs -- I don't know, John, maybe we can talk about this at some point -- which are really kind of interesting to me, whether we could start to employ adaptations of those designs to pain studies, as well.

I think trial execution is important. But the last point is this. It really is a collaborative partnership that is the quality among sponsor, the biopharmaceutical company, the CRO, if

Page 174

there is one, the site, and, most importantly, the study subject partner, which I'll emphasize again.

I think with that, when you realize that it's really a partnership and that you have to work together, observing how to do that well is I think what really will ensure the quality of studies and the quality of the output, and hopefully the quality of drugs that are going to get to the patients who need them.

Thank you.

(Applause.)

Q&A and Panel Discussion

DR. McDERMOTT: As always happens, we're sort of tweaking and modifying the schedule. Not John Farrar. John Farrar, stay there, and Markman come up here.

We're doing some slight modification. What we're going to try and do is the panel discussion we're going to save until after lunch, but rather take a few minutes to have an opportunity to ask questions of our two presenters, and then we'll break for lunch in 15 or 20 minutes.

Page 175

Just a caution that I should have mentioned in the housekeeping details. This is being transcribed. So, David, to say you get fun from drugs, I'm not sure you want that to be in the record.

(Laughter.)

DR. McDERMOTT: It could be a concern. For those who have looked at their program and are expecting Sharon Hertz to be here, and I'm not her, in case you were wondering, she, unfortunately, was ill and wasn't able to come in today. So I've been filling in for her.

So let me start this off and then I'll let any questions. One thing that both of you pointed out, and I think really maybe we need to underscore even more, is we've talked about patient training and we've talked about site staff training. But I also heard both your two coordinators mention this, and, David, you mentioned this, is the recalibration.

Doing this once at the beginning of the study is probably not sufficient. This may be the

Page 176

kind of thing that should be planned into your protocol, or could be considered to plan into the protocol, both the site training and potentially -- and we can talk about it later -- about how you want to do this with patients, should that be planned, that is not to wait until it's a problem, but to actually plan that recalibration.

I know we're involved with some procedures that do require some physical examination, and what we've learned early on is that training our physicians to do these evaluations, if you follow them up two or three months later, they were fine initially, but they start drifting from that. So that that becomes important.

So I don't know if you want to add anymore to that or expand, but I think it's something that you both mentioned, and I think it's something that we haven't talked enough about.

DR. MARKMAN: We mentioned sort of the paradigm of prevent, identify, and manage. And it would seem to me that doing it up front is the

Page 177

prevention piece. But in the protocol, if you sort of had a remediated component, if you identify an issue that you think raises a question about how the rating is going, based on some statistical analysis or other observation, and then a dose of education to manage it, I think if that were prespecified, that might be a way of not creating the issue which we were concerned about yesterday, where education might be started on an ad hoc basis, and that might be even be introducing variability from subject to subject.

So I think maybe making that a little more standardized with some contingencies built in would be the way to go.

DR. HEWITT: I think it's huge. Obviously, we're talking about pain. But if we were talking about Alzheimer's disease, this wouldn't even be a question. Rater drift is recognized in a lot of areas within neurosciences as a really big issue, and you have to keep training people.

It can even get more complicated. You could have central raters who look at videos to make sure

Page 178

people are doing the physical exam correctly, or you can record people to see if they're really presenting the instrument correctly.

So how you do it, even how you talk to the staff about interactions and whether you're sort of gaming the system to increase the placebo effect, you can record sessions, you can videotape them, you can send them in to the central reviewer and see how that works.

But I do think whatever the -- particularly for the outcome measure, it's important to get training and re-training. And I think for a lot of times, even though we're doing electronic diaries, we think they're going to be better. I have to say, if I were a patient and I were doing it, there are going to be some times if I'm filling this out every day that I might -- I'll put down a 6. This is a 6 day. And I might not give it the consideration that it was really worth doing.

So I think it's very important to stress that with patients.

DR. McDERMOTT: Let me precaution, I can't

Page 179

see all that far back. So if, for some reason, you are sitting in, I'd say, the last two-thirds of the room, I may not see you by name, but I'll sort of point at you to call on you.

But in this part of the room, I think I saw already a hand going up. Yes, Rob?

ROB: John, one thing you mentioned I wanted to just follow-up on, and it resonates with some of the work that Nat has done about sites that don't enroll a lot of patients. And you hinted that the first patient or two that come into your trial, you're learning on.

It would be interesting to know if there's actually an evaluation of that first patient at each site and whether -- we do all learn with the first patient. Are there more mistakes made? Is there more data missing? Is the quality of that first patient or two worth looking at?

One thing we did when I was with a recent company is we ran a number of pilot studies, and you almost wonder do you want to run the first patient in as a pilot at each site to get the site

Page 180

greased. They have all been to investigator meetings. We know they don't pay attention until they have that first patient, and that's when they work out the kinks of the study.

But I'd be curious to know if, one, the analysis has been done. Even sites that enroll lots of patients, what do those first one or two patients look like when you separate them out?

Again, I know, Nat, you've done an analysis that suggested if you're a site and you only enroll one or two patients, that's not a very good site. You want the sites that have more patients. But I'm just curious about your thoughts.

DR. MARKMAN: I'm not aware of any data on how -- basically, how the detection machine is affected by that initial patient. But I agree with you, there is learning on that patient.

I think of the analogy as to surgery, basically. You don't want someone to do your Whipple procedure who does two a year. You're right. You want to go to someone who does a lot a year. That's how we practice medicine, and that's

Page 181

how we think there's a relationship in medicine, and we think there is some level at which quality improves after exposure.

I don't know what that is, and I don't know what the effect of that is on analgesic signal detection. But I do think it's an interesting question, and I've observed it myself and in our team.

I think the other countervailing factor, though, is I do think that the first patient you enroll, in our experience, tends to be a super buttoned-up crisp patient. Right? Because you've been searching for that patient for a while. Right? It's sort of like dating, and now you've got Mr. and Mrs. Right. And you could have screened 200 people to get that person.

In some ways, though, I think that that is a very crisp patient, and you tend to have less creep with that first patient. So I don't know, might be a couple of ways.

DR. McDERMOTT: David, do you want to comment? I'll get to you, Nat. I just want to see

Page 182

if David wanted to comment.

DR. HEWITT: No. I think that that's a really good question. In my experience, I have looked at the data in a few studies, and I haven't noted that the first patients are that much different from patients who come in later. I have seen studies where it definitely feels like the quality drops off over time and that people may be more cautious. Particularly, sometimes the CRA might be really close at hand, and they may be being guided by the sponsor very carefully to the CRA in those first few patients.

With that said, I do think there is a learning curve and people get better and better over time, particularly for good sites.

DR. MARKMAN: You're very invested in that first patient, though. You want that first patient to make it through. It's sort of like not making the sale at the local market to the first person who comes to your stand. It's a bad omen.

DR. KATZ: Just a quick correction in response to Rob's comment. That was actually Neil

Page 183

Singla's study with folks from Pfizer, where they showed that in three studies of pregabalin for acute pain, sites that only enrolled a small number of patients did not separate pregabalin from placebo, whereas sites that enrolled more than a certain threshold of patients --

MALE SPEAKER: Talk into the mic.

DR. KATZ: Sites that enrolled more than that certain threshold, patients all of a sudden on pregabalin did look better than placebo. I don't know, Neil, if you want to add anymore comments to that.

DR. SINGLA: No. I don't mind being confused with you, Nat. That's a compliment. We look so much alike, too, that's the thing.

(Laughter.)

DR. SINGLA: That's the summary of that study exactly.

DR. HEWITT: I just want to say one other thing, though, since this is about quality. I really believe that after the first two or three patients get in on a site, the site stops

Page 184

enrollment. So the CRA has to be out there, and they have to look at the data to see if it's quality data or not.

I think this idea, and I have seen this before, where people let patients enroll at five or six sites, just go crazy with enrollment, because there is such speed to get the studies done, is an example of quality being diminished.

I didn't put that in there, but if I was going to redo the slide deck, I'd say you need a visit after the first two or three patients.

DR. McDERMOTT: I can't see.

DR. JUGE: Dean Juge. In these discussions, you have academic and CRO as kind of two pieces, but in the last 10 years, I've seen kind of a hybrid, and let me explain that. It may be an issue to the companies and their sponsors when they're getting studies done in that you have a lot of academic areas or the academic PIs that belong to an outside CRO entity to get their research done.

So the patients are coming from an academic

Page 185

institution, but for whatever reasons -- and the biggest reasons I've seen, primarily two, is that it takes so long for an institution to get an IRB reviewed and approved, or the institution wants to collect so much funds from that for the institution, and not as much is coming back to your department, that it's easier to take outside if the rules of that institution allow it. But that could be an issue for the research organization.

For instance, at the University of Iowa, we were doing a study in a company I worked for in the past, and we had a sleep apnea study going on. So within the pulmonary department, they had to use internally, and it took forever to get the IRB, that they just didn't meet the timeframe and we had to drop them and contract on that end.

Yet, with the psych department in a TBI study that we were doing with the same product, they were using an outside IRB institution and conducting it through an outside organization.

So the principal investigators are part of an outside group, but yet the patients and their

Page 186

day job is within the institution itself. So it was really like they are there an academic, however, the studies are done through an outside institution.

In some of those studies, when you go look up their address or their affiliation in the study protocols, you'll see that affiliation is not the institution, but the affiliation is the organization they did the study through, although they are an academic institution and the patients come from that.

So I don't know if there are issues with quality or sites or whatever that happens there, but those are things that I have dealt with in the field, and I was wondering if you could respond to that.

DR. MARKMAN: I think that's exactly -- I think it's a very astute comment. I think that is what I tried to get at with this notion of what it means to be an academic is changing. And I think there's going to be all sorts of hybrid models and other versions which evolve, because, again, I

Page 187

think as the focus of these large academic medical centers and their networks change, which they will, I think their research priorities are going to be recalibrated. I don't think that's going to be at all 119 in the same way, but I do think that these changes are happening.

I think in our own institution, we use WIRB for a lot of our sponsor trials, which is pretty efficient, and our IRB is now run from someone who came from industry. So that's been incredibly helpful, or at least the administrative component, and that's been a real sort of accelerant.

But what has not improved, in my experience, in fact, may be just as challenging as eight years ago, is contracting. I just find that to be just incredibly laborious and frustrating, and it's literally months, and there is no reason it needs to be.

I think that, from my perspective, that's the biggest disadvantage to our own institution, and I think that, as I mentioned, the larger groups at our institution have their own attorney. And so

Page 188

for them, that person really is accountable to that department, but I'm using the attorneys and I'm using the individual contracting infrastructure that those other 300 investigators are using, and I'm just one group. I think that is the biggest rate-limiting step in our process.

DR. McDERMOTT: Let me just add one thing that might be related to this. I think you're right and you're right from what we're seeing. At the University of Washington, there is a new concept that they've been using. Somebody mentioned that they were a highly funded research facility. And that is the concept of dollars per square foot.

That is, they look at clinical space, and they want to know how much revenue is being generated on that clinical space. And the amount of revenue that's generated on clinical space on these types of trials is substantially lower that we developed if, in fact, John was doing a procedure on a patient.

So I think you're going to see more of that

Page 189

partnering out, extending out. I think that is the wave of the future.

Mike Rowbotham, I saw your hand up.

DR. ROWBOTHAM: I just wanted to say something about the contracting process. Dave, maybe you're in a position now to do something about it. But our contracts for Sutter Health are all done centrally, one office for all 27 hospitals, all the physicians. And we work hard to try and develop master contract templates with the major sponsoring companies. But then when they send the study to a CRO, the CROs insist on their own contracts, and that just hugely slows things down.

There has been an effort in California that I've been a part of for a few years. I'm not sure if it's ever really going to get off the ground. It's something called PACT, Partnership to Accelerate Clinical Trials. And that was bringing together 13 major research organizations from the state. So it was all five UC biomedical campuses: Stanford, USC, Sutter Health, Dignity, all the

Page 190

biggest healthcare organizations, to have a single contract template, a single IRB, and really try and streamline things and get all those barriers out of the way.

But it has actually been very hard to get industry to actually fund these kinds of streamlining efforts on a regional basis and to try and reduce the number of thinkers messing around with the contracts that make it take so long.

DR. HEWITT: Yes. Look, I think that one of the biggest impediments to -- actually, this is a good point. I didn't make a point of it on mine. I think that impacts quality, because when there is a huge time for site startup, particularly the time between the investigator meeting and the site being ready to get their patient in, I think that's a huge negative. I talked about this when I was at Merck, and I talked about it at inVentive, as well. I do have some power to make changes like that, so I have started to do that.

I think that one of the issues is it's not always as simple as it may appear because most of

Page 191

the time, if the CRO is going to be the one that manages the contracts, that's fine. But I think when it gets into trouble is when you have the CRO, the site, and the biopharmaceutical company all wanting to play with the -- because the sites change it. Most of the time, it's the sites that want to change things. And then you spend a lot of time with the churn.

So this has been a big issue, and I'm always working on this one. This is a continuous issue. So it's important.

DR. McDERMOTT: Roy?

DR. FREEMAN: A couple of things. I think I was little bothered yesterday with Nat's very creative fall below the threshold, intervene, coach, retrain, system, although I think there are ways of getting around that and doing it in a way that does not induce bias.

I liked John's Catholic high school graduates way of coaching each time the patient came in, provided that that is done globally across the trial. And it was fascinating that they picked

Page 192

on the average daily pain and we train them what it is. I'd love to hear how they explain the average daily pain.

MALE SPEAKER: Dr. Dworkin will be happy to talk to you at lunch about how you do that.

DR. FREEMAN: But the point I do want to make with that preamble is that training is obviously so very important, and I think patients need to be retrained, and I think sites need to be retrained. And here I'm going to -- the question is really directed at David.

