ACTTION SCEPTER II - Clinical Trials to Evaluate Safety Outcomes in Procedural Sedation

---

November 19, 2016

A Matter of Record
(301) 890-4188

ACTTION

SCEPTER II MEETING

Clinical Trials to Evaluate
Safety Outcomes in Procedural Sedation

Saturday, November 19, 2016
7:57 a.m. to 2:47 p.m.

Sofitel Washington DC Lafayette Square
Washington, D.C.

C O N T E N T S

C O N T E N T S (continued)

P R O C E E D I N G S

(7:57 a.m.)

DR. WARD: Good morning, everyone. Thank you for being here for the second day. I thought yesterday was an incredibly productive and exciting day. I tried to organize this so that yesterday morning, we spent our time talking about what the adverse events were out there. So it had a little bit more of a QI flavor because how are we measuring adverse events from procedural sedation? What's the problems clinically with the current sedatives and the way they're used?

In the afternoon, we spent some time talking about the difficulties in performing clinical trials, in which we want to measure the frequency and severity and types of adverse events that take place.

What I want to do this morning is now we're going to focus a little bit on some more specifics. If we're designing clinical trials for adverse events, what are the adverse events that we want to measure as the outcomes? How are we going to

Page 5

measure them?
We're going to spend a little bit more time on respiratory because that's obviously the most common and most devastating adverse events that happen with sedation. But we're also going to cover gastrointestinal, cardiovascular, neurological, and liver-kidney.
Then this afternoon we're going to try to pull it together with some discussions. When I say panel discussions on the agenda, it's really discussions of everybody here.
I know we all have flights to catch this afternoon, so what I will propose to do is to come back from lunch at 1:00, and then we'll have a two-hour session looking at recommendations that we can be making -- because I'm going to go home and start writing the paper to send around to everybody -- and get through at 3:00. And there'll be some snacks out there that you can take with you as you go out the door. And that way people who have 5:00, 6:00 flights can make them without a problem. So if that's agreeable to everybody,

Page 6

we'll come back at 1:00 and adjourn at 3:00.
So we'll start out this morning with a review of the literature that we have started, what's out there for clinical trials, and Mark is going to start out with that.
Presentation - Mark Williams
MR. WILLIAMS: Thank you, Denham. Welcome and good morning, everyone.
So as Denham said, we started with a review of the literature just to give us a platform of what trials are out there, what safety aspects they are using, trying to cover. And the initial search was performed -- we did eventually, unfortunately, end up with quite a restrictive search because the search terms "sedation" and "safety" bring up hundreds of thousands of articles.
To bring this down to some manageable amount of articles, we ended up looking at procedural sedation studies, which were clinical trials. And we ended up keeping it quite restrictive into randomized clinical trials, which were double-blinded. So we understand we missed, obviously,

Page 7

many trials, but we also included a lot of trials. And I think it gives us a state of the play of what these trials are including when they are trying to look at safety.
We obviously excluded -- for this meeting, we excluded ICU trials, general anesthetic, general anesthesia, and retrospective studies. We ended up with 500 studies, of which we excluded about 400 initially. It left us with about 141 which met the exclusion criteria, and we looked at 133 articles for this analysis.
No information was put in 1 percent of the studies. We looked at what -- or generally classified them into two different classes. Some studies just said, we recorded vital signs as safety assessment, and others were much more specific on what they were recording.
About 50 percent, 54 percent, of studies were specific on what classified as safety measurements, and of those 54 percent, 46 actually also included interventional techniques or interrupting procedure as items included in their

Page 8

safety assessment. And 11 percent of those 54 percent used a standardized tool, such as Beck or World SIVA, to try to assess their safety outcomes.
Reviewing which safety outcomes they looked at, as mentioned, respiratory, every one of the studies that defined their safety outcomes used a respiratory item as a measure of safety. Surprisingly, only 69 percent used some cardiovascular measure; and then fewer, about a third, for neurological and gastrointestinal items.
These were two studies for the efficacy of fospropofol, but they also captured safety data. And this is similar to some of the studies that went through safety. These are actually similar across the studies, looking at desaturation.
Here they used less than 90 percent for more than 30 seconds, and heart rate less than 50, requiring an intervention, or blood pressure less than 90 or so, requiring an intervention; so similar across the fospropofol studies here.
These are the dexmedetomidine studies for

Page 9

the FDA license for efficacy, also including the safety. These are slightly different, definitely different from the fospropofol, slightly different between the studies. These use the time of blood pressure outside of the range. I think this was an area-under-the-curve study.

Here they're using a percentage of blood pressure, 30 percent or more change from the baseline. And the other study was 25 percent -- and for 30 percent, and for respiratory rate was less than 8 breaths, 25 percent. So a bit of a difference across studies and between studies.

I believe this got sent around. This is just some of the studies that we went through trying to just capture what they looked at, again just showing the lack of consistency across studies in not only timing, but the adverse events also, the actual values that were used.

So the main items, as we're well aware, most studies captured maybe a saturation or a resp rate. A lot of studies are starting to use end-tidal

Page 10

monitoring. What that value actually means isn't often clear, but they at least capture it, and then other important respiratory items, cardiovascular, usually blood pressure, heart rate. And then GI, predominately nausea and vomiting.

Neurological was often a wide span if they include. Sometimes excessive sedation was classified as an adverse event. Even looking at not unpleasant recall but recall, were they able to form a three-object recall later on, or did they have persistent amnesia for events or mini-mental scale assessment decrease. Then generally capturing other side effect data.

Severity, unless using a designated tool that we discussed -- it was discussed well yesterday -- often ambiguously classified. Sometimes it interfered with a patient's function -- that was classified differently -- or met a certain intensity or a certain deviation from normal physiology. Otherwise, classified as an interrupting procedure or not, or requiring treatment or not, those seem to be the usual

Page 11

classifications.

Serious adverse events, not often discussed. Serious adverse events from the good clinical practice definitions were not often discussed.

Then, as we mentioned yesterday, time frame, how long do you allow someone to be outside of these parameters before you classify its ranging, in these studies that we looked at, anyway, from 10 seconds to 1 minute, and then some recorded the cumulative time, cumulative repeated times outside of the range.