There seems to me to be a reluctance on the part of CROs to actually get involved in the retraining process. And I know specifically with the bedside QST that I've introduced to a number of studies and have wanted for reproducibility, reliability, all of the obvious reasons, have the CROs get involved in training, and somehow there has been a block. They haven't wanted to do it. Perhaps the pharmaceutical companies felt it's too expensive and have made the CRO be the bad guy and say, no, we can't do it, but there has been some

Page 193

reluctance.
So I wanted to get a sense of the notion of the CRO not just looking at data and putting yellow stickies in, but actually training the sites on measures and assessment tools. That's the one.
Then the other thing is who actually trains the trainers? How well do your guys actually know the protocols? You're sending people out to sites, but do they know the story?
DR. HEWITT: What I would say is that, obviously, ours, there aren't that many CROs that have a rater training group embedded within them. As a matter of fact, I think we're the only one. So we put a lot of effort into rater training, and we always recommend it to every single study that we do, because I think it makes a big difference.
There are people who have been doing this rater training for a long time. They've had a lot of experience in the instruments, whether they are pain instruments like the BTI, or they're instruments for Parkinson's disease, like the UPDRS. So they have that experience.

Page 194

We, as a company, really believe in the power of retraining in terms of having quality data. I can't really speak for other CROs in this regard. It is true that if the sponsor doesn't want to have it, it won't get done.
In terms of who trains the trainer, it really depends a little bit on what the instrument is. So, for instance, this came up about QST, and we would get somebody who was like an expert in QST to train the trainer, like you, for instance.
So we would make sure that whatever it is that you do -- I mean, we'd probably have you do a video, probably have you at the investigator meeting, and we would probably have you and some people like you that you've trained and that you have -- because this is a really big issue. QST can be very different in different hands. You need to make sure everybody is doing it the same. So we've talked about this in studies where they've had QST.
In the year I've been there, nobody has done QST, but this is something that I'd recommend, and

Page 195

then I recommend that we have a way of checking it. So it's a really good point.
In terms of -- yes. I answered both questions.
DR. McDERMOTT: Laurie?
MS. BURKE: Right. I would like to take this discussion one step further, because it is really standard now that every measurement instrument is accompanied by a user manual and a training module for the reporter. That is part of what is required for review at -- well, required.
I can talk about required now because I'm not at FDA. Okay. But the user manual and the accompanying training module will have an impact on the measurement properties of any assessment tool. So in order to evaluate the measurement properties of an assessment tool, there has to be those accompanying modules, and they have to be standardized.
So it's not a recommendation. This isn't like something that's a good idea. It really is a best practice bordering on if you don't do it, you

Page 196

really aren't even implementing your assessment tool in a way that is scientifically valid and reliable.
So I'm having a really hard time in my current life because I'm working with companies who are developing clinician-reported outcome measures, taking tools and assessing patients. And we create a training module for the clinician, and they think I'm being unreasonable, and they don't want to sit through it. They all know how to do this thing. This is part of what they got trained in medical school, everybody knows how to do it. But we know that they all do it differently, and in a clinical trial, this is going to be a problem.
So I think this paper really needs to be much more strong than what we've actually been discussing here, and say that if you don't do it, it's contrary to standard recommendations. The PRO guidance talks about this. The qualification guidance at FDA talks about this. All the best practice papers put out by ISCOR on outcome assessment and clinical trials, all state this. So

Page 197

it's not just a good idea.
DR. HEWITT: Let me clarify that. I think the difference we should make is the difference between being trained on these things the first time, which is what obviously -- versus the ongoing continuing training over time. I think that's the distinction I was trying to make in response to that question.
MS. BURKE: Well, I think that that should be in the training manual. Whether you have found that you need to do this only the first time at baseline or whether this is something that needs to be re-administered throughout over time, how long does it stick in terms of the training.
DR. McDERMOTT: I'm going to take two more questions, and then we're going to go to lunch. And then we will come back, so that those questions that don't get picked up during these next two, which his going to be Bob and Nat -- and, Ajay, I saw your hand up, but you were the third. But definitely when we come back, we'll have an opportunity.

Page 198

Probably the best thing about question-and-answer sessions is when people leave with more questions, because that means that there's a lot of interest and enthusiasm. So Bob, and then we'll go to Nat.
BOB: David, I'm all for training, as you know. I think, first of all, there are two huge problems. And this has been alluded to, but I'll say it explicitly. One is it's not evidence-based. I take Laurie's point that it's important, but we don't have any evidence of any kind of improvement associated with training, that I know.
Secondly, every SRO does it differently. What you guys do at inVentive is not what they do at Premier, is not what they do at INC, is not what they do at Quintiles. So we have this incredible variability of what exactly we mean by training, and however it is being done, no one has ever systematically studied the impact of it.
I don't have an answer to that but I think it's a huge set of issues.
DR. HEWITT: I think that's an interesting

Page 199

point, because I've worked with -- when I was at Merck and J&J, I worked almost exclusively with CROs, and so I've worked on a lot of studies. I don't remember there being that much variability in the pain world, but you may be right. And I think it's worth looking into further.
In terms of Parkinson's disease, it's pretty uniform. Whether you work on a Parkinson's disease study done by -- you all have to have training videos, and you have to show them what a Parkinson's disease patient looks like in the on-state and the off-state.
Part of that training is it goes beyond what you do in pain. It really goes on to can I identify a patient who is in the on-state or the off-state. And you'd need to talk to the patient about whether their dyskinesias are troublesome or not troublesome. And so you need to talk to them.
So there is this training. There are two different types of training. One is the training that looks at intra-rater reliability or whether all sites are doing it the same way and when there

Page 200

is variability there, and do you follow that and test that. And you do do that. There are people who spend a lot of time looking at intra-rater reliability to make sure that there isn't discordance over time. I think that's one thing.
But the other thing is just to make sure people know how to do the study in the first place, they need to be trained. So if you don't train somebody for a study, they're untrained. It's not going to go well.
I'm not sure if you need to do -- I mean, there are some things that are kind of so obvious, I'm not sure they need to be trained.
BOB: I wasn't clear. I really meant training about our efficacy assessments, our zero-to-10 scales, and the other secondary endpoints in a clinical trial. And I think that varies all over the map from no training at all -- the patient is given BPI --
DR. HEWITT: Right. I see what you're saying.
BOB: -- to what Nat and Neil and we have

Page 201

begun to do, which is to look at the kinds of things that the FDA would require.

DR. HEWITT: No. I mean, I think that's a good point. I think that's a good point.

BOB: And it's huge variability, and it isn't evidence-based. I think at this stage, variability is good because one could at least imagine a study where Nat's training program is compared to our training program is compared to Neil's training program, and that would be really cool, and which of the three training programs have the best performance.

DR. HEWITT: I see what you're saying, yes. I agree.

BOB: None of that is being done. Right now it's all totally ad hoc.

DR. McDERMOTT: Nat. And then we're going to have lunch. But before we're going to go to lunch -- and don't run out the door -- I'm going to ask Valorie if she's got any comments. So Nat, then Valorie, then lunch.

Nat?

Page 202

DR. KATZ: I wanted to follow on Laurie Burke's comment in that I completely agree with her that anything that you want -- any instrument that you want the site to utilize to assess a patient, if it's important to assess that aspect of the patient's state, it's important to train the people doing it how to do it.

So I agree with that. But I wanted to point out that there is a certain trap there that is worth recognizing that I've run into once or twice, where I've run into people who say, "Yeah, we want to do a training" -- it's fine to use a training program, but if you're going to do that, you have to revalidate the entire instrument from the very beginning because the training program might somehow -- if you were going to put an 8, maybe you would have put a 7. It changes the performance characteristics of the instrument sufficiently that we can't really sign off on the use of that instrument.

If we wanted to stifle improvement in quality, that's the best way to do it, I think,

Page 203

would be to require that instruments have to be revalidated from the get-go.

So I wonder what other people think about that and whether that's worth addressing in this paper.

DR. McDERMOTT: Well, first, back to these gentlemen. Do either one of you want to comment on that observation?

DR. HEWITT: I guess one of the comments I'd make is this. We've had a lot of clinical trials over the years, and we've had a lot of drugs to get approved by the FDA. So I think one of the things that I think this kind of raises, to my mind, is what we're really trying to do is do it better.

I think the question we always have to ask ourselves is are there any drugs that have not gotten approved because there was not adequate training on an instrument. I think that's a question we can ask ourselves. Do we know that that -- have there been drugs that have not been approved, and then, conversely, are there drugs that have been approved that shouldn't have been

Page 204

approved.

BOB: David, I have to interrupt. What do you think happened with topiramate for DPN? You were involved in that. Your trial was positive, three other trials in Europe, as I recall, were negative.

DR. HEWITT: Worldwide.

BOB: Do you think if those European trial investigators had been trained, topiramate would now be available?

DR. HEWITT: I think that's a good point and I think -- yes, I mean, I guess just prove me wrong.

BOB: You couldn't have been more central.

(Laughter.)

DR. HEWITT: That is a good point, but -- I guess that is true that you can -- that that has happened. But the flipside is --

BOB: Three European trials killed your drug.

DR. HEWITT: I know, I know. That is true. That is true. But the thing is that even with

Page 205

that, I will say this, that all those studies, it could be -- I mean, to be honest, it could be that topiramate really isn't that good for neuropathic pain, and that our study was the outlier, and that the other three studies were better.

From an evidence-based point of view, I have to be -- obviously, I wanted my study to be the one that's positive, but the way the world goes is it's not necessarily the greatest -- I mean, sorry for those people who do topiramate. But part of the problem with that study -- we could go on a lot about it -- had to do with dropouts due to taste differences and to the fact of CNS, what they call Topamax.

So there is no question that the drug -- who came up with it? There was a great line that somebody had today about effectiveness, that it may not be -- it might be an efficacious drug, but it may not be an effective drug.

In that regard, it probably isn't a bad thing that Topamax isn't approved. It's getting used for it, as well, but good point.

Page 206

DR. McDERMOTT: I don't think we need to get into the details of the studies here.

Laurie, let me just have the last word -- you need to respond. Okay.

MS. BURKE: I need to respond to Nat. I think that there is revalidation and then there's revalidation. And so there may be some confusion about this. If you implement a training program that you think really tightens up your measure, then, of course, you would want to have additional test/retest reliability on that measure so that you have a better idea of how you interpret your clinical trial results. You would want to have that. You also would want to do exploratory work during your clinical trial to correlate with other related measures and do some construct validity.

I completely would argue with anybody who says that you have to start from scratch and revalidate from the get-go and go through your whole process. But there are some -- it is going to modify the measurement properties of the instrument. And, therefore, in order to know what

Page 207

the new measurement properties are, like test/retest reliability, you would have to re-measure them.

DR. McDERMOTT: Let's hold it there. You want to go back and forth, and that's great, but we do need to go to lunch.

So, Valorie, any comments that you want to make to people?

MS. THOMPSON: If you want to order a taxi, we will have a notice at your desk when you come back after lunch -- [inaudible - off microphone].

DR. McDERMOTT: Valorie was not in here when I thanked her and Andrea earlier. So I want to, again, thank you for all the efforts that you and Andrea have done to make this meeting useful.

(Applause.)

DR. McDERMOTT: So we're going to be going to lunch, which is back where we had lunch yesterday. Come back here at -- let's make it 1:30, to give you a little bit of extra time, and we'll then have the panel discussion.

(Whereupon, a luncheon recess was taken.)

Page 208

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

MODERATOR: Okay, please take your seats so we can get started again. And for those of you that are sort of wandering around, come in or you won't be able to get your certification. It would be required. Only your level 1 certification. Remember, you have to go up to -- it's provisional until you do the paper.

Okay. So we want to pick up with the panel, but I realize that as I was sort of rushing you to lunch, because I could see hunger in your eyes, that I may have short-circuited a bit the conversation that was going on with Laurie Burke and with Nat Katz.

So Laurie, I just want to give you the chance to -- in case there was something you wanted to finish off that I didn't give you the chance to. And then if Nat has any comment on that, and then we'll go back to the panel.

MS. BURKE: Sure. I think I had the opportunity to say most of what I wanted to say, but I think that there may have been some question

Page 209

mark in some people's minds about exactly what I did say.

The point is that Nat's question was do you have to revalidate the instrument once you have changed it in terms of attaching a training program and a user manual to it. The answer is --

(Laughter.)

MALE SPEAKER: The FDA talks.

MS. BURKE: -- yes, but not in the way you might think. The answer is you would want to know what the new measurement properties of that instrument are. You would want to know, in particular, what the test/retest reliability of that measure is, and what the variability, then, estimates are to help you in the interpretation with clinical trial results.

If nothing else, at a very minimum, you should do this testing at baseline in the week before randomization in your clinical trial. So therefore, you have an estimate of variability in your patient population. You'll be able to use that in the interpretation of your treatment

Page 210

effect.

At the end of the day it will give everybody more confidence in the particularly small-ish differences between treatment groups, and it can be a real advantage. And also, the other types of validation like construct validity testing can be done as exploratory analyses in the phase 3 trial.

Of course, optimally, you would want to do this in a stand-alone study, but few people want to fund that, and it's not necessary.

MODERATOR: Nat, do you want to amend at all your comment?

DR. KATZ: No, we agreed on that, so we're friends again now.

(Laughter.)

MODERATOR: Okay. Well you know when Laurie Burke wants to talk, everybody stops.

MALE SPEAKER: We're all going to listen.

MODERATOR: So let's go back to the panel. Just before lunch, we had had David Hewitt and John Markman sort of give perspectives again from academic and from the CRO industry. But I think

Page 211

what we heard was these are coming closer and closer together. So what I thought I would do is start by having Ian Gilron and John Farrar, any comments they have about either what they heard in those presentations in particular, also to chance if they want to drift off other places we might let them. But try and focus, at least initially, on what you heard from David and from John.

DR. GILRON: Okay. I was just telling John earlier I usually give unsolicited opinions, so I've actually been asked for it, so it's kind of exciting.

MODERATOR: I'm asking for an unsolicited opinion.