So there's not really good consensus on which items are important to capture, what should be mandated to be captured throughout the studies, which parameters to use, what time frame to use, or how to classify them for severity.

So with that lack of consensus, we hoped we could get some consensus across the group with your feelings across some questions. So I'll hand it over to Denham, and he'll take it.

Q&A

DR. WARD: I think we're going to open up

Page 12

for questions first from the others. I have a couple.

How common in those papers were there an active control?

MR. WILLIAMS: An active?

DR. WARD: We pointed out yesterday, did we run into a clinical trial that --

MR. WILLIAMS: Yes. So we only looked at papers with active controls in those groups, yes. So that was the randomized, double-blind, controlled trial, yes.

DR. WARD: Do you know what kind of control group, what drugs and procedures they used for the controls for some of those?

MR. WILLIAMS: It was usually -- if they were looking at ketamine, it might be enzo and -- enzo, that they were looking at, yes, and other sedative medications.

DR. KOCHMAN: Mike Kochman. So the problem with these constructs of safety outcomes -- sorry -- of the adverse events that are being looked at is that often they can be

Page 13

relatively insensitive to what we would look at and say are really important, as opposed to physiologic changes that may occur.

So in GI, one of the things, for example, that we look at is there are a whole bunch of different scales to grade pancreatitis. And we can argue that those scales are insensitive for the detection of significant pancreatitis, moderate pancreatitis, and mild pancreatitis. We throw those terms about, but we don't know what they actually mean.

So one of the things I'll just throw out is, does it become an AE, an SAE, when we have to alter the intended procedure that the interventionalist is performing, or you have to change your anesthetic plan from what you intended to do, a priori, with an intervention that was not part of the goal of that anesthetic plan?

MR. WILLIAMS: Good question. I don't know.

(Laughter.)

DR. KOCHMAN: I'm looking for help here.

DR. WEISS: Like many things at least about

Page 14

the anesthetic plan, the answer is, I would think that depends on the situation, and what the patient is presenting to us at the time, and what the needs of the proceduralist are presenting to us at the time.

There are times where we certainly have to change the anesthetic plan, or what we're doing, because now we're in trouble. And there are times where you may not be in trouble, but I know that I can give a proceduralist a better feel, which will allow him to do his job, which will better take care for the patients, even though the patient isn't having any drugs given at the time.

So it seems like mainly -- it seems like there's no clearcut answer to that. It all depends on what the situation is on the ground at that time. And when we do these procedures, we talk about this all the time. So we're communicating this back and forth all the time. I think that's going to be, too, as well.

DR. WARD: Any other comments on what's out there in the literature? One thing you should do

Page 15

with a systemic review, besides the computerized literature searches, ask the experts in the field if they know of any other papers or unpublished papers, or we'll send around the list of papers that we have found, and for you to add any to them that you think should be added.

Rick?

DR. RIKER: There's a study in press in JAMA called the SIESTA trial. It's ischemic stroke thrombectomy, general anesthesia versus conscious sedation. It's by the Heidelberg group. And I think to me it brings in a couple of issues.

Excluding general anesthesia would have lost that study, although maybe it would have popped up with conscious sedation. But it was a very interesting study because they looked at long-term safety and efficacy, and actually showed an improved, modified rank in 3 months in the general anesthesia group. So it may change our focus a little bit on what we consider the duration of effect of a sedation or a procedure.

There were some time differences between the

Page 16

groups. It took 10 minutes extra to get an arterial puncture. Maybe the anesthesiologist was having lunch or something. But time from puncture to reperfusion was actually quicker in general anesthesia.

So some things you don't necessarily think about. And counter-intuitively, the incidence of pneumonia was higher in the general anesthesia group, kind of the opposite of what I might have expected. So I don't know.

DR. WARD: Yes. I guess the active control could be general anesthesia. Besides, we think of another sedation protocol. But it could be general anesthesia, too. I was interested in -- because I hadn't really looked at all these papers, of course, as you had -- how well did they describe their measurement techniques? I'm thinking particularly of respiratory rate and apnea.

MR. WILLIAMS: Some of them, not.

DR. WARD: Because clinically, assessing apnea of greater than 30 seconds is not necessarily a very easy thing to do.

Page 17

MR. WILLIAMS: No. And some didn't define how they were assessing these. They just said that they were assessing them. So obviously, the methodology is important.

Going back to how long do you follow them up, it's a very good question. If you're looking at adverse events, if you don't have them within the first 5 minutes after they recover, is that okay or do you have to follow them for months?

DR. RIKER: Yes. I guess the other issue is long-term cognitive issues. Right? I know this is a very controversial area, but probably should at least mention it.

Group Discussion

DR. WARD: I asked the group if you had any questions you wanted to survey the rest of the group. I didn't get many responses, so I made up my own. So if you don't like the questions I made up, you can just blame yourselves for not sending me better questions.

So I thought we'd just go through these one at a time and discuss them as we go through. And

Page 18

as I said, I make no claims of validity of these questions, and the other is more or less random.

So one question I asked was, should at-risk patients be included in phase 3 sedation clinical trials? We talked about that a little bit yesterday. Do you make sure you've got patients with -- kids with big tonsils when you're doing sedation trials, or OSA in obesity?

Most of us hedged our bets and said, it may be necessary for some trials, although a substantial proportion said, yes, always, there should be trials with that. And fewer thought that it really is not required in a phase 3; you can really put that off to a phase 4 trial.

Comment, Jerry?

DR. LERMAN: The one minor problem in pediatrics, of course, you can define and observe and diagnose obesity well enough. But OSA, way more difficult. So you're liable to trap OSA patients in your studies whether you think you're doing it or not, deliberately or otherwise, and they're going to be in the study.

Page 19

They may be an issue in that an opioid that was administered is going to have a more untoward effect in those patients. And that can surface as an adverse event.

So excluding those patients is somewhat important. If they actually have documented polysomnigrams and nocturnal hypoxia.

DR. WARD: Well, you can probably do better trying not to include high-risk patients than you could being --

DR. LERMAN: Well, certainly, but that's what I'm addressing is the fact that you're going to trap them, generate AEs, SAEs, unwittingly.

DR. KOCHMAN: I would argue that the question in the phase 3 trial is what is that phase 3 trial? Is that the pivotal labeling trial or not? If we're looking at a pivotal labeling trial and exclude certain patient populations deliberately, then what is the onus later?

To get the manufacturers to go back and change label so that we're operating within the label becomes very, very difficult. And I would

Page 20

argue that properly trying to size and properly subgroup, or even if you end up with the issue of unintended events, be able to record those issues so that you don't end up boxing yourself out at a later date of being able to use product in somebody even though we know in a phase 4 it may work.

DR. WARD: Right. I think the issue becomes if you start including some of these high-risk patients, you're going to start to get a higher incidence of maybe not SAEs, depending on what you're measuring, but certainly getting what we would be classifying as AEs.

DR. KOCHMAN: But it also makes your power easier, too.

DR. WARD: Right. Exactly. It may be something you may not want to know at that point. But we need to know.

DR. BERKENBOSCH: So this is a trial with a comparator. You may have an increase in adverse events --

DR. WARD: In both groups.

DR. BERKENBOSCH: -- but you've got both

Page 21

groups. So you've got that comparator to balance the increased risk.

DR. WARD: Exactly. If you were trying to compare propofol with dexmedetomidine in obese patients with sleep apnea going for a GI procedure, you'd have both groups.

DR. BERKENBOSCH: If you've got appropriate comparators, it may not. Those continue to wash out a little bit more.

DR. WARD: So how about a noninferiority pragmatic clinical trial with a new sedative agent protocol compared to a current regime? By pragmatic, I mean, you let -- with the current regime is you let the people do what they usually do, and not as this -- not specify quite as rigorously as you would for a non-pragmatic trial.

So pretty well split between yes, you should and recommended but not required. Dan?

DR. SESSLER: Suppose you do a study testing a new agent against placebo and you find that it's better than a placebo?

DR. WARD: Can you ever do a sedative study

Page 22

against a placebo in a late-phase trial? Certainly, in phase 1 that's what you're doing. So that's why I said compared to current regime.

AUDIENCE MEMBER: You can use a multi-drug regimen and you certain can use a placebo.

DR. WARD: Right. That's sort of what I was getting at because the current regime is usually a multi-drug kind of regimen. We've done one in the GI suite and let the gastroenterologist use what he wanted to use, and we saw a huge variety of different combinations that it wasn't really sure why he picked that combination for that particular patient. It wasn't even consistent across patients.

DR. SESSLER: Well, along those lines, routine treatment is a terrible control. We have to use it sometimes, but it's inherently bad because it varies tremendously among institutions and even within institutions. And it's very hard to characterize.

So if you're reading a paper where the control is routine treatment, you don't actually

Page 23

know what the routine treatment is. So generally, it's best to specify the controls or the actual [inaudible – of mic] period.

DR. WARD: So you'd recommend that if we were going to do a clinical trial with a new, to actually not do this kind of trial but to really specify.

DR. SESSLER: A specified control is always better.

DR. WARD: Frank?

DR. DEXTER: I know the goal is the safety issues, but yesterday when people talked about the time reduction, one of the other problems with the routine care is that it can differ endogenously related to the work flow.

In other words, it's routine care at one organization because their work flow versus another one, but that work flow is oftentimes not characterized, not described in papers. And so then it's so difficult to figure out whether the economic result, the time reduction results, would be generalizable or not.

Page 24

DR. WARD: Rick?

DR. RIKER: I would say we'd have to be careful of not throwing the baby out with the bath water here because it really varies with the question. The FDA is really pushing for longitudinal safety studies, which essentially are pragmatic registry trials. So I think the goal would probably be to help the FDA couch these in the appropriate manner, given their mandates.

DR. WARD: Other comments? I designed these more to create controversy than necessarily to get consensus.

So is there an appropriate scale that should be used to measure emergence delirium? So most are saying no. Some didn't think there was one, but a few people said that there is. So those of you who said there is an appropriate scale to measure delirium --

DR. LERMAN: Well, since I entered it and published it and it's in "Anesthesiology," I would think that's the only validated scale. If you don't know about it, well, that's fine.

Page 25

(Laughter.)
DR. WARD: You haven't adverstised it enough, Jerry, because obviously we don't know about it.
DR. LERMAN: That's not my job.
DR. WARD: You want to describe the scale?
DR. LERMAN: Yes. It's 5-point scale. It's a little labor-intensive to administer for a nurse. It's easier to do a scale of 1 to 4, which two other groups have developed. And unfortunately, the only diagnosis in the other scales of 1 to 4 are thrashing and unruly behavior.
Our scale actually specifies three scales on the DSM delirium criteria, which includes inability to associate with the environment, fails to perform purposeful movement, and out of touch with the environment, I think.
The other two may be overlapping with pain, and that's part of the little issue that's going to come up here, which is if you measure or attempt to measure emergence delirium in a procedure for which there is post-op pain and you haven't controlled

Page 26

the pain completely with a block or some other means, you risk some interaction between post-op pain and an emergence delirium.
But by and large, the criteria we use, which is a value of 12 out of 20, gets you almost -- 10 or 12 out of 20 -- gets you most of the emergence delirium.
DR. WARD: Okay. Appropriate scale to measure nausea?
DR. LERMAN: Not in children. Most children don't know what nausea is.
DR. WARD: Correct.
DR. LERMAN: Excluding adolescence, perhaps.
DR. WARD: In the chemotherapy world, I would think there would be some scales. But it turns out most of us don't know. So those of you who knew there's an appropriate scale to measure nausea?
DR. TONG: But there is a facial scale that probably people typically use for the younger children. I'll show it when I --
DR. WARD: Okay. You're going to show that

Page 27

in your talk? Okay. Great. Because nausea, really, as opposed to what? Vomiting is more of a measurable outcome. So I was trying to get at an outcome that was a little bit more continuous than the dichotomous vomiting or no vomiting.
Frank?
DR. DEXTER: The issue in terms of nausea, "Anesthesia and Analgesia" had a series of articles. One of the challenges is that not only would it be an ordinal scale, it's a highly skewed ordinal scale. And so that when you see, for example, is that even when you try using relatively valid instruments, you will see, for example, a difference by a Mann-Whitney test. I get the medians are identical or near identical.
So I think that one of the things that is better in the past two, three years is mathematical analyses are improved. That's the good news. The bad news is that those methods are at least currently uncorrectable for covariance.
So in a purely randomized trial, reasonable, but beyond that, it's a real challenge to deal with