DR. GILRON: Yes, okay. Well I'm just going to start just with some general comments talking about the dichotomy that Nat started talking about and John discussed further. And I think of it, as well, from the perspective of, as John described, someone who does trials more from a proof of concept, academic perspective.

I think of the onus on the scientific part

Page 212

of things really to be exploratory and to provide appropriate guidance towards coming up with new treatments, whereas the regulatory approach is really one of public health responsibility and looking at cost effectiveness, meaning not to approve treatments that are only marginally effective and possibly very expensive, but also a big emphasis on safety.

So I was wondering that -- I don't know if we have spent enough time talking about quality with respect to safety assessment in reporting, and I know ACTTION has very involved in sort of waving the flag of improving safety reporting. So I was just wondering whether we need to, in the paper, make more noise about quality with respect to that part of things.

MODERATOR: John?

JOHN: So, I was taken in the presentation by John Markman of the comment of the obviously very talented coordinators that he has, relative to the conversation or multiple conversations that she'd had, obviously recently, relative to blood

Page 213

pressures of 141 over 82, versus 139 over 79.
It brought to mind something that I think was brought up a little bit yesterday, which is that we need to be very careful as we move forward, to meld a couple of ideas that IMMPACT and ACTTION have been working on, which is -- one of them is obviously to make trials responsible and conducted in a way that makes sense and that provides valid results.
But on the other hand, not to impose on those trials a burden of control, regulatory control, that doesn't actually benefit the trial and adds additional time and effort, frustration, to the process.
I'm reminded that many of our young investigators that come to us, when we ask what they want to measure, they basically say "everything." They want to know how the patient's feeling. They want to know a host of different things, and we need to remind them that we have only so many things that we can ask in any one trial. And I think that applies through a number

Page 214

of different components of what we've talked about today.
All of the things that we've talked about in terms of looking for potential signs and markers for different kinds of problems -- careful monitoring, visiting maybe after the third patient, doing central monitoring on ongoing basis -- are all important. But we do need to take them with just a few grains of salt and ask ourselves the harder question, which is, are we going to gain benefit from what we're actually doing?
Is the trial going to be better at its ability to differentiate a real effect between a treatment, either two treatments or a treatment and a placebo, assuming that that treatment actually exists?
So I really feel very strongly that we need to be cognizant of the cost benefit value of the kinds of things that we might impose to improve both the ability to detect efficacy and safety. I would just push that as a component of how we go about implementing all of the various steps,

Page 215

realizing that nothing is ever perfect and that we are always going to have to make some compromise.
We just need to decide where those compromises are and be efficient and comprehensive and careful to try and make the studies answer the right question and provide us with the data that we need.
MODERATOR: John, did you want to comment on that? I saw you look at me.
JOHN: Well I just thought -- I think John's point about a little bit of flexibility is important. I think that's some of what you're suggesting.
The other question I've had, and I've been really trying to get my head around is, at some level when you're doing one of these smaller trials, too, you're thinking about the next trial, because your hypothesis comes out of the observations of doing the trial you're currently doing. And I think that -- and I thought of this kind of vis-à-vis the issue of blinding a site to all, and blinding every piece to it.

Page 216

I mean you're really -- how much are you really giving up a serendipitous observation, which will be the germ of your next hypothesis, by disengaging everyone at the site from their little local observations?
It's sort of a little bit of tangential point, but it also kind of plays on this notion of how much flexibility and, again, the differentiation between a public health consequence of a trial versus a trial which is exploratory.
MALE SPEAKER: But I think your point is exactly the right one, which is that the question is what's the goal of the trial? If the goal of the trial is to approve a drug that's going to be used in millions of people, then you damn well better be sure it's safe.
If the goal is to know whether that drug works at all so that you might use it as a model for another drug that you might develop or you do formal testing, or if it's a pilot study, then the requirements are somewhat less.
So keeping in mind what the goal of the

Page 217

trial is, and you presented three different potential goals I -- one might argue that there are four or five. But if we understand what the question is that we're asking, we're much better at figuring out how to answer it.

MODERATOR: An important point that you're sort of implying, and I was thinking about this when I was sitting there was that we, in some sense, have been talking about and thinking about these trials as if they are, once you do it, you design it, you run it, you're done. But what we learn from those trials potentially influences the things we can benefit.

So when Bernard talked about the placebo issues and when we've heard issues about fraud from Eric Devine, as we learn those things, that feeds back to -- so that we improve the science. So take the information you're gaining, and in addition to whatever you find out for that one study, that information then becomes --

So if we talk about what can you anticipate and try to prevent in the next study, you'll use

Page 218

this as an opportunity to learn something and not about the study, study design going forward; and not just, okay, we've now done our study, we're done with that, and move off to our next study and start essentially as if nothing was really acquired.

I don't mean nothing, because obviously the outcomes people will be aware of, but to realize other things that came up along the way, problems, issues, difficulties, things you need to control, some of which may be controllable, some of which you may have to understand that are not controllable.

But I think we really need to think about programmatic research, where we used to talk about when we were promoting professors, it wasn't that they had one article or one well-known paper, but what's their program of research. And maybe we should be thinking more broadly about what's the program of a research related to clinical trials. And maybe that goes back to the discussion we had, which I think went into much more detail on.

Page 219

But if you're educating people about being qualified to be clinical trial investigators, what are we learning that we could then feedback to those people?

So unless any of you want to comment on each other's -- David?

DR. HEWITT: I just want to make a comment about what you said --

MODERATOR: Sure.

MALE SPEAKER: -- because I think that's really key is -- and I was kind of getting to this with the Six Sigma comment I made yesterday, is that this is really an iterative process, and it's actually iterative during the course of the study when you're retraining people.

But it's also iterative in terms of taking the lessons you've learned and making sure that you disseminate them. And so it doesn't matter if it's a bid defense for me or a clinical trial, I'm really kind of a stickler for having a lessons learned meeting after whatever's happened, as soon as possible. And I think in that way, you can

Page 220

disseminate that information across an organization or a group.

But a lot of it has to do with how groups and organizations work and how they get better at what they do. And I think you're right. It's if you reinvent the wheel each time for -- you shouldn't. You should have methodologies, and you should be realizing this is where you're going to have the potential to really screw up, so don't -- there's no reason to screw up twice.

MODERATOR: Well, I think this gives you a sense of why Bob and I and the rest of us who've been involved going back to the early days of IMMPACT and to ACTTION, was to say not only do we want to have meetings and discuss things and get things around, but how do we get that information out there so that -- and we'll be talking very shortly about a manuscript. It's because we want to make sure that the information gets there, so that whatever wisdom we've picked up along the way from the presentations, from the discussions, gets disseminated, and then hopefully better trials will

Page 221

come along and people will learn from this experience.

So okay, let's open it up for the audience. Any questions for our panelists? Yes, Rob?

ROB: So maybe this has been emphasized enough, but certainly one size won't fit all. I mean, in the drug development process at Wyeth -- no longer Wyeth, now called Pfizer -- we had this concept of learn and confirm, and whether those terms mean anything.

Learn, from my perspective, was explore. You're early in the stage of drug development. You want to work with small numbers of sites, very scientifically based questions. Confirm, not to label it as a regulatory requirement, but that really was the burden of a drug development program to look at all the safety issues and to confirm what you've learned in earlier trials.

The burden and the responsibilities of each of those two phases takes on a very different profile. You might have a Web-based medical investigator meeting once you've established sites

Page 222

and you've established the product, but early on you want to have face-to-face. You want to have very personal relationships with your investigators and your study staff.

Again, as we start to think about solutions, there may be some that run across both a learn and a confirm phase of drug development or even a post-approval. We haven't even talked about how you control trials after they've been approved.

But I think we need to think about -- there may be common threads, but there may be distinct differences.

MODERATOR: Anybody want to comment? Everybody nodding agreement.

MALE SPEAKER: There is one comment I want I want to pick on, which I think one first made, which was the difference between academic studies and pharmaceutical studies and about the drug.

The fact is, all the drugs that are out there, for the most part, that are done within academic centers are drugs that have been approved or -- I think for the most part. I think it's very

Page 223

hard to get a drug before it's been approved and do a study on it. You can, but I think it's a big challenge to do that.

So in a way, once you get through all that safety data, right, and the efficacy data with the original approval and the regulatory process, things ought to -- to John's point, things should be a little bit easier. There should be less -- not less rigor, but maybe less neurosis around the compulsion to cross every T and dot every I, because the drug's already been approved, and we know that it's safe, from a safety point of view.

Now from a quality point of view and demonstrating that your efficacy really is your efficacy, that goes back to I think the point you're making. But I think that's important.

MODERATOR: Let me just add one thing, and I'll get to you. We keep talking about drug studies, and I'd be interested in knowing whether there's anything that we've talked about, if we took the word drug out and put complementary

Page 224

medicine or put rehabilitation, other than unique features that are specific to a particular drug's side effects or the worry about other medication they're taking, is there anything different about what we've been describing for what would go into a well-done, well-designed, carefully-controlled, high-quality clinical trial, anything that isn't necessarily a drug?

MODERATOR: John? You've done some of those studies, complementary medicine.

JOHN: Right. I think that exactly the same principles apply. Procedural studies, however, add an additional complexity. If you have trouble blinding the patients to a particular procedure, if you have trouble figuring out how to standardize the application of acupuncture in your trial, if you are trying to compare surgical outcomes and you're using five differentiation surgeons, you know that you're going to have slightly different techniques by those various surgeons.

So I think there may be some value, and I'm not sure we have time today, but some value of at

Page 225

least thinking whether there are some additional issues that we might want to consider in those trials.

When talked about blinding, we all agreed that blinding in a pharmaceutical trial and making the pills look the same, maybe even -- I mean there was talk, back in the day, in the testing of opioids, of trying to use a benzodiazepine as a placebo, or putting a little benzodiazepine as a comparator, because in fact you wanted to make the patients a little sleepy to hide the side effects.

But in situations where it's really not ethical to do that, and certainly surgical and many procedural things fall into that category, I think there may be some other things to consider and how to deal with those.

MODERATOR: Blinding the control groups may be more challenging in a surgical study or a physical therapy study or an acupuncture study, but as far as the need to pay attention to the kinds of things we're talking about --

JOHN: I agree.

Page 226

MODERATOR: -- those would be -- let's say -- and the only reason I'm pushing that is that, to the extent possible, I try to make the papers that we write be as general and then refer -- when there are specific callouts to refer to those. So I would hope that we'll be able to at least talk about the importance of these in clinical trials and not in just in drug clinical trials.

I'm sorry. That's all I have.

MALE SPEAKER: So I would say I disagree.

MODERATOR: Okay.

MALE SPEAKER: I think that a big chunk of what we do, like biggest part of it is safety. It's safety, safety, safety. And the safety issues associated with a clinical trial I think are going to be different if you're looking at yoga, or bicycling, or tango. I'm giving you examples of things from Parkinson's disease world that have shown to be effective.

I think it's much different. I think it's a good thing. Now some of the things there are going

Page 227

to be overlap. Does every tango dancer instructor need to teach the person how to tango exactly the same way for the results of the trial to be important? Are you comparing hip hop versus tango? Will it make a difference?

But I think the point is, is that I think that the questions are different and the methodologies are different. In terms of writer training and quality and consistency across centers, yeah. I mean, I think it's -- you know, you want to try to do that.

Hopefully everybody does -- like in Rochester, they had a big tai chi effort, didn't they, way back when? You want to make sure everybody does tai chi the same way.

What interests me, and I've mentioned this before -- and in emails as well, which maybe I shouldn't have -- combining these drug studies with alternative therapies like yoga or physical therapy maneuvers, because it does interest me whether -- because we don't control for that.

One of the things we haven't talked about

Page 228

for a lot of this is we're not controlling for activity. And we say don't change your activity level, but I would say everybody who gets into a drug study should probably -- particularly if they have a severe pain, you may want to add some physical therapy modality and just make that constant through the study, so everybody's getting that, so you control for that.

I think it's really problematic when you're just studying a drug, and then you just leave it up to whomever, whatever they want to do, in terms of what -- if I were in a study, all of sudden I'd be like, "Oh, I want to get healthy," so I should start doing healthy things. I'm going to change the way I eat. I'm going to go to sleep at a better time.

There are all these other things I might do because I'm now taking care of myself, because I've just enrolled in a Farrar study, and I really want to make sure that I do a good job. So I start doing these things that potentially would help me get better because I think that's the goal. But

Page 229

the goal isn't to get better. The goal is to make a controlled trial and keep things as much as possible the same.

MODERATOR: But wait a second. Let me push you a little bit on that. So the issue of blinding can be more challenging. The issue of alternative comparative treatments can be more challenging, but we still think about this.

But the safety issue, did I hear you say that you don't think safety is a concern for rehabilitation studies, or for physical therapy, or for surgery, or for acupuncture?

MALE SPEAKER: I think they're different -- I think the regulatory issues are much different. The safety -- the amount of hoops we jump through in the pharmaceutical industry to follow the safety of a drug is much more significant.

You won't get lab values, I don't think, for most rehabilitative processes. You won't get chest X-rays or -- you know, there's a lot of really invasive stuff we do. I don't know why you would

Page 230

get an ECG on somebody if you're looking for long-QT syndrome and things like that.

MODERATOR: Is Ann Costello here still -- can't see her -- who works with devices? I'd be interested in her perspective of whether safety is an issue for devices within the --

MALE SPEAKER: Well, they are. I mean devices are a little different though.

MODERATOR: Okay. But we did some studies early on for self-disclosure in a rehabilitation program, and one of our people fell off the exercycle and twisted his ankle. Is that a safety issue I should have been concerned about? Was it reported then?

MALE SPEAKER: No, absolutely. I'm not saying that there aren't any safety issues, but the way we approach safety and we assess safety is different in a drug trial than it would be in one of these other trials.

MODERATOR: Is the rigor different?

MALE SPEAKER: Yes. I mean, you're doing much more surveillance of safety issues.

Page 231

MODERATOR: I disagree, but okay, that's another discussion.