Page 28

it because the methods designed for the ordinal variables rely upon a very symmetric distribution within the rank scale.
DR. WARD: And what do you mean by -- how was the nausea scale skewed? A lot versus --
DR. DEXTER: Yes. And in addition, what you're seeing is that you have people who have large amounts of nausea move towards the center. It's not a shift, is what's happening. And therein lies the challenge.
DR. WARD: Rick?
DR. RIKER: One of the challenges we've had in this area in the ICU is how do we score these things. In other words, let's say we have a valid data and a well-accepted nausea scale. Now what do we do?
Are we going to say it's the presence or absence of nausea, using the scale, any degree of nausea gets you into it, and it's a dichotomous variable? Is it going to be the intensity of the nausea? The duration of the nausea? I don't know if there's any patient-centered data here. What do

Page 29

they value? What's important to them?
So I think for every one of these things, it gets a little complicated to know how do we actually quantify and score, and what's the important aspect of it to us and to patients.
DR. ROBACK: Maybe it's my pediatric orientation, but like Jerry says, it's really hard to measure in children. And so we go back to interventions, and was it enough that they received an antiemetic for it? And that would say that maybe it's significant. But again, you don't get gradation of what the actual degree is.
DR. WARD: Again, I was trying to get at non-dichotomous outcomes to help. But I thought that in fact, when the adverse events may be rare, then you're trying to find out some non-dichotomous outcome.
DR. DEXTER: So I have the paper, the one Devine, where we had a series of studies on the statistical analysis and so forth that was part of -- studies of aromatherapy and stuff like that. The thing with the -- clearly, it's all going to be

Page 30

adults, and it very much becomes one in terms of almost binary.
The reason being is because you're going to have patients with large amounts of nausea, and if you think about, for example, a Wilcoxon test, all the ties completely go away. So it essentially is going to become mathematically as if it were to be binary. Think of it as extreme amounts of nausea become average or zero, is the way that essentially works.
DR. WARD: Okay. So in pediatric patients, how long should the follow-up be after the sedation? Regression in milestones, nightmares, long-term neurological. I guess the majority wanted to go at least a month for follow-up, with only a few as 48 and a week. So that speaks to somewhat longer follow-up than what we normally see.
Jerry?
DR. LERMAN: Oh, I'm sorry. I thought you were finished.
DR. WARD: No. I'm just saying I don't

Page 31

think we see very many studies for procedural sedation that has given follow-up for a month or more.
DR. DEXTER: Yes. That's what I was going to ask, if Mark came across any papers that actually went to postoperative behavioral outcomes, nightmares, cognitive issues, up to a month afterwards.
MR. WILLIAMS: Not up to a month, but some certainly looked at the immunological outcomes over the next 48 hours, including nightmares. But yes, I didn't come across any out to a month.
DR. LITMAN: Denham, the best we have in this area are some very old studies that looked at not sedation per se but hospitalizations, and that the whole hospitalization data and what happens to these kids' fear of doctors, regressive behaviors, it's just so foreign from the hospital environment that there is today, where back in those days they didn't even let parents in, and kids got a lot more shots than they do now. It's so hard to know.
DR. WARD: Yes?

Page 32

DR. LIGHTDALE: And there's the fact that kids, unlike adults, are probably undergoing procedures because they are ill. They're going to have so much confounding in terms of that month interim as opposed to, let's say, a screening colonoscopy, which everybody needs.
DR. WARD: What I'm sort of hearing, though, is different than what everybody responded.
DR. LIGHTDALE: Or you had a bunch of adult --
DR. WARD: Yes. Yes.
DR. BERKENBOSCH: I think there's potential value to looking our further. But the question is finding a control group where you've undergone a procedure with maybe -- there's going to be very few procedures, at least in kids, because it's probably more of interest in kids than it -- well, the question's in kids.
You almost have to look at that age group where you could or you couldn't get the MRI with or without sedation, and look at those that underwent MRI without sedation in a similar age group to

Page 33

those who underwent with, to try to tease out some of those confounders -- procedure-related anxiety, not just the sedation-related stuff.

But I think that unfortunately takes out some of the age group that's really at risk for these longer-term behaviors, at least as has been described in previous anesthesia literature.

DR. WARD: Dan?

DR. SESSLER: I guess I'm a little confused. If you compare the novel treatment with the usual treatment, if you're using a randomized design, there is no confounding. Both groups will have some baseline effort, but --

DR. LIGHTDALE: You're going to have so many --

DR. SESSLER: -- you're looking at the difference between the two.

DR. LIGHTDALE: I'll put it back. There's heterogeneity in what's going on with the patients. And so unless you --

DR. SESSLER: Sure. Sick patients are inherently more variable than well patients, yes.

Page 34

Variability is one of the factors in sample size. We need to study more patients [inaudible]. Nonetheless, you're comparing it to some other treatment, and you're looking at the difference between the two.

DR. WARD: So if the patient population in a procedure would be appropriate -- I'm going to take MRIs -- and you're comparing a new drug -- I mean, say if you were even comparing ketamine to propofol. Pick two old drugs, one of which may have some longer-term effects.

Are we saying that if the procedure and timing -- so where you got reasonable normal kids getting MRIs versus hospitalized for three weeks, a lot of other things going on -- that it would be appropriate to have follow-ups of at least a month? Is that kind of what the group is --

AUDIENCE MEMBER: It would be novel information for any drug nowadays.

DR. LERMAN: Well, what you're doing is escaping any criticism that you failed to detect some long-term or longer outcome events. The

Page 35

probability of yield, very small. But you won't be criticized if you've made it a 24-hour cover; then you may end up missing something. You don't know, with any new drug.

DR. WARD: We're trying to come up with some recommendations with this group.

DR. LERMAN: Yes. Well, I just throw out a week and say a week is a compromise time-wise and for efficiency of doing studies, and see what everybody else thinks.

DR. WARD: Jenifer?

DR. LIGHTDALE: Wouldn't the phase 1 study really need to focus beyond nightmares, even, on neuropsychiatric testing, pre and post? That really needs to be sorted out early on with this controversy going on.

DR. WARD: Is there a need for clinical intervention? Reversal agent? Treatment? We've been talking about this a lot; particularly a lot of the QI databases are really run by intervention, intervention plus, and something, and something else.

Page 36

Is that the best primary outcome measure for a sedation clinical trial? Is reasonable consensus there that it's not?

DR. BERKENBOSCH: So you have to define "need" for an intervention a priori. That has to be a threshold. If you need an intervention, you have to almost provide a threshold for it. A lot of people will receive an intervention whether or not they needed it.

So if you want to use the word "need," I think you have to define what that threshold will be to intervene first because then you cross the threshold. By definition of the trial, you needed it. If you're using it at clinician's discretion, I think you report that they received it, and that may or may not be as clean an outcome to look at.

DR. WARD: Dan?

DR. SESSLER: Also highly context-dependent. A jaw tilt is an intervention, but it's also trivial. Who cares? I can use my example from yesterday, if you have to stop an MRI to put oxygen on, that's a big deal. So if you're going to use

Page 37

that sort of outcome, you need to define it very carefully, and they have to be interventions that matter with enough content.

DR. WARD: So I would say if you've got a clinical trial in which you're collecting the data that generally would indicate what the need for the intervention is, do you really need the intervention piece?

You've got the saturation data. You've got the respiratory rate. You've got the nausea scale that you're measuring. So I'm really asking it, is the intervention -- which is the clinical outcome; right? The quantitative data that we're measuring is not the clinical outcome. That's a surrogate for it.

I agree it'll be dependent on the procedure and the risk tolerance of the provider. Somebody may put oxygen on. There's a lot of things that -- other psychological research says if your previous patient had a saturation of 85, then your next patient you'll probably tolerate a lower saturation.

Page 38

That's been well-documented, that we have biases based on our situation. So there's a lot of difficulties with a straight clinical intervention, but yet that's the real outcome, is having to intervene.

DR. RIKER: I think if we moved in that direction, I would echo Dan's comment. It's got to be a clinically significant outcome that we talk about, not something that is 5 seconds in duration and has no long-term impact.

DR. WARD: Or something that interrupted the procedure.

DR. RIKER: Interrupted the procedure, extended beyond the procedure, caused harm in some way.

DR. GREEN: Yes. I'm certainly a fan of interventions because then they let you know what is clinically important. We keep talking about this hypoxia. Some people do procedural sedation without supplemental oxygen. Other people do it with high-flow oxygen.

So trying to just say, was there a

Page 39

desaturation, if you're using high-flow oxygen, you could have a couple minutes of apnea with propofol and never lose that. But yet an intervention such as assisted ventilation is a tangible, objective measure of something with a clinically important respiratory depression.

DR. WARD: So again, differentiating between a QI-type database and designing a clinical trial, so could you design a clinical trial in which -- if you define the intervention, if there's an intervention using MRIs, then putting on nasal prongs would be a significant intervention. If it's in the GI suite, that might not be. But could a clinical intervention be the primary outcome of a clinical trial that you are designing?

I see Dan shaking his head yes. We said no here, but maybe my question wasn't very clear.

DR. SESSLER: The question was, is it the best?

DR. WARD: Okay.

DR. CARLSON: But if we don't choose intervention, which is measurable, then we have to

Page 40

agree upon whether the biophysical markers, of which that nets. Maybe intervention is a proxy, not perfect, but actually coming to agreement about what is the measure would be really hard if it wasn't an actual act.

DR. WARD: Dan?

DR. SESSLER: Maybe the consensus here is that physiologic measures by themselves aren't very helpful because they often change in ways that don't really mean anything; saturations, for example. And the interventions will not be important, and it's highly context-dependent.

It's probably not going to be possible for us to come up with a simple solution. We're going to have to recommend to the investigators we have reasonable endpoints. In a given context, even, the procedure, the type of sedation, and who's doing the procedure.

DR. WARD: Yes. The difficulties we had with midazolam decades ago illustrate that, when I was a junior faculty -- I was a junior faculty and you as a resident, and we were squirting midazolam

Page 41

when it first came, we didn't have any problems with it. But then they started using it with some fentanyl out on the [indiscernible] suite, and then there was a lot of problems with it.

DR. SESSLER: And along those lines, different drugs have different expected complications. So if you're evaluating ketamine versus midazolam, you're looking for different sets of side effects.

DR. RIKER: Generally, when we're deciding on a primary outcome, it's an efficacy outcome, not a safety outcome. And I would think for a sedation trial, the successful completion of the procedure or something along those lines would be the more likely primary outcome, and then there'd be a number of safety outcomes.

Am I off base? I don't know.

DR. LERMAN: I would agree with you. I'm just interested that you listed that as a primary outcome measure. I would think that these are not primary outcome measures. It may be primary safety outcome or a secondary outcome outright.

Page 42

The primary outcome is that you're delivering sedation, how the procedure is completed, and the patient left the hospital. The secondary outcomes are everything else that happened. And you can call them mild, moderate, and severe.

I think it's crucial that we record these interventions because anybody else who is beginning to give this drug would like to know, for example, that 50 percent of the patients need a jaw thrust. A, they may not want to use that drug, or B, they'll be better prepared to intervene in an airway before a disaster actually happens.

DR. WARD: Suzie?

DR. KARAN: I think you're going to control for the intervention, though. And you're going to design a priori the clinical trial, then measure respiration other than just counting it, like put monitors on to measure respiratory activity or flow or something, and then measure the intervention. And then you can confound for who the person is, so you get it before the intervention has happened.

Page 43

You're going to design it from the beginning.

DR. WARD: I'll talk a little bit more on respiration. But should a provocative test be used in early clinical trials for a new agent to provide more information about possible propensity for adverse events? And you see most of that in the respiratory studies.

Kind of split, most with no opinion. And let me expand on that when I give my subsequent talk. The difficulty of that just is that very few of the provocative tests that we have, have any validity in connection to adverse events in clinical use.

DR. DAHAN: And they're very difficult to perform, especially in patients. It may be impossible.

DR. WARD: Well, some of them. I mean, hypercapneic response is the traditional one that people have used.

DR. DAHAN: Still difficult to do.

DR. WARD: The end tidal CO2 as an outcome during sedation clinical trials? Maybe outcome's

Page 44

the wrong word there, but should it be reported in sedation clinical trials? I was surprised by that. I thought that would be much more yes than a resounding maybe.