MALE SPEAKER: Unless you're getting chest X-rays and CT scans and lab work every month or something, looking for particular things -- maybe I'm misunderstanding what you're saying. I might be being too literal. I don't know.

MODERATOR: Okay. Well, I'm going to close on this one and take my prerogative and turn it off.

(Laughter.)

MODERATOR: We're doing a geriatric study, and we're putting people into some type of exercise program, am I going to be concerned about the potential safety issues with whether people should be doing tai chi? If they have limitations in their ability to walk, do you have to modify the tai chi? Do I have to record that? So I obviously need to pay attention to that.

MALE SPEAKER: Well, I'm not saying you don't pay attention; I'm just saying that the safety issues are different. You still need to

Page 232

follow safety. It's just that they're different.

MALE SPEAKER: So I just have to interrupt and say that we focused this meeting on quality of efficacy outcomes, and we obviously haven't had any presentations, haven't had any discussion about quality of safety outcomes. And so that's really off the table.

Though Ian, thank you for an idea for another IMMPACT meeting, which would be just like this one, but addressing safety outcomes and trying to resolve the difference of opinion between David and Dennis.

MODERATOR: Okay. And I think, Michael Rowbotham, I think I saw your hand.

DR. ROWBOTHAM: This may seem a little off topic now, but I think it's a good segue to the next section. I'm going to give a talk, an informal talk, at the end of the month, and I've decided to entitle it about the -- what I learned here, and it's the four "F" words, which are fraud, fabrication, failure, and futility.

So as we go through the discussion, I'd

Page 233

really like to get an idea, from the panel and also everybody, really how big an issue -- especially when you start putting them all together, are these different things we've talked about: people not taking their meds on time, people lying to get into studies, study sites fabricating subjects or fraudulently double enrolling them, and all these other things.

When we put it all together, does that mean that many of our trials are doomed to failure or futility because there's so much data and so many subjects you just have to exclude as not being usable for data analysis?

MODERATOR: I'm tempted, but I'm not going to do this. You're going to give this talk in a month, we'd be more than happy to have you practice.

(Laughter.)

MODERATOR: Obviously, you're waiting to get more data to do it. Does anybody want to comment on Mike's comment?

MALE SPEAKER: I would, because this is

Page 234

something I think about a lot. And I was thinking about this across -- I've done a lot of clinical trials in industry over the years, and I would say that on average, I see about one or two sites per study where there are real concerns. And these are large studies; they're large phase 3 studies.

So I do think there is an issue there. I don't think it's an overwhelming issue, but I do think it's an issue. And what I see at those sites are not necessarily fraud; sometimes it's absurdity, like people who keep their medical records in the basement of their house. You're supposed to hold on to all of this stuff for like whatever the rules are until this drug's approved or two years after or whatever.

The point is, I've seen a lot of -- I've seen not a lot of -- I've seen that happen. I haven't seen -- I don't always know when the data's being fabricated. That's the problem.

There was a study in India where we used Topamax as one of our drugs for migraine, which it is approved for migraine, and nobody responded.

Page 235

Nobody responded at placebo; nobody responded to Topamax. And it was like a profound thing.

So when nobody responds to a drug that you know that it works -- and they had like 12, 14 people, is that fraud? I don't even know. I still can't figure that out. So I don't think these are easy questions to answer.

MODERATOR: Lee?

DR. SIMON: So I think that the problem is pervasive, but it is not ubiquitous. So I would bet you that every trial, no matter who's running it, where it's being run, will have some problem. It probably isn't consistently fraud, but I do have to just share one unbelievable story with everybody in the room.

So there was a new formulation of methotrexate that was being developed by a guy who spent 22 years in Louisiana at the university doing this work, extraordinary idea, physical chemical property difference, and he got his buddy in Peru to actually do a clinical study.

Adequate and well-controlled by design.

Page 236

They did this trial in Peru, and they had -- we use in rheumatoid arthritis something called the ACR20 as the primary outcome. And typically, you get 60, 40, and 20. Sixty percent of the patients will have a 20 percent ACR20 response with a drug that works. Ninety-eight percent of the patients had an ACR20.

The way this trial was designed, the patient came in, got picked up, they actually included them in the trial, they got the drug, and they disappeared for 12 weeks. And then they came back at the 12-week mark to get analyzed.

Basically, I did due diligence on this product for a company. Despite what I told them, the company said, "We're going to buy this." They bought the product. They actually started to study it, and it was totally non-bioavailable. You would take the drug orally, and there was no drug in the body, and yet they had a 98 percent ACR20 response.

I don't actually happen to believe that it's fraud. I actually happen to believe they probably took whatever they took after they left the clinic,

Page 237

and they were never seen again until the 12-week mark, and may have had a response.

If you do a trial in India, if you do a trial in China, if you do a trial in some other places, this is the kind of data that you get. Every autoimmune disease trial in China, every patient in China takes thunder god vine, off the shelf from their, you know, naturalist physician, despite the fact that they're told not to.

So I think that every trial has a problem, and every system has a problem, but I don't believe it's probably fraud. There'll always be something you'll find. The more we can improve the quality and the more we can improve the PI's behavior associated with what's required, the more likely it is that it'll get tighter and tighter and tighter.

MODERATOR: Laurie, did I see your hand up?

MS. BURKE: Right. I'm responding to what I heard on the panel before Lee's comment, and that is a reason -- I want to support the idea of training and certification and some sort of initiative that this group could lead in terms of

Page 238

good practice in clinical trials and teaching clinicians and investigators and whoever wants to be participating, because there is a big huge movement to train people in another type of science, and that's what sometimes is called comparative effectiveness research. It's sometimes called pharmacoeconomics.

It's the real-world ideas of not really looking for a treatment effect, but looking for the effect of a conglomerate of issues in different types of environments, in the clinical environments. And I think that this has really taken hold and is being talked about a lot because of the billions of dollars of the Cory funding, because of HTA and AMCP being convinced that they need real-world, non-clinical trial data.

So in terms of weighing the amount of information out there for well done, randomized, controlled trials, with the amount of other types of information that are now being generated because of all this interest and money being poured into it, I worry that there's going to become more

Page 239

disregard -- less regard for what we're trying to promote here.

So the idea of getting it discussed out there in the public, producing these materials to explain why rigor is important, why these ideas that we're talking about are important to think about, are critical. So I just wanted to make that point before the end of the day.

MODERATOR: Before I ask the panel to respond, Mike Rowbotham, I know you've done a lot of thinking about comparative effectiveness research or effectiveness research. Do you have any response to what Laurie was raising a concern about that type of research?

DR. ROWBOTHAM: I think the hardest part is to get it funded, because it's just not -- it's not the new-new. It's not going to lead to a regulatory approval or anything else. But really, it's so important, we should be looking at everything that we do in clinical medicine for whether or not it's better than something that might be less expensive or easier or safer.

Page 240

So I don't know how much we're going to try and put comparative effectiveness research into this particular document, but it really gets to the kind of things that you do with pragmatic trials, where some of the issues that we've talked about actually go away because the patients are treated within their usual practice setting, and then it's embedded within the electronic health record.

So a couple of the biggest issues we've talked about in terms of potential fraud or fabrication are people exaggerating to get into a trial that just disappears because they're already in it, and they may not even know it, because it's just really how their care is delivered. And that's one way of doing comparative effectiveness research, almost a little bit in the background, is through pragmatic trials.

MODERATOR: Thank you. Anybody? John?

MALE SPEAKER: Laurie, I understand where you're coming from, from the perspective of a drug approval, but I agree completely with Mike Rowbotham that comparative effectiveness trials are

Page 241

actually very important in terms of understanding how drugs end up being used in an environment and understanding the interaction of complex medication combinations and complex environmental medication interactions.

But I think the point to be made is that we're talking about clinical trials, and I think we're talking predominantly about phase 3 clinical trials. So we need to keep focused on the question we're trying to answer.

Clinical effectiveness research is trying to answer a different question and has a whole different set of issues. And so if we --

MS. BURKE: And that's exactly my point. And I'm sorry if I wasn't clear. But yes, there's a place for both, but they're very different.

MALE SPEAKER: Yes, no question.

MS. BURKE: And I think that we hear people say they devalue clinical trials because they want to see the real-world pragmatic stuff. And there's not enough information about the value of clinical trials and why, in fact, you have to have a

Page 242

randomized, controlled trial that's not real world, that's not pragmatic, in order to be able to detect a treatment effect.

MALE SPEAKER: Yes. I think there isn't a person in this room that would disagree with the concept that we need clinical trials in order to demonstrate effects. But I guess where I was going is that I think that the answers to some of the questions that we're going to try to put forward in this paper are going to be different, in different circumstances, in different situations. And how broad to make this and how restricted to make it, I'm not sure. But it seemed to me that the purpose of the meeting was really to focus on sort of large phase 3 trials where we're really trying to demonstrate efficacy.

Just to sound like a broken record, I think we need to include in that the ability to do them efficiently, effectively, and validly, looking for issues, but keeping in mind something, which is that we all know somebody who has something, who has something that's very rare. And so we've heard

Page 243

a lot of stories about fraud and about the professional patient and a bunch of other things, and I think it behooves us to look at the relative risk of those problems to the cost of trying to detect them.

I'm all in favor of a little bit of monitoring going a very long way. So I'm very much in favor of doing a lot of the things that we've just talked about over the last couple of days, but to be a little bit careful about over-emphasizing some of the stories that we hear.

MODERATOR: I think we are going to narrow this down, so it's probably not going to be broad to cover both of these. It doesn't mean we won't have a sentence or a small paragraph saying that some of these principles are relevant and they've been modified, but that's not the purpose of the paper.

So I don't think we're going to go -- unless Bob Dworkin tells me otherwise, I don't think we're going to focus on comparative effectiveness, but it's not to say sweep it under the rug and say it

Page 244

doesn't exist.

MALE SPEAKER: I thought one of the most provocative comments in the last two days was Bernard's comment regarding medication adherence when you said, we think we're doing an efficacy trial, but we're really doing something between efficacy and effectiveness. I thought that was just beautifully framed.

To Laurie's point, I think it might just be important to say that we think these types of -- answering the questions in these ways are different. They give us different answers for different methodologies to get there. And the reason to get this right is because we want to be sure that we're asking the question in this world, not in the comparative effectiveness world. So I thought that was what was particularly compelling about how you framed that.

DR. VRIJENS: I think the answer to that in the future -- because the problem today, we try to navigate between the two. We try to be selective, to show something, to show a difference against

Page 245

placebo, but we still want to be somewhat representative of real world. So we try to navigate in between the two, and we don't answer either of the questions.

I think the future is, at the beginning, we will need to be even more selective and even better train the center, and do everything even better to show efficacy as soon as possible with a limited number of patients.

Then we need to go to the more broader population, and there will be other -- in the effectiveness measurement. And we have to measure the sources of viability because we need to understand when it doesn't work, why it doesn't work.

MALE SPEAKER: Exactly.

DR. VRIJENS: But we need I think to answer the two questions a bit separately. And I don't know if it still fits the phase 1, phase 2, phase 3 design that we are used to. But I think that's the way to go in the future.

I was at the AU Union, and there was a

Page 246

president of the [indiscernible]. This is an association of GPs of any AUs from the U.S. It was striking. He said, "I have not a single patient who fits a clinical trial. I don't trust clinical trials at all."

Those messages become very dangerous because they don't want trials at all. They don't trust them at all anymore. And I think navigating in between makes us -- they don't trust us anymore. So I think we need to be more selective at the beginning and more broad at the end, and answer both questions.

MODERATOR: I think we're reached our first consensus. I saw every head nodding in agreement, so I put that into the paper.

MALE SPEAKER: I just want to follow that just by saying that working clinically in the operating room, we have problem rounds every Friday morning, and we talk about the horrendomas that happened all week. And it's really all we think about. We don't think about there were a couple of quiet nights

Page 247

We talk about fabrication and fraud, and that's really scary. And as Lee said, I don't know how -- it's not going to stop, but we don't know how common of a problem it is.

So I'm just wondering like, for example, whether we should change the title from "Ensuring" to "Improving" and make sure that we're not giving the impression that there's a crisis here in data quality.

MODERATOR: Bob?

BOB: Yes. So I want to take issue with that. You know I haven't done these kinds of studies, but I presented yesterday three very different groups that have been looking in a very focused way on duplicate patients: that Rabinowitz' IMI EU Israeli initiative; Mitchell Efros, who's based in Long Island, New York; and the guy in Southern California, Shaevitz. And they've all come up pretty much the same estimate of 8 to 10 percent of the patients in CNS and kind of symptomatic trials are duplicates.

That, to me, is a huge number, because if

Page 248

you then add to that 10 percent of patients who are participating in the same clinical trial in Los Angeles and San Diego, all the stuff that Eric has found, to me it suggests -- and I know, as has been said, it's terrifying -- there could be 30 percent of the patients in a trial are doing something seriously funky.

So I'm really struck by the Rabinowitz, the Efros, and the Shaevitz coming up with the same 10 percent duplicate. And I want to say one other thing. By the way, that 10 percent was exactly the figure in the IOM report for that schizophrenia trial with 300 patients. And when the FDA investigated, 30 of them were duplicate.

So to me, that's not a kind of minor problem. That seems to me like a huge canary.

MALE SPEAKER: Inasmuch as that's a problem -- well, I mean I don't want to minimize that. But is it systematically biasing results in a particular direction?

BOB: Well, it's hard to imagine how duplicate patients, if they're intact at all,

Page 249

aren't taking twice as much medication and probably aren't taking any medication at all, it's hard to imagine how that'll give you a false positive. But it's awfully easy to imagine that it's responsible for false negative results.

In fact, there was the analysis that Rabinowitz did showing that when he removed post hoc, 10 duplicate patients from a schizophrenia trial, the p-value, for what it's worth, went from 0.08 to 0.03.

MALE SPEAKER: I want to make one quick point, which is that it wasn't that the study went from being positive to negative. It went from being statistically significant to not statistically significant. I would be willing to bet that the effect size was altered a little, but it didn't change direction.

BOB: But if that was a phase 3 schizophrenia trial and the missing data were handled in the way that the FDA requires, in one case the trial doesn't get the drug on the market, and in the other case it does.

Page 250

MALE SPEAKER: No, no, no. I'm not arguing that issue. What I'm saying is that we talk about them as negative trials. And the best way to present a trial is that the trial shows an effect, but that it did not reach statistical significance. Because to talk about it as a negative trial suggests that it showed no effect, and it doesn't. It's a pet peeve of mine.

BOB: I'm using it as shorthand for the FDA's not --

MALE SPEAKER: I understand.

BOB: -- going to consider it, in most cases, as evidence that counts towards approval.

MALE SPEAKER: The point I'd like to make is that I think that a little bit of monitoring would go a long way to avoiding duplicate patients. It makes a great deal of sense. I would argue, though, that our target is not zero. Our target is 5 percent or our target is something; that if we said that this is a huge problem and we need to focus on getting it to zero, that we're going to over expend resources on trying to do that.

Page 251

So I would argue that it's well worth looking at; that there ought to be some way to do that within HIPAA regulations that would allow us to compare. And that a little bit of monitoring that way would go a very long way to reducing the number.

It's never going to get to zero, and so we need to just be cognizant of the fact that we need to focus on looking at each of these problems, trying to make them smaller, and being efficient about them and not being onerous in terms of the regulations and other things that we impose that would make clinical trials harder.

MODERATOR: Ajay?

DR. WASAN: One thing to add about this pragmatic versus efficacy trial issue is that I know we've all been keeping in mind phase 3 clinical trials with our comments, but I would say almost every single issue we've talked about actually applies to large pragmatic trials as well, and the thinking about them and how you would design them.

Page 252

So I really see our comments being kind of broadly applicable, not necessarily just the phase 3 trials. Plus, I think the vast majority of sort of published clinical research in pain medicine are blended trials that have both aspects of effectiveness and efficacy in there, if not for the only reason of adherence, for instance.

So I think that's part of what we want to come across, too, is saying that these are -- to get to your sense of trying to be as general as possible, but mention some specifics, I think that would be important, too, for what we're coming at. We're not just giving this narrow lens. So I'd like to see that go forward as a group with that.

MODERATOR: David?

MALE SPEAKER: Yes, I just want to kind of pick up on what Bob has said is the general theme of the conference which is, I do think a lot of this is really terrifying. I just don't want to minimize it. I want to make sure some of my comments aren't out of hand. I mean I've been dealing a lot with Parkinson's disease lately, and

Page 253

Alzheimer's, and fraud in those is probably different than fraud in pain.

But I do think -- I am a very big advocate of -- and Mitchell Efros has reached out to me many times. And in many bid defenses and many proposals, I've suggested using that in the past, was to try to find duplicate patients.

I think what Bernard came up with is probably the most terrifying thing that I've heard yet, and I don't even -- I'm almost stunned and speechless by that presentation. And I think that it really behooves all of us, particularly those of us that are doing a lot of clinical trials that are for pivotal studies, where there's really a lot of dollars on the table, to think about whether we can really go home and be happy knowing what we've heard and not be really, really disturbed by what Bernard said.

I think that all of these things together makes me wonder -- in certain areas wonder like why is there placebo creep? And one of my favorite areas is migraine, and there's this placebo creep

Page 254

in migraine. And I always thought, well, migraine is this great area, because of course they're always recycling these patients.

I guess Nat left. But the point, the question is it ever right to have a cadre of professional patients? What I would say, in migraine maybe it makes sense, because they've certainly -- they know what it is to take a triptan, and these studies go very quickly, and I think they're pretty accurate.

But it could explain why there's this placebo drift, and our drugs seem to be less effective for migraine, like if you look at a Maxwell study from 20 years ago, and now it seems to be a little bit less effective.

Bob, you were going to say something?

BOB: No. what you said there makes me think one thing we've left out for the last two days is thinking about this from the perspective of the patient in the clinical trial who's actually a partner and is doing his or her best to provide high-quality, valid data.

Page 255

If a clinical trial subject like that was at this meeting, they would be very upset, because they're putting themselves at risk of side effects, of getting placebo. And what we've spent two days talking about is all these things that are working to make the data uninformative.

So I think we need to somehow get into the article not only how terrifying this is to us, as the people doing the trials, but how terrifying this should be to the patients who are participating, who are the straight-shooting patients and trying to make a contribution, but there are all these forces that are working against them.

MALE SPEAKER: But I mean it's an ethical issue, I think.

MALE SPEAKER: Exactly.

MALE SPEAKER: They go into clinical trials with the idea that they're going to suffer, they're going to have inconveniences, maybe have bad side effects, but with the idea that they're results are going to have meaning and are interpretable. When

Page 256

we have all of this going on, it impedes that. And if we don't address it, there is an ethical dimension to it in my mind.

MODERATOR: There's going to be two more questions. I'm only going to take two more because we've got a lot of things we want to cover. But it seems to me like, number one, I would imagine that Penney Cowan got extremely excited when she heard the comment from Bob about, "Gee, we really need to look at the patient's perspective on this," or the person who has the problem and since you might want to use it.

Also the lady in Palatka on her website, I don't know if she at all talks about that perspective that might be of interest for us, those that haven't -- I'm not familiar with it to go look at it and see what she has to say.

There are going to be two last questions. One, Eric Devine's had his hand up, and then Mark Jensen. Then I'm going to cut it off at this point. If you have additional questions, hopefully we're still going to have a lengthy discussion and

Page 257

maybe they'll fit right in there. So Eric, you're first in line.

DR. DEVINE: Oh, thank you. So despite the numbers that I saw in my study with a high level of deception and fabrication, and the numbers that Bob is referencing, I don't have the perception that this is a crisis across lots of phases of research, because I think it has to do with the vulnerability of the study.

Studies with criterion that are diseases that are assessed by subjective assessment versus objective, like Amy said earlier, in an oncology trial where there's no reimbursement and people already have free access to healthcare, the chance of people gaming for some sort of study enrollment is very low. And while there could duplicate entry, because people are desperate for care, that's a little bit different than the population that I'm noting.

So when you think about how do you allocate resources to combat this problem, you really have to look at the vulnerability of the study. Is it

Page 258

paying money? Is it a condition for which subjects can fake their way, and we know that they can. Do they have access to it through clinical trials? Is it something that's a network where they can go from site to site?

If you have a narcotic pain relief that's part of being in the study, that bumps up the vulnerability, because the street value is just too tempting. So you get reimbursement plus a little recreational drug use, and maybe some money on the side from selling what you don't use. So vulnerability is what we need to --

MODERATOR: So obviously, some balance. Yes, Mark? Last word on this section.

DR. JENSEN: So I'll just respond a little bit to what David said about it, the issue of adherence that Bernard mentioned. We've talked about training patients for better assessment, and I don't know that it was specifically talked about whether we should include in the paper training for adherence, that we would want to include issues about that. We always do that in our psychosocial

Page 259

events, is we use MI to train for adherence, but I don't know how many drug trials do that now.

MALE SPEAKER: The only thing --

MODERATOR: Last word.

MALE SPEAKER: Okay. The issue about adherence, I think we need to be a little bit circumspect from the perspective that I'm convinced that pain trials have different adherence issues than an Alzheimer's or a blood pressure trial, or things where patients are not symptomatic. And it's somewhat telling that they're -- unless Bernard knows of a larger population of studies. But there are very few studies that look at adherence and anything related to pain. Maybe those should be done.

I'm not at all suggesting it shouldn't be mentioned in the paper. I'm simply saying that I think we should be cognizant of the fact that, at least in my patient population, the issue is not taking too few rescue drugs, it's taking too many. And so the problem of taking their drugs on a regular basis is not as much an issue.

Page 260

One could argue that in drugs where we're giving them and they don't see a dramatic effect immediately -- pregabalin would be an example -- I certainly have patients who come back and say, "I took two pills. It didn't help" and I have to educate them on taking them regularly. So I would agree with teaching in that area.

MODERATOR: I think you better be careful about saying that there's no studies on adherence in pain. Just as a crude area, in the opioid area, there have been a number of studies that looked at urine tests on people to people who are supposedly being prescribed opioids, and by far, many more underuse the medication than overuse the medication.

So it's not as it, number one, we don't do this, and number two, there are plenty of people who underutilize. So let's end this session.

Now the fun begins. This is the part you've all been waiting for. The exam. This is the exam. This is the did you earn your provisional qualification? Dr. Dworkin, do you have the page?

Page 261

Thank you, gentlemen.

Consensus Discussion

DR. DWORKIN: Okay. So you're in the home stretch. There is no formal coffee break this afternoon, so it won't hurt my feelings if you wander out. We have a very hard stop at 4:00, because there are taxi arrangements and everything else.

I only have a couple of slides. And really, for the next hour and a half, or however much time is left before 4:00, or we might finish sooner, it's going to be discussion and argument.

I thought a good place to start was with the two definitions of quality that Nat put up yesterday morning; the one from the FDA presentation and his own. This is more or less what we've been talking about. And as you all know, these two days really are to provide the raw material for an article with recommendations. Or if we don't have recommendations, because there's not a lot of evidence, there'll be considerations or recommended considerations. We've used all of

Page 262

that language in the past.

Dennis, I thought made an important point about this effort, these recommendations, is that to the extent possible, we want them to be as applicable to a clinical trial of yoga or acupuncture or cognitive behavior therapy or hypnosis as they are to drugs, and that's always been our hope with IMMPACT articles, that the recommendations are kind of generally promiscuously applicable to clinical trials.

The other thing -- and that came out this morning -- is we'd like the recommendations to be as relevant as possible in non-regulatory settings, you know, academic settings, foundation clinical trials, et cetera, Bill and Melinda Gates studying treatments for malaria in Africa.

To make that point, I wanted to show you a photograph I took of -- this is the outside of John Markman's pain clinic. That's Meg Ryan who was his very first study coordinator, showing up for work one day. And you can see this is a very small operation, and they're very sophisticated, they're

Page 263

very knowledgeable about clinical trials.

This is David Hewitt's clinical trial shop. This is obviously on a much larger scale. And if you look really hard, you can see Tom Hanks showing up for work. Tom Hanks was one of David's early study coordinators. But so you all know, this is from You've Got Mail.

But I think this is important for us to remember as we go to the next slide, which is really the important slide. We would like these recommendations to be as relevant to the small boutique, clinical trial, academic operations as they are to the much larger situations.

So what we've tried to do on this slide is to kind of summarize what a few of us thought were the high points of the last two days. And so this is really my last slide, except for one more that is at least an initial summary, an initial kind of scaffolding of the article that will be drafted.

So it seems that we've really been talking been talking about sources of discordance, discrepancy, mismatches, between what a clinical

Page 264

trial protocol is intended to do, the study the investigators intended, the objectives they had mind, and the way the study gets executed that can adversely affect the quality of efficacy data.

As I said a moment ago, we really haven't focused at all on safety. That's a whole other meeting. And so what are these kind of discordances, discrepancies, between the intent of the protocol and what actually happened when the rubber met the road?

This is really the summary of those sources that seems to have come out of the last two days. There are patient sources, site sources, so characteristics of the patients, whether they have the disease that the clinical trial is studying; has there been some exaggeration of their symptoms so they can get randomized; are they hiding, as some did in Eric's study, treatments from the investigator?

There are sources of discordance involving outcome reporting. This is, of course, Mark's presentation yesterday morning, the intentional

Page 265

unblinding that I talked a little bit about. Of course what we were all very, very troubled by, the lack of medication adherence that Bernard talked about, and then a set of site characteristics; I'm not going to go through them in detail.

One of the themes it seems over the last two days has been what can we do to prevent these discordances. Quality by design is obviously an approach to this. You build in, as much as possible, safeguards into the protocol, but of course nothing is perfect. And what can we do to identify these mismatches between the intention of the protocol and the study execution, as they're occurring.

Then -- this is something we've danced around a lot about, and I don't know that we've got a whole lot of answers here -- once you've identified something funky, what do you do? What can you do legitimately, to address it in the middle of a trial; and if you can't, afterwards in the analysis?

So when I look at this slide -- and forgive

Page 266

my inability to work with Excel -- it really should be a 9 by 4 grid that we would be filling in. So there are kind of 36 cells on this slide that you can't see, but hopefully you can imagine, which are these four ways of addressing discordance across these nine aspects of patient and site arenas, domains for discordance. Bob?

BOB: Immediate reaction, I think this is actually terrific. Two things that come to mind: one is another dimension, which was in the definition of quality, may be possible to both address, I guess, the importance of these factors in the integrity of the study and the ability to produce reliable results and protection of human subjects.

DR. DWORKIN: Yes.

BOB: So bringing both those. And then the only other thing I would say is I wonder -- I'm very interested in the medication adherence thing we had a presentation on. I don't think we really focused much discussion on that. I think Dennis just mentioned it. I think there's a lot more that

Page 267

we didn't talk about, so I'm not sure how that will be brought to bear here.

DR. DWORKIN: So four cells -- so in a row obviously -- Bernard had to leave early. Obviously, we will rely on Bernard to fill in the four cells here: prevention of medication mis-adherence.

Identification, he talked a lot about that yesterday in terms of the electronic approaches to identifying medication adherence; and then of course the issue of --

One question I talked about at the break with someone is if Bernard's electronic system says the patient stopped taking their medication a week ago, is it appropriate for a coordinator at the site to call the patient and say, "Mr. Smith, we noticed that you stopped taking your medication a week ago; what's going on?" Or is that that kind of midcourse correction not appropriate?

I know what I think, but I'm not sure we could get an answer today.

Then, of course, what Bernard talked a lot

Page 268

about is after the fact when you've got the adherence data, you can do a secondary post hoc analysis to look at the relationship between whether the patient was taking their medication and efficacy and et cetera. So that's the 36 cells that aren't on this slide.