DR. PRATIK: [Inaudible – off mic].

DR. CONWAY: But it's not really the number of the end tidal CO2 that you're wanting to report in the sedation clinical trial. It's the changes and the wave forms or periods of apnea, increase or decrease of 10.

DR. WARD: Right. But should it be reported? If you do a clinical trial, should one of the measurements that you make in the clinical trial be end tidal CO2?

DR. LITMAN: That would make it a lot more difficult to do, and that's a consideration.

DR. LERMAN: You don't measure CO2 in every patient?

DR. LITMAN: Because it's not always completely accurate using either a nasal cannula. There's a lot of kids especially that don't tolerate nasal cannulas. It's very common that we

Page 45

use blow-by, oxygen blow-by, to get through an MRI or whatever. Sometimes we put the kids on their side.

DR. WARD: But we're talking about a clinical trial, right, so we're talking about something that you're designing --

DR. LITMAN: Who volunteer. Volunteers.

DR. WARD: You've got volunteers in patients. But you've designed a clinical trial to look at a new sedative. Are you going to require that you have end tidal CO2 in that clinical trial?

DR. LITMAN: If you could control for it easily. If you think it would be accurate. I just -- when I think about children, I think about how hard that is.

DR. LERMAN: Well, I think it's more or less mandatory. And part of the reason is the implications of sedating perhaps a child with a particular disease, who has already pulmonary hypertension or may have a congenital heart difference, who cannot tolerate an increased CO2, if we knew that sedative regiment caused an

Page 46

increase in CO2, we would clearly avoid it.

So the issue is, knowing what the effect of the sedative regimen is on end tidal CO2 compared to the standard regimen, assuming -- and we generally do in every case, including MRI -- for children having sedation. So I think it would have to be part of the study.

DR. LITMAN: I agree with everything, Jerry, but there's no reason to think that other parameters won't be predictive of end tidal CO2. Like I can't think of a situation where your ventilation would be normal but your end tidal CO2 would be higher than you expected.

DR. LERMAN: That's right. But how do you measure tidal ventilation, except with my little monitor, not invasively? Because there is no technique except the one that --

DR. SESSLER: Impedance ventilation.

DR. LERMAN: I'm sorry?

DR. SESSLER: You can use impedance ventilation.

DR. LERMAN: You can, but you're not -- for

Page 47

example, in an MRI scanner, and it's not -- never seen it used in a child or in a pediatric study. It may be used, but it's just not practical to recommend it for a clinical trial in sedation in pediatrics, in my opinion.

DR. SESSLER: I imagine the most common response was generally, but not always.

DR. LERMAN: I suspect --

DR. SESSLER: You can say it should be done if you can, and recognize that in certain cases, it may be difficult.

DR. WARD: James?

DR. MINER: I think if you're looking at -- we're always going to use apnea and airway obstruction as major outcomes in any sedative trial. And if you're expecting an observer to record those based on their impression, it would be so biased it's almost useless, especially apnea. Nobody sees apnea because they don't want to.

End tidal CO2 always picks it up. It doesn't matter if you're doing blow-by. It doesn't matter if they're standing on their head, it's

Page 48

going to pick up apnea accurately and without anybody's impression being biased into the measurement.

DR. WARD: This is from Aaron, actually, suggested this. Transcutaneous PCO2 be reported? It gets around some of the problems that Ron was talking about.

AUDIENCE MEMBER: There are always problems with that, too. It's more serious problems.

DR. WARD: I'm going to move ahead here because I want to get to my talk.

(Laughter.)

DR. WARD: So these are applied to adults. Critical value of end tidal CO2 be considered an adverse event? And most people wanted greater than 55, which goes to speak that CO2 is not a particularly good biomarker of an adverse event. High CO2 really by itself causes few if any problems, long-term problems.

AUDIENCE MEMBER: One of the problems we get into with CO2 in trials is that when people obstruct, it makes more room air. And if they're

Page 49

using nasal cannula, the value drops. I don't think it works really that accurately for picking up a high-end value.

DR. WARD: Right. I'm just assuming you can get an accurate measurement, and there are obviously a lot of technical difficulties with measuring end tidal CO2.

AUDIENCE MEMBER: I think recognizing that's not the only CO2 value that's of importance. I mean, the maximum value isn't the only thing of importance. Again, if you're obstructing and the CO2 suddenly plummets, that's also something that's worth looking at.

DR. WARD: Right. That's more of airway obstruction, which end tidal CO2 is probably not a good way to measure airway obstruction.

DR. MINER: But it's consistent, though. If your value all of a sudden drops, it's consistent from obstruction.

DR. WARD: A lot of other things, too, a lot of mixing of not a good exhale.

Is there a difference in CO2? We want to

Page 50

have CO2 that'll go up pretty high to be not any worse than existing regimen. Blood pressure change, 20 to 25 percent? Heart rate change? And most people said change was not important.

DR. LITMAN: Dexmedetomidine. Some of those kids get pretty low, and we just watch them and they're fine, and their blood pressure's fine.

DR. WARD: So how large a change? So a pretty large change, and most people would say a change isn't important. It's the absolute values.

DR. LERMAN: The difficulty here is you're looking at a change, and it depends on your baseline values because compared to a patients' baseline, an un-premedicated child who comes into a strange environment and you do vital signs, you're going to get extreme values.

If you look at the pediatric norms published for that age child, there could be huge discrepancies. And so depending on your reference point, these percentages have some or no value.

Presentation – Denham Ward

DR. WARD: So most people that are adults

Page 51

were still less than 40, would be an absolute change, entire cardiac greater than 120. So both outside, the usual normal range of 60 to 100.

Okay. So that was the opinions that people had. And again, I put those out more to help people think a little bit and be a little controversial than to really collect anything that we were going to go from.

Could I have my slides? So the rest of the morning we're going to spend talking about more specific measures. We've gone through the very high level yesterday, and we're going to get down into what were the kind of things -- if you're designing a clinical trial for a new sedative, what were the things that you'd want to measure?

Respiratory, so I allotted more time to respiratory than the other areas because that's clearly the one that we have spent the most time on. And I'll talk some about it, and Albert will talk some about it, too.