BOB: One other minor -- well, maybe not so minor -- I think we did focus on unintentional -- or excuse me -- intentional unblinding. That really isn't unblinding, and including unintentional --

DR. DWORKIN: Yes.

BOB: There are things at the level of prevention, et cetera, related to -- well, prevention of --

DR. DWORKIN: You know, I agree, Bob. I'm not sure --

BOB: -- not intentional.

MODERATOR: I'm not sure why I made this intentional, because as one of the speakers -- I forget who mentioned -- assessing whether patients became unblinded from side effects is a very

Page 269

reasonable thing to do. Consider the word "intentional" withdrawn because it's very reasonable to ask patients at the end of the trial, Which group do you think you were randomized to?

Just because some of you may not be able to see it, at the bottom, I tried to emphasize that this all applies to eligibility criteria, efficacy outcome data, adherence data, follow-up, and subject disposition, but not adverse events. David?

DAVID: One of the things we didn't mention, I don't think, was the idea of overdose, and that's kind of the opposite, right, of this adherence issue. But certainly, it speaks to quality issues. And how we give out drugs, whether we use blister packs or bottles, we didn't get into that, I don't think. I think that becomes very impactful as well. In addition to missing doses, is overdose.

Of course, with that is always the concern, particularly for some of our drugs, that there may be diversion of these drugs as well if there's a concern that there is a positive reinforcing effect

Page 270

of the drug and sharing it outside of the confines of the study.

DR. DWORKIN: So you see overdose an adverse event?

DAVID: It can be an adverse event.

MALE SPEAKER: So shouldn't that be off the table for this if we're focusing on efficacy outcomes?

MALE SPEAKER: It depends what you're attempt on the trial is, what's your -- the importance of the trial. There are trials that have overdose as a primary outcome.

DR. DWORKIN: Absolutely. Right.

DAVID: Well, it's an adherence issue, I think, because you're taking too much. Not all overdoses are assigned the designation of an adverse event. If there's no adverse event associated with it, it's not an adverse event. It's not adherence.

DR. DWORKIN: Right. Mark?

DR. JENSEN: Just another complication, assuming that this table will be filled with

Page 271

recommendations, there's going to be different levels of evidence, the strength of evidence before those recommendations, so we need to incorporate in the table or in the text how strongly we think you ought to be doing the thing that we're saying.

So I think there's a consensus, no paper diaries to assess adherence, but there are probably other ways to assess adherence that we think are pretty good. But in terms of training, for example, we don't have the evidence yet. We think this would be a good thing to do, but we don't have -- so just some way of indicating the level of evidence in the table or in the paper.

DR. DWORKIN: Yeah, we're going to need to do that. I actually think that we're probably going to end up -- my guess based on previous impact articles is that we're going to end up calling these "considerations" rather than "recommendations" because there is no evidence, so we can't really have evidence-based recommendations.

I think even for electronic versus paper, my

Page 272

sense is we all agree that at this point in time, electronic is preferable, but it would be really hard-put to cite something showing that you get better quality data from an electronic device than from a paper diary. I don't know. What would I cite?

MALE SPEAKER: I would cite that the paper diaries we know are bad. We know they're bad. The electronic diaries, we don't know that they're bad.

DR. DWORKIN: So what would you cite to show that paper diaries are bad?

MALE SPEAKER: Oh, there's plenty of evidence, observations of people doing the hood, the lack of consistency in their responding. We have papers. I can find them for you that have the evidence.

Clearly, we'll have some qualifiers saying these are recommendations. But I think there may be some recommendations we feel more strongly about than others, and I think we should just make that clear. That's my point.

DR. DWORKIN: You know, I think we should

Page 273

have strength of recommendation or considerations, strength of evidence. And those could be -- we could have a strong recommendation with weak evidence because it's just common sense. And any reasonable, thoughtful person would agree with it, even though we can't cite chapter and verse of randomized trial.

I think that's brilliant, kind of strength of evidence, and it's often going to be not very much at all. But we can also give the strength of our recommendations. Yeah, Mike?

MIKE: I just have a question about this intentional unblinding aspect. If it's really intentional on the part of the site to either have access to something that they're not really supposed to be looking at, that's really site misconduct. But I've always had concerns about explicitly asking subjects what treatment they think they got assigned to, partly because it helps unblind the staff who may not have thought very deeply about it, so in the process of querying the subject, the subject will say I felt this and I

Page 274

felt that; and when I put all these things together, it convinced me that I was actually receiving the experimental drug.

I would want to have some caution in this about explicitly asking subjects too many things about what treatment group they were assigned to because, cumulatively, if they're guessing more correctly than by chance, it will start to unblind the investigators in the studies.

DR. DWORKIN: I never thought about that, and I don't know that anyone else has in print. One way to address it would be to have the assessment being done at the end of the trial electronically with the study staff being blinded to it. The patient says in some kind of electronic capture did they think they were randomized to drug or placebo and why, and the site staff never see those responses.

MIKE: Can I just add a quick follow up? I'm sorry. There was something in the news some years ago where they were reporting on venture capital and other equity groups who were pretending

Page 275

to be patients and getting on to these kind of patient websites, and then inquiring about side effects and sort of going by what was maybe in clinicaltrials.gov or something else. Then getting to know electronically other subjects, and then asking them about efficacy, and trying to pick out the ones that they thought were really on the active drug and getting an early read as to whether or not they should buy this stock or dump the stock.

MALE SPEAKER: That's brilliant.

MALE SPEAKER: Oh, yeah. It makes a lot of sense.

DR. DWORKIN: If you will send us a reference -- Mike, I promise to include this in the article if you can send me a reference to that.

(Laughter.)

(Crosstalk.)

DR. DWORKIN: That is too titillating a tidbit to ignore. Ian?

DR. GILRON: Thanks to Mitchell Max, how we did this when I trained and learned how to do

Page 276

trials with him was to routinely ask people to do blinding questionnaires. We routinely, at the same time, asked the research staff to see what they thought.

John Markman showed obviously a very dedicated and insightful research staff who pay a lot of attention. You can't have it all ways. We want good quality and stuff, but they also -- they think about what they're doing, right?

Our experience has always been that the research staff has always been more unblinded than the patient. We do it at the end of the study. And to be honest, I can't say that we've protocolized the order in which it's done.

I think Mike's concern is that it's going to somehow affect allocation concealment. I'm not even sure what the problem would be. If you're asking at the end of treatment -- if you think somehow that the study subjects' response is unblinding the research staff for that particular patient, I'm not sure what the implications would be there.

Page 277

If they're thinking about the block randomization, it might affect their allocation concealment for its subsequent patients. But I'm not sure what the liability is for doing unblinding questionnaires.

MIKE: For example, there were some drugs in development that caused this distinct change in taste in the part of the subjects. So really, elaborate procedures were done to change the way the pills looked, or other things, to get it into clinical trials and not have inadvertent unblinding. And that's fine. You may want to test the adequacy. But it really is more the interaction between the subject and the site personnel.

So if you're doing it after the fact, completely separately from like an independent grader, completely separately from these active clinical staff, that's fine. I already found the reference -- at least one. There's lots of them from 2002 Wall Street Journal.

DR. DWORKIN: So Ian, I'm not sure I

Page 278

understood your question because isn't the risk that the study staff become unblinded because they start to -- and the expectations that they then have somehow unintentionally, nonverbally get communicated to patients if the staff over the first few patients learns that half the patients seem to have dizziness and half don't.

The patients are kind of saying, I think I was on drug because I was dizzy, but my pain also got better; that the staff develops an expectation that patients who report dizziness are going to get better, and that somehow augments the drug effect and decreases the placebo effect when there's no dizziness.

DR. GILRON: I understand that. But I mean, except for a phase 1 trial, every consent form is going to have AE information. I mean there's always a potential that patients are going to be unblinded and --

DR. DWORKIN: But I assume that the patients either don't read the consent form or forget it within 15 minutes of leaving the clinic. Please

Page 279

don't repeat that outside of this room.

DR. GILRON: It's in the transcript.

DR. DWORKIN: John?

JOHN: In thinking about this table, I think it's a great table in terms of the implementation. But we heard a number of presentations that talked about things and issues related to the design. Maybe you're going under that. But the other one --

DR. DWORKIN: [Indiscernible] my last slide.

JOHN: Okay.

DR. DWORKIN: So I sent Nat -- because Nat had to leave early, I sent him the previous table, and he said it's leaving something out, which is kind of my bunionectomy example of designing the trial to minimize -- and Amy, and Nat, and I tried to come up with a term, and the best we came up with was minimize experimental noise in the way the trial becomes conducted.

We thought about whether the word is "covariates," so see if this is what you were mentioning. This is Nat's slide really, not my

Page 280

slide.

In the bunionectomy example that he presented yesterday, remember when, I guess it was Scirex, first started doing that as a phase 2 design, they hadn't really learned -- and Rob, I think you were there -- the different factors in the procedure with the patient who was being assessed while they were lying down or sitting up. And over time, they standardized all those experimental procedures. and all of a sudden, assay sensitivity went up.

So Nat said what's left off the previous slide is this set of considerations about these sources of noise in a trial that should be addressed ideally in the design.

So is this what you were thinking of?

JOHN: Partially. But for example, it's alluded to even in the second statement there, which is the factors that affect the primary endpoint.

An issue that was brought up, I think, very nicely by Scott Evans in his comments on the panel

Page 281

was the issue of designing the trial to avoid missing data. I'm not sure how that fits here, but it's critical.

I like to say that if you want weekly data, measure it daily; if you want monthly data, measure it weekly; and if you want quarterly data, measure it monthly because that way, at least you get something that you can then average if you're missing a little bit.

If you design your outcome as a very complex multi-leveled questionnaire, you're going to get different kinds of answers. So I think that the issue there is -- I could see it fitting here, but I don't see it there. And I'm wondering how you'd see it there.

DR. DWORKIN: I was going to say, Shouldn't missing data be number 6 on -- well, either number 5 --

JOHN: I understand.

DR. DWORKIN: -- number 5 under patient. So wouldn't we consider having missing data one of these discordances between the intention of the

Page 282

protocol --

JOHN: Yes.

DR. DWORKIN: The intention of the protocol was that everybody gives you complete data, but the execution, of course, some patients drop out; there are missing data. I think missing data is omitted here, inadvertently, and it should be number 5 under patient. And we build into the trial everything we can to prevent missing data, but we also want to identify it, and then how do we deal with it?

JOHN: Correct. I don't know if this is exactly right, but you start off at the bottom, which I can only partially see, I think with eligibility. I think it actually backs up to the issue of project design. and I'm not sure how to include that exactly.

My point -- I mean, I think you get the point, which is that I think we need to think about the design issues that lead to improper data collection for a variety of reasons. These are what the patient does and what the site does, but

Page 283

some of that is sort of built in to this process.

DR. DWORKIN: Yes. Laurie and then Andrew.

MS. BURKE: I think that the bunionectomy example is part of the assessment; it's part of what you would do with this training to make the assessment in the assessment tool. Maybe outcome reporting is just part of that assessment.

I think it's combined in there. I think it should be more than reporting. The whole assessment process would take care of that missing piece, don't you think?

DR. DWORKIN: I mean, I agree, but I also think it's sort of assay sensitivity. I think about like third molar extraction, the type of extraction, as I understand -- I don't know much about it -- is associated with the assay sensitivity of the model.

It's something about outcome assessment and also the model. So I think we're going to have to struggle how this set of issues -- can it be incorporated into the previous slide or is it a kind of separate set of issues?

Page 284

Amy, and Nat, and I did struggle with it for about a half hour over lunch, and this was the best we could do, to put it on a separate slide, but we will be working with this. Trudy?

DR. VANHOVE: I was thinking could you not put it under site because, really, you would want the procedure to be similar between sites, the bunionectomy procedure.

DR. DWORKIN: Standardization of --

DR. VANHOVE: A standardization of the procedure or whatever it is that you're looking at.

DR. DWORKIN: And that's partly training but partly protocol. So that's right.

DR. VANHOVE: Partly protocol, exactly, yes.

DR. DWORKIN: Yes, it's partly protocol, partly training.

DR. VANHOVE: It's not a training issue, but it is an assessment issue as which size are you going to enroll.

MALE SPEAKER: It could be a separate section.

DR. VANHOVE: Selection?

Page 285

MALE SPEAKER: You have patients and site; you could also have protocol, or design, or something that would just fit therein, and then you could put the missing data and the other pieces right into that. It might be a way.

DR. VANHOVE: But it's a site selection issue.

My other comment would be, could we -- well, would it be possible to replace "mid-trial" with "during the trial"?

DR. DWORKIN: Yes, absolutely.

DR. VANHOVE: Okay.

DR. DWORKIN: I think we will have a whole lot of back and forth with our colleagues at FDA about this, what is appropriate, reasonable to do during a trial versus -- and I think Sharon was very clear about one thing.

One example of this yesterday, where Sharon said quite clearly that saying to a patient, during their participation in a trial, "Notice that you said your worst pain was less than your average pain. You need to think more clearly because that

Page 286

really isn't logical." Sharon said that's unacceptable.

So that raises the question, I think, in a lot of our minds, well, then what is acceptable? Is it acceptable to say to the patient you haven't taken your medication; you haven't completed any of your diaries? Obviously, we're going to have to get those issues ironed out after this meeting.

Andrew?

DR. RICE: Bob, I wonder if we can put one other factor under patient, and it's been the issue that I was aware about before I came, but it's been really emphasized. That's the issue of the healthcare setting or the country in which a patient is recruited.

So this concept of a professional patient is totally news to me. We've had a lot of interesting discussions about the fact that patients can earn money and income from participation in clinical trials. That just wouldn't happen in Europe. And therefore, the motives of the patients entering the trials are different. I suspect that's true of

Page 287

most of Northwestern Europe.

There's also a very interesting comment that was made over lunch that there's a very close relationship, generally in the UK and many other European countries, between patients and their doctors. And you'll often hear a form of bias creeping in where you suspect patients are giving more positive answers because they don't want to upset the doctor about his nice new drug.