Unfortunately, I thought I'd have to go back and do a little respiratory physiology with

Page 52

everybody because there was a few missed concepts that I was hearing yesterday.

Not to get everybody too upset first thing in the morning here, but to remind everybody that the respiratory system and control of breathing is an incredibly complex, interactive, closed feedback loop system.

It goes well beyond our medical school understanding of essential chemoreceptor, carotid bodies, brain stem, airway muscles. The ones in blue are the traditional what we were taught in medical school. And we know now that that's much too simplistic.

Not only do we have cortex and subcortical, cortextual, with a limbic system that affects breathing, but we have to take into account the spinal and hypoglossal motor neurons with the upper airway and chest wall components to maintain ventilation. And obviously, there's a potential for drug effects up and down this system in a myriad of different ways.

I thought for our purposes a little better

Page 53

way of thinking about it for procedural sedation is to think about ventilatory control coming from three areas. There's the traditional metabolic control, the CO2/O2 chemoreceptors that when your O2 goes down or your CO2 is up, or your pH is low, there's stimulation through the closed metabolic loops for the chemoreceptors.

Obviously, voluntary control. What I'm doing right now illustrates voluntary controls, that I can manage to not only walk and chew gum at the same time, but I can breathe and talk at the same time. And I can do it fairly efficiently, without very much disturbance of my blood gases. If you measured my blood gases right now, they'd be essentially normal. My CO2 might be a couple millimeters low. So I got voluntary control.

But probably the most important, if I'm not talking, probably the most important control is behavioral, wakefulness control, that in spite of being taught in medical school that it's the intersection of the CO2 response and the metabolic hyperbola that controls our resting ventilation,

Page 54

that's probably not true. What's controlling my resting ventilation is what Ray Fink termed back in Columbia in the 1960s as behavioral wakefulness drive.

So in the normal resting and maybe during light exercise, this mixed control really controls all three, but the primary effect is a behavioral control. And obviously, the procedural drugs that we use can have effects on any of those three areas, and they do.

Examples to start off with is with opioids that severely depresses the hypercapneic and hypoxic response, may reduce their wakefulness control a lot, but voluntary control is still intact. You can ask an apneic narcotized patient to take a deep breath, and they can still take a deep breath for you.

So I want to divide up the idea of testing sedative drugs into two areas, with an overlap here: The clinical, basically monitoring, as opposed to a laboratory, more provocative. And we've talked a lot about patient and procedural

Page 55

selection.

So monitoring ventilation is something that we can do clinically. Methodologically, that's not necessarily the easiest thing to do. We've heard a little bit about CO2 monitoring. Albert yesterday mentioned humidity monitoring. Thermistors are used by the pulmonologist to measure air flow. In the laboratory, there's ways with pneumotachographs and face masks and mouthpieces and things that you can measure it accurately.

In the clinical situation, measuring ventilation, tidal volume, and breathing rate is not a trivial task. Measuring apneas and differentiating between central and peripheral apneas, measuring the CO2, measuring the saturation, depending on whether or not you have supplemental O2, and recalling without airway obstruction with a patient that has supplemental O2, saturation is going to be a very poor monitor.

In fact, a concept called apneic oxygenation, if you've got an open airway and you're on 100 percent O2, you will never get

Page 56

hypoxic because sufficient gas will flow because of the gas exchange ratio.

Then the apnea-hypoxic, hypercapneic index, so the respiratory disturbance index, looking at ventilatory arrhythmias; and then the need for clinical intervention. These can all be done just monitoring a patient without any provocative intervention.

In the laboratory, and these are generally more phase 1 type trials, hypercapneic ventilatory response, so you can give patients a traditional CO2 rebreathing study, where you can have subjects rebreathe CO2 and look at the slope and intercept; hypoxic response. The pulmonologists look at negative airway pressure, called Pcrit, negative pressure that causes the airway to obstruct; and you can find that between normal sleep, general anesthesia, and sedation, there's a change in what that Pcrit or negative airway pressure could be.

Pain arousal, audiovisual stimulation, and exercise airway load, all those factors change not only the breathing but change the effect of a drug

Page 57

or a sedative on breathing.
Other measurements in the laboratory were the EEG or EMG for the genioglossus muscle, esophageal pressure. And then finally, most of the time it's not a single drug, so how are you going to measure the drug interaction between two drugs, looking at PK/PD or response surface kinds of modeling.
So we've got lots of options. If you've got a brand new drug and you say, I want to characterize this drug's respiratory effects, we've got lots of options, both early at a phase 1/phase 2 trials and maybe phase 2 to 4 kind of trials.
So a clinical intervention index, something that we put together, not validated, but one that we wanted to have. But it's based on other studies. We used this in a study that we did in GI, that Suzie did, looking basically at a clinical intervention score.
It's based on the kinds of desaturations, things that we talked about, the need for a chin

Page 58

lift, desaturation of a greater amount, so some sort of scoring system. That's been mentioned a couple times; there are a lot of areas for intervention; pancreatitis.
We have scoring systems. We don't really have any validated scoring -- any that I'm aware of -- validated scoring systems for a clinical intervention need in sedation. So we kind of made this one up and used it in a study that we did. One of the things we could do is to go back and do some validation of this study.
If you're just monitoring a patient, this was a study quite a while ago that I did in the laboratory so I can measure everything accurately. So this was at 5 minutes after we gave an IV infusion of the drug, and we see tidal volumes and ventilation, and we see saturation.
This is a laboratory oximeter, so you see a lot more variations. Most of your clinical oximeters average over 3 to 10 beats, so give quite a bit of average signal. This is an un-averaged signal, so it shows more variation in the

Page 59

saturation than perhaps you're used to seeing.
Then this is what's known as a RespiTrace, so it has bands around the abdomen and the chest wall, that measures the excursions of both. So normally they're in sequence, and as there's degrees of obstruction, they become out of sequence. So here we've got three different ways of measuring a respiratory effect, a tidal volume measurement, saturation measurement, and a synchrony between the abdomen and rib cage.
Here we see, just monitoring the patient, that we have some apneic episodes here, with a subsequent time delayed decrease in saturation there. And it's a little hard to tell without blowing this up whether this is a central or a peripheral apnea.
Anybody want to guess what drug this is?
(No response.)
DR. WARD: Nobody wants to guess what drug, what sedative this is?
DR. RIKER: Midazolam?
DR. WARD: No.