I suspect the motives may be different also in other healthcare settings. I've done trials in the developing world, and that to do not financial gain but to gain access to healthcare.

I guess we've got two choices here. We can either talk about these professional patient issues and say this is just about -- I'm not sure it's just the USA or it includes Canada as North America. And these issues are pertinent to those settings, but there are very different issues with regards to other healthcare settings.

I think it's a really interesting research question just to try and document what incentives

Page 288

might be in different countries.

DR. DWORKIN: I agree that we need to have something in the manuscript that whatever we end up putting in this 36-cell table is very likely going to vary by region, country.

It might vary importantly by whether it's phase 2 or phase 3, whether it's a single-site academic study or a kind of multinational, multisite phase 3 protocol. So region would go in the category of moderators of these factors and what can be done about them.

DR. RICE: But I do think it's important to have a discussion about this professional patient and earning from a clinical trial issue because I think a lot of people in Europe just don't understand that. It's not a concept that they grasp very easily.

MALE SPEAKER: I would agree, Andrew, and I think it's not kind of up there because we did spend a lot of time. But under patient, you could call just "patient misconduct," and that would cover the duplicative patient.

Page 289

I'm a little surprised that you would think in today's world it'd be so easy to have software for the CROs or sponsor, but certainly the CROs. And it should be picked up at screening this patient has already been in a trial, and that would be a huge service to those guys, where it doesn't disrupt the trial or turn a significant trial into a non-statistically significant trial as you said.

You know what you could do? You could carry it one step further. Is it inappropriate? If a patient has done that, that's a willful act. That's sort of -- me, that's a one and done if it's a urine analysis. Certain urine analysis in my clinic, a little THC, I sort of forgive the patient, ask who their supplier was, and then move on or whatever.

(Laughter.)

MALE SPEAKER: But a duplicative patient in the same trial, that is a pretty much of a serious thing. The CROs could blackball that person from ever going into any clinical trial again, at least within that context. John is shaking his head no.

Page 290

I don't know. But it seems to me like you could penalize, to some extent, that patient. That's a pretty -- most of the people, I would think in this room, would not want to risk, if they could, that patient participating, subsequently.

DR. DWORKIN: I completely agree. I think the HIPAA issues and confidentiality issues have been resolved by Rabinowitz, and Efros, and Shaevitz . And we thought about inviting all three of them to this meeting. But since they're all doing this separately and presumably are competing with each other, we didn't want this meeting to turn into a kind of slug fest of who has a better online system for identifying duplicate patients.

But I agree, that all three of their approaches seem very straightforward, and it's hard to imagine a reason why you wouldn't implement it because these are people you don't want in the trial. Dave?

DAVE: Yeah, just a couple of things. One is, with all due respect to England and all the medicine being better there, which I'm sure that it

Page 291

is, until you do the study, you don't know whether there is, in some level, fraud going on there. The motivations could be completely different. I mean I think it's a study that was worth doing.

Certainly, it would have different motivations. I mean, it wouldn't be for money. But I'm not sure how much of it is all about the money in the United States either. I think that's one question. I think that's something to consider.

DR. DWORKIN: Dave is suggesting that ACTTION fund Eric Devine to go to London and redo the Boston study there.

(Laughter.)

DR. DWORKIN: And it looks like Eric is all for this idea.

(Laughter.)

MALE SPEAKER: I just wanted to find out if this is going to be part of the paper as well, is that when you address all these areas, one of the things you may do at the end -- and it will be interesting to opine on -- is whether you're going

Page 292

to increase the assay sensitivity of the study.

With that, what is the implications for moving forward in terms of our historical data and how we power studies. I mean, it might change things in a very fundamental way, and maybe we'll be able to get away with smaller numbers of patients to be able to do some of these studies. It's an interesting thought.

DR. DWORKIN: I think that's the hope. Raymond?

DR. CHEUNG: I notice with Nat's -- the slide that you showed -- Neil and I both have bitter experience with failed clinical trials, specifically in the post-op space where standardization of the procedure would affect the baseline pain if you didn't do that.

If you didn't standardize the procedure and the post-operative analgesic regimen -- and I guess Nat also pointed out some other factors like how you actually ask the patients their pain. I think maybe there could be a category of the condition of the procedure that could affect the results.

Page 293

DR. DWORKIN: Yeah. No, I think -- John suggested this, that there's probably a third block on this slide that is something about the model, the design, and that's where we'll put in things like missing data, standardizing the procedure, the assessment; partly, that's training as well as Laurie pointed out. So yes, we will add some third category here to address Nat's bunionectomy and related issues.

Neil?

DR. SINGLA: Yes, just one quick point regarding the site factors; there's five factors listed. This is just my opinion, but I think that it's more actionable right now to help sites get better quality by improving their processes, and that most investigators out there are not fraudulent, and that the FDA -- the whole construct of clinical trials that are being done for industry right now very much looks for fraud a lot instead of looking for true quality.

So if we're trying to improve the quality of clinical trials, it probably makes sense for us as

Page 294

a group to talk a lot about the last two and not so much about the first three because they're all the same in a way: fabrication, falsification. And if you make it like a police article, where it's all about how to police more, that's not really, in my opinion, what we need. We need to just be better at what we do.

(Applause.)

DR. DWORKIN: I agree, Neil. I originally had a big red box around these two, and we could put a red box around this also because that's where training is targeted. Training is targeted at the patient not doing a very good job of reporting their pain. And then, of course, training is targeted at the carelessness, poor training, recording errors, misunderstanding, incompetence down there.

So I think we actually, as you heard from the applause, we all agree with you that we need training; we need standardized training; we need evidence-based training, and that's hugely important.

Page 295

Lee?

DR. SIMON: I'm interested that no one has actually referred to a problem, which is ubiquitous in the orthopedic community in pain trials, particularly when you're using devices that have been invented by the person doing the study.

I've been involved in a couple of trials where the inventor of a drug was a study site, and the patients he recruited actually knew he was the inventor of the drug and wanted to make him happy, and all had a response, including those responding on placebo. Therefore, the studies failed.

This is a training issue. We should know that if we are invested in such an event, that we should not be the person carrying out the trial of studying that product. But nowhere up there is actually this been said. And because it's ubiquitous in the orthopedic community, it perhaps isn't something that people recognize because it's clearly not the right thing to do. But nobody keeps saying it.

DR. DWORKIN: Somewhere up here is kind of

Page 296

making sure we've done our best so the patients have realistic expectations, and this is something that's making --

MALE SPEAKER: Removing bias.

DR. DWORKIN: Exactly. S that's up here somewhere. We can put that in.

Laurie, you had your hand up.

MS. BURKE: I was just going to say that I think under site 1, 2 and 3 really belong under 5. They're like subsets of systematic error that need to be addressed. I don't know.

DR. DWORKIN: Yeah. We based some of this on publications in the literature, and so we'll go over that. That's right, they are systematic but they seem in another level of kind of illegal.

MS. BURKE: I'm reacting to the suggestion that we don't want to make this all about the fact that there's so much fraud in the clinical trial world. These are exceptions rather than the rules.

DR. DWORKIN: One thing we can obviously do is combine 1, 2, 3 and make it just one subsection, kind of, fraud, fabrication, misconduct rather than

Page 297

splitting them out. I just wanted to put down the definitions for fabrication, falsification that comes from, I think, the article by Biogen. But in the article, we will not make it look so lopsided as Neil and you both pointed out.

Other omissions, additions, et cetera? Bob?

BOB: Well, in the service of just maybe stating the obvious, the word "fidelity" isn't up there, and maybe it's similar to "quality." The basic premise of designing a study, developing a protocol, and then following it, and knowing that you've done what you said you were going to do in the service of producing replicable methods and results, I think that that really is a core principle of value -- or I mean of quality, excuse me.

So it comes to things like it's really about designing the trial and developing a protocol that's going to prevent problems as a really fundamental premise about this enterprise.

Then in terms of correction, identification, you want a protocol that's going to help you

Page 298

identify problems so that you can correct them. And then the correction, I think, it's important that it's transparent when that occurs and that there's thought about that.

Of course, it's important at a local site to be able to move clinically with good clinical practice in mind and the patients', human subject, protection in mind to act, to deviate from a protocol, if you will, on behalf of the patient care.

But having said that, I think that it's fundamentally important that these mid-trial corrections, if you will, are really carefully considered in the context of that overarching concern about the fidelity of the trial.

DR. DWORKIN: I think that -- right.

BOB: So it's maybe just restating --

DR. DWORKIN: I think the most challenging -- the most challenging column here, in some ways, is exactly this midstream correction, and it should be prespecified ideally, transparent. What we heard yesterday from Paul, from Sharon is

Page 299

it's got to be extensively, clearly documented. This is going to be the most challenging part of the manuscript to draft, I think, that column of 9 or 10 recommendations. Dave?

DAVE: Just to be clear, you're not going to include the company, the planning, the sponsor -- it's kind of interesting because it means that you're really putting all the onus on the patient and the site, and that you don't really think any of the risk to quality sits with the biopharmaceutical company.

DR. DWORKIN: No. Isn't that this? Have they designed the right study that prevents --

DAVE: Okay. I'm sorry.

DR. DWORKIN: Yeah. No. I think the company is -- actually the company is responsible for all of this because this should all be in the protocol, right?

DAVE: Okay.

MALE SPEAKER: Just to follow up to David's point, obviously, the data comes in from the patient. The site does something with it. Then it

Page 300

gets put into a database, then it gets programmed, then the statistical analysis, then the endpoint, the final results. We're not really going -- to David's point, we're not thinking we can correct those aspects of data quality and data process.

DR. DWORKIN: Right. So the first version of the slide had after the word "execution" in parentheses, it said "not analysis and interpretation."

MALE SPEAKER: Not analysis. Okay.

DR. DWORKIN: Because analysis and interpretation, I think, would be a whole other meeting. So this is really just about what happens before the database is locked and the statisticians take over.

Trudy?

DR. VANHOVE: Bob, where would you put like a bad medical monitor? Because you have a brilliant protocol, and the medical monitor, however, gets calls, and he lets in patients that really don't meet the eligibility criteria. But he's kind of like -- he's not very strict. It's

Page 301

just a bad medical monitor. Where would that go?
MALE SPEAKER: Carelessness.
DR. DWORKIN: Mike!
MALE SPEAKER: I was just going to say, the last two comments remind me that maybe it's patient site, but you also have the people who are responsible for overseeing the study. It isn't just designing it. It's overseeing the conduct. And I think that's what you're starting to hearing from people.
DR. VANHOVE: Yes.
DR. DWORKIN: So there's a third or a fourth, depending on what we do with design, category of oversight, absolutely. That's an omission.
MALE SPEAKER: Bob Kerns mentioned the fidelity with the protocol, which is essentially capturing that point. Bob Kerns talked about the fidelity of the protocol, as are people following the protocol. A bad monitor is not following the protocol.
DR. DWORKIN: No. You can have a rogue

Page 302

monitor who isn't doing his job, and that's not -- the job of the monitor isn't really specified in the protocol. It's specified I guess in SOPs of the CRO.
FEMALE SPEAKER: Oversight makes sense.
DR. DWORKIN: I think oversight makes a lot of sense. Yeah.
Lee?
DR. SIMON: So just to go back to your issue about this mid-trial column -- and Laurie and I are probably the only leftover people from former FDA as opposed to any FDA people here. It's really critical not to make anybody who reads this paper to believe that they have carte blanche to manipulate issues that come up or become evident in the mid-trial or ongoing review.
I can't tell you the numbers of times that I've actually had to see, on both sides of the table, where we see a data set that it suddenly dawns on them something is not right. And they don't understand that the trial then is obviated based on how much they do or what they do.

Page 303

So it really is -- we have to be really careful not to give people the sense that this kind of tweaking is de rigueur and welcomed. It may not be welcomed.
DR. DWORKIN: That's right. I could even imagine -- it's late in the day; we don't have to think about this -- that we could recommend, in certain settings, we as a group think it's perfectly reasonable to call up a patient and say, "Mr. Jones, it looks like you didn't fill out your diary yesterday or you didn't take your medication yesterday." And we might say that, but it's going to be followed by, "However, for registration trials, the regulatory agency needs to be kind of contacted to ensure that this is acceptable." We will stick that in after any recommendation where we think it could be problematic at either FDA or EMA.
DR. TURK: So that's like a black box warning?
DR. DWORKIN: that's our black box warning, exactly.

Page 304

Rick?
RICK: That's what I was actually going to address, is that mid-trial correction piece is I think going to be very problematic. Analysis, sure, we can talk about what we did wrong and how to address it for future studies. But for mid-trial correction, we want to be careful, as has been said, what we can correct.
We can retrain monitors if they're letting patients in as exceptions in the protocol. We can retrain -- there's a lot of things we can retrain, but there's a lot we can't. So we just have to be very careful when we write that section, what passes muster for mid-trial correction and what we should address.
DR. DWORKIN: So just out of curiosity, how many people in the room think it would be -- forget about FDA for the time being. How many people in the room, just as investigators, researchers, think it would be reasonable and acceptable to call the patient and say, "Mr. Smith, yesterday, you didn't complete your pain diary, and we hope that, you

Page 305

know, from now on, you are more careful about that."
MALE SPEAKER: If it was written in the protocol.
FEMALE SPEAKER: You can write in the protocol --
MALE SPEAKER: If it was prespecified in the protocol.
(Crosstalk.)
DR. DWORKIN: Okay. But Sharon also said yesterday that in no circumstances would it be acceptable to call Mr. Smith and say, "Hey, yesterday, you said your worst pain was less than your average pain."
(Crosstalk.)
DR. DWORKIN: I'm not going to say they're -- well, I'll tell you, if I'm drafting this article, I'm not going to say there are two different issues until Sharon says to me that they're two different issues because of exactly what Lee said. I don't want us to make recommendations that it turns out the FDA doesn't

Page 306

agree with.
MALE SPEAKER: But Bob --
DR. DWORKIN: So we all agree that there's no problem with calling the patient and saying you didn't complete your diary.
MALE SPEAKER: If it's in the protocol; only if it's in the protocol. It's got to be transparent, and you can plan them.
(Crosstalk.)
DR. VANHOVE: Exactly. I mean, very often, it will say if the patient hasn't filled it out for two days, there's going to be a call. It's written down, it's prespecified, and you follow that.
DR. DWORKIN: But Trudy, what I'm saying is if -- the way I understood Sharon yesterday is even if it was written in the protocol that you call the patient to say the patient's worst pain cannot be less than average pain --
MALE SPEAKER: But that's asking them to change the data.
(Crosstalk.)
DR. VANHOVE: So she had a problem with you

Page 307

going back and saying, "Hey, does this correct a score?"
MALE SPEAKER: And she backtracked a little bit on that, Bob, in the break. You should get it from her.
DR. DWORKIN: It might be that you're all right. I'm just saying I don't want to write that and publish it until we confirm it.
MALE SPEAKER: Of course.
DR. DWORKIN: We all agree.
(Laughter.)
Raymond?
DR. CHEUNG: I think in the conduct of the study -- and we talk about there are opportunities for training -- you don't need to necessarily reference that I know that you're doing it wrong. But as part of the training, that you can always remind people, are you taking your medication; are you filling out your electronic diary? I don't think that that would -- I think that might be less of a problem.
DR. DWORKIN: That was clearly not a

Page 308

problem, but what many of us wondered about is that kind of retraining on a regular basis all of the patients is obviously much more cumbersome, costly than targeted intervention. But targeted intervention might not be acceptable.
Dave?
DAVE: Yes. [Inaudible – off mic]. Part of medical monitoring in so many clinical trials to do -- there's a difference between training and coaching. And I think what's happening right now is we're combining those two things.
If you say to somebody, I saw what you did, right, and you're doing it wrong; let me tell you how I think it should be done, that's coaching. I think training is to just say, "You have to fill out your diary every day." That's fine to get a notification that you need to fill out your diary.
Those are different things. It's making sure that they're adhering to the protocol design is absolutely legitimate. And I think if you need to spend some time and say, well, I want to make sure you understand the difference between least

Page 309

pain, worst pain, and average pain, you can do that. You can train them on that. That's a difficult concept for some people to get.

But you can't go, "Oh, I saw that you wrote something that was really crazy yesterday; go change it because it was wrong." I think that's what Sharon was referring to.

MALE SPEAKER: Yes, exactly.

DR. DWORKIN: I blame Lee for all of this because --

(Laughter.)

DR. DWORKIN: I was just agreeing with Lee that we don't want to make a recommendation that's going to end up biting some sponsor six months down the line because they read our article, and they think something is acceptable when it isn't.

So I hope that we can all agree that we just want to make sure that our recommendations are either acceptable or unacceptable, and that we know what they are before we make them.

Ajay?

DR. WASAN: I think it's really important in

Page 310

this section to define context. And obviously, the context of an FDA phase 3 trial, such iterative processes, you have to have very tight parameters. But on the other hand, there's the opposite view, not for the FDA registration trials but some other kind of trials. Let me give you some good examples.

Obviously, there's agreement that the best science is done as an iterative process. Let's say your outcomes you're looking at are physiological outcomes, so QST changes or FMRI changes. Those are some of the studies that Rob and I do for instance, and that you use the clinical trial as a mechanism to look at changes in physiology, and that's your main outcome.

If someone is not adhering and you found out, it's kind of good that you talk to them about it. If they don't do their rating scales and something's bizarre about them, actually, since you're not primarily testing efficacy, you're actually trying to look at physiological outcomes, it's actually better that you have these iterative

Page 311

changes.

You got to be careful too, right? You don't go overboard and fundamentally change and compromise the primary outcome you're looking for, which may be physiological.

So I think we have to be really just careful on the context and define the context in which --

DR. DWORKIN: That's exactly the kind of language we will have. Depending on the context, targeted intervention retraining may be appropriate, but in regulatory contexts, don't assume it is without getting approval from the regulatory agencies. That's the kind of language I'm imagining. It's what Dennis said; it's a black box warning. Other comments? Laurie?

MS. BURKE: I think it might be an overkill to try to have a mid-trial column. You might just want to have this mid-trial considerations paragraph, and then -- the considerations are to change -- change your processes midstream are usually a bad idea, but there may be a reason to do something if you notice something that would deep

Page 312

six your whole program.

DR. DWORKIN: SP I agree that I hope that we end up with fewer than 36 cells --

(Laughter.)

DR. DWORKIN: -- either by deleting a row -- I mean a column, as you suggest, or by combining, as I just suggested, some of the rows. I would hate for the ultimate manuscript to go in with an Excel spreadsheet that I can't do myself with 36 cells in it.

Yes, John?

JOHN: To say something that may already be obvious, but I think the point is that studies can be designed to monitor certain things and implement certain changes if things are found. You design the study -- I mean as David was just saying, it's completely reasonable to encourage a continued enrollment and filling out the forms. And if you know that people are not filling out forms, that you contact them.

I don't think anybody would object to that, but it ought to be written in the protocol, which

Page 313

brings me to the second point, which I think Laurie would -- is partly what she's saying. And certainly, Sharon said yesterday, is be sure that you're upfront and transparent about what you do. And if you're in a registration trial, before you make any mid-trial corrections, you damn well better talk to the registering agency.

Honestly, I don't know how clinicaltrials.gov works in this score, but if you change the protocol halfway through, somebody is going to be upset unless you -- and I think you need to go there and actually make the change there as well as a change. I'm not sure. But my point is transparency is really key.

DR. DWORKIN: Other comments?

MALE SPEAKER: The only thing is you can't always anticipate, right? When you're doing science, you can't anticipate all the problems, so you just -- that's the other caveat too. You can't prespecify -- unless you say a general term, "If there's something that comes up I can't think about right now, then I reserve the right to make some

Page 314

changes."

(Laughter.)

MALE SPEAKER: I think John is right. I mean, everything you do -- the protocol is it. Everything needs to be in the protocol, but part of being in the protocol is that the study will be monitored. And part of monitoring a study is to ensure that things are followed.

As you were just mentioning, you may not be able to know everything that won't be followed, but part of the job is to make sure they adhere to the protocol. That's what the protocol is.

If they don't adhere to the protocol, that actually is really bad, and you don't need to have somebody -- you don't need to have extra provision from the FDA saying that you have to -- it's okay to adhere to the protocol and make sure that the sites and the patients adhere to the protocol. That's what a protocol is.

What I would say is, under there rubric John mentioned -- even if you can't predict it precisely, the point is we know with the protocol

Page 315

is trying to do. I'm hard-pressed to find an instance where you could do an intervention that would be -- as long as you're not changing the data or coaching the patient, I think you're fine.

DR. DWORKIN: Trudy?

DR. VANHOVE: I totally agree. I would say data falsification, if you identify that but you can't correct it, or you don't let the FDA know that, hey, I've got these patients that reenrolled 10 times or whatever it is, or misconduct, then what? Okay, you identified it, and what are you going to do?

DR. DWORKIN: I don't know.

DR. VANHOVE: You can't correct anything.

MALE SPEAKER: That's what's going to be fun about writing this paper.

DR. VANHOVE: I totally agree.

DR. DWORKIN: I think the issue -- and Paul, you've been silent. But the issue of what's appropriate when these things are identified in an ongoing trial -- I said this already -- I think it's the most challenging part of this paper to

Page 316

write. Just because you've anticipated some of it in the protocol, it doesn't mean what you say in the protocol was automatically right. If I write a protocol today saying I'm going to deal with missing data using LOCF, that doesn't mean I then -- I'm going to get a drug on the market using LOCF.

So yeah, we have to put the best things we can in our design, but we need to make sure what it is that's going to be acceptable to the regulatory agencies, unless I'm missing something. This goes back to Lee's point. We just can't run the risk of misleading people that something is acceptable when it might not be.

MALE SPEAKER: I think we definitely need Sharon to weigh in on some of this. Usually, statisticians aren't asked for this type of level of what's appropriate for a mid-trial correction. I think hitting the high points here have been -- we're more -- in the past, we've traditionally been much more focused on the site investigator and trying to train that individual

Page 317

rather than the actual subject. So that might be something that needs to be discussed. How far do we want to go in terms of interventions with subjects?

I personally happen to agree that it's fine to remind somebody, you haven't been taking your drug; you haven't been filling out something. Going further saying, do you need some help filling out your patient-reported outcome statement is a little -- that's starting to stretch things, and it's going a little too far.

DR. DWORKIN: Well, that's inevitable. What if they haven't completed it for three days? Can you remind them that they haven't completed it?

Trudy, I completely agree that -- so Dennis and I are doing a trial now on fibromyalgia. It's NIH-funded. Of course, we would. All I'm saying is I don't know with a hundred percent certainty that Sharon would say of course. She might and then we're all in agreement.

I think we're beating a dead horse here. Are we beating a dead horse, Dr. Turk? Yes.

Page 318

DR. TURK: Yes.

DR. DWORKIN: Rob?

ROB: Again, just to remind, in any trial, you may uncover violations, or deviations or errors. You can always, first of all, query. And as long as you're transparent in everything you do, you can identify. If you were to identify a patient who you thought was fraudulent, you could still transparently suspect that, put that patient and their data into a separate list and say, look, we suspect or we're worried about the data for some reason. And you could analyze it separately. Obviously, it's not the intent to treat.

But if you come with violators or deviators of any soft because of urine drug screens or faulty data, if you can document it and be transparent about it, you can analyze it, do a sensitivity analysis. I think the FDA would welcome that.

But as long as you're transparent about any errors -- and you can make mid-trial corrections as long as you're transparent, I think, almost at any level. But it may have implications depending on

Page 319

the impact or the type of corrections you're making in a study.

DR. DWORKIN: Other comments?

(No response.)

DR. DWORKIN: Are we done early? Andrew?

DR. RICE: It was just the issue of unannounced blood sampling as another measure of adherence.

DR. DWORKIN: Yes.

DR. RICE: We might ought to put that just as a one-liner. You can reference David Simpson's study where we did it. He was the first author under pros and -- we discussed the pros and cons of doing that.

MALE SPEAKER: Yes, and Bernard actually referred to some data suggesting that they were kind of dramatic important differences between what you got when you did announced versus unannounced.

DR. DWORKIN: Phil?

DR. CONAGHAN: Bob, I'm just a little concerned about the generic versus specific pain issues and almost the selling of this paper, as it

Page 320

were, how you make it related to pain.

A lot of the things we discussed are not just relevant to pain studies. They're relevant to lots of different trials. So to make this pertain to a pain audience, I'm assuming some way you're going to have to make examples that always relate to a pain study when you're writing your manuscript.

Is that what you've got in mind already?

DR. DWORKIN: Yes, absolutely. I mean that's right. A lot of this is very generic about clinical trials and not pain. Some of it is going to be very pain-related like training people how to do zero to 10 pain diaries, et cetera.

The issue you identified hasn't been a problem with other IMMPACT papers that were 100 percent generic. We have an IMMPACT paper with recommendations for how to deal with multiple endpoints in a clinical trial, and it's really all about statistical approaches to multiplicity.

I don't think there was anything specific about pain in that article. And I think those

Page 321

articles are -- there haven't been many of them -- are largely generic or maybe 70 percent generic, are viewed by the reviewers and editors as educational. That's I guess the way we've thought about it.

But this will have, I think, 25 to 35 percent pain, specific material in pain examples. So I don't know that it's a major problem.

DR. CONAGHAN: The other element that I think is part of good recommendation papers is to highlight at least some of the priority research agenda, which you normally like to get in. And if you got some time now, I'm thinking of a couple of things.

For example, even the training issues you've brought up to me are not well evidence-based. The issues of training people to use VAS or NRS scores or whatever, we need to see the evidence base to just make some difference after you've psychometrically adjusted these scales.

That's just one example, but perhaps while

Page 322

you've got people here, a quick thought of what the juice is for research recommendations would be good.

DR. DWORKIN: Yes. We'll definitely have a table with a research agenda that tends to kind of write itself, because as we're writing most of the paper, there are all these places, as Mark was saying, where we're going to saying the evidence is minimal or lacking. And that then becomes an item in the research agenda.

So there will be definitely be a research agenda that's driven by the holes in the evidence underlying our recommendations or considerations.

Dennis?

DR. TURK: I was going to respond to the first part of your question, which was about if it's broader than just the pain, will it get sort of seen or will it be picked up or observed or would that information get out there?

In the past IMMPACT and ACTTION papers, all of which have -- 99 percent of which have appeared in the Pain journals, they've ended up getting

Page 323

cited in over 600 different journals across the entire spectrum. So somehow or other, even though we're putting it in the Pain journal, it gets picked up much more broadly than we maybe anticipated.

DR. DWORKIN: Actually, that's the opposite of what, I think, Phil was suggesting. If the paper is 70 percent generic, it actually has a larger audience than -- so that's actually an interesting kind of --

DR. CONAGHAN: Those issues are really important for all trials.

MALE SPEAKER: But my point was that even though we're putting it in a Pain journal, it gets picked up.

DR. DWORKIN: Yes. Other comments?

(No response.)

Adjournment

DR. DWORKIN: All right. I wish I had one of these timers that counted down five seconds. You will be hearing from us because the way this works, and many of you are very familiar with this,

Page 324

is a manuscript will be drafted. Everyone who's been here will be listed as a co-author. You are completely free to ask us to take your name off it, and we're happy to do that. If you don't, you will be involved in multiple revisions of the paper before it gets published as an article somewhere.

Thank you very much, Valorie and Andrea for coordinating a wonderful meeting. Thank you all, and especially the presenters, for your participation, your ideas, your thoughts, and your terrific presentations, and have a good safe flight home.

(Applause.)

(Whereupon, the meeting was adjourned.)

B

C

F

I

J

M

N

P

S