Page 60

AUDIENCE MEMBER: Opioid?
DR. WARD: No.
AUDIENCE MEMBER: Dexmed?
DR. WARD: Dex. This is dexmedetomidine. This is 2 micrograms per kilogram given over 2 minutes. It's a pretty high dose given pretty fast. This was a phase 1 trial that I did in dexmedetomidine.
So even though we say dex doesn't have any respiratory effects, if you actually have a high enough dose given fast enough -- I'm going to show you what the heart rate did in that study -- there are respiratory effects. And this is probably an obstructive episode.
So not an easy thing to do. But we have accurate measurements of how long the apnea was, what the saturation was, and what the airway was.
So this is a blow-up of a different drug. This is propofol. So this is now the flow, so here's inhale/exhale, inhale/exhale, inhale/exhale. And here's the RespiTrace. So here we've overlaid the abdomen and the chest wall tracings.

Page 61

We see this patient, or subject, was going along breathing quite nicely, and then had an apneic episode. But you see it's a central apneic episode, that the tidal volumes were decreased, but they stayed in synchrony. So we can differentiate a central and an apneic episode. This I think it was with midazolam and remi.

The central episode came on for a number of seconds. So we can actually quantify the apneic episode. And as we've pointed out, I think we're relying on clinical observation for length of an apneic episode as pretty subjective.

But an interesting thing on this particular one was when we started to get a respiratory arousal, we started out obstructed. So we started out with the abdomen and chest wall out of sequence for three breaths, and then the patient cleared the obstruction, started getting an arousal, and started breathing again with airflow and synchrony.

Or you can do this even more so. This is from one of the studies that Dr. Karan did, where you can look at a full montage EEG, a thermistor

Page 62

measurement of flow at the mouth, and the saturation, and the respiratory thorax and abdominal movement.

He had no flow with this. You can clearly see they're out of sequence, with the thorax going in and the abdomen going out, thorax going out, abdomen going in, and the saturation going down, until we get an arousal saturation, and then it comes back up again. And you can see the spontaneous breathing now back in synchrony, with a resumption of flow, as measured by the thermistor. And you'll see an arousal event in the EEG.

Obviously, this would be a phase 1 type trial, but just to illustrate the kinds of exact measurements that you can make for respiration, and to differentiate them from the kind of very inexact, more subjective, the patient a 30-second apneic episode that you'd try to get in a clinical trial.

So for provocative tests -- Albert said these are hard to do; they're not that hard to do; they're hard to do well -- you can look at

Page 63

increasing CO2 with a change in the slope and intercept of the hypercapneic response, and you can decrease the PO2 and look at a linear increase in ventilation with a decrease in saturation. Each dot's a breath to breath, with pretty good correlation.

Sedative and analgesic drugs have been well classified with these. You can go back to Santiago and Edelman, but maybe even before that, 0.2 milligrams of morphine IM and looking at the shift in the -- and the decrease in slope of the hypercapneic response, and the decrease in slope of the CO2 response. Ventilation went down. Tidal volume went down. Respiratory rate went down just a little bit.

These are pretty substantial shifts. But I don't think anybody would say that 0.2 milligrams per kilogram IM morphine in an adult is a dose that's going to be particularly a problem for multiple adverse events. So these are very sensitive measures of the respiratory system, and they're done almost invariably with any new

Page 64

sedative analgesic drug.

But we don't have any correlation with how much this changes to how much that decreases ventilation, and again, that's because this is a metabolic controller, and you've probably left the wakefulness drive and the voluntary controllers intact.

My mentor, though, a long time ago, Jay Bellville, in a classic article with four subjects that would never get published today, pointed out that if you combine morphine plus sleep, you get a tremendous increase in the respiratory depression because now you've taken away the metabolic drive with the opioids, and now you've taken away the wakefulness drive with sleep. And you get an even bigger decrease. And he pointed out then that the altered state of consciousness may be more important than the particular drug effect on the metabolic chemoreceptors.

We've talked about patient state a lot, and arousal and pain makes a big difference. The study that Suzie did looking at hypoxic response with

Page 65

remifentanil. Right, Suzie? Yes, with remifentanil, and looking at the difference between resting, eyes closed, pain, or audiovisual stimulation.

We found different effects depending on whether the subject was laying there with their eyes closed, they had a painful stimulus, or they had audiovisual stimulation, with interesting audiovisual stimulation being more than a painful stimulation.

We looked at midazolam and propofol on collapsibility using dynamic negative airway pressure. In that case, we were actually lowering the pressure at the mouth and seeing what that did to ventilation, so lowering this down to minus 4. The flow was decreased, and we had partial obstruction. We started to see some out of synchrony there.

Lower the pressure even more, we started to see a further decrease in the flow and more dyssynchrony in the ventilation with a partial airway obstruction. And finally, lowering this

Page 66

down to minus 10, we saw complete obstruction and complete asynchrony in the thorax and the chest.

What the pulmonologists then do is then plot this, or with each of the pressures, and measure where the airway-closing pressure would be. And there is work in general anesthesia done by Eastman in Australia that looks at that critical closing pressure in general anesthesia as a propensity for airway obstruction. So does a drug that has a more or less negative Pcrit -- in a sleep apnea patient, this Pcrit's positive, and that's why CPAP works. They actually obstruct with a positive pressure.

So we talked about midazolam a little bit yesterday as an illustration of a drug that gets approved and then has difficulty after it's approved; and a study that Peter Bailey did years ago looking at midazolam, fentanyl, and fentanyl plus midazolam.

This is a ventilation of CO2 at 50. It raised the CO2 to 50 to see how much the ventilation is, so 35 liters a minute with an end tidal CO2 of 50. And with midazolam, there was no

Page 67

change. Ventilation started recovering here about an hour later, so that would be consistent with the pharmacokinetics. Midazolam by itself, and there were no desaturations.

Fentanyl by itself, 2 mics per kilo, caused a big decrease in ventilation, as you'd expect, and 6 out of 12 of the subjects desaturated. Adding midazolam to it, there was no further decrease in ventilation. So there was no synergy between fentanyl and midazolam as far as the ventilation to hypercapnia was concerned. But now 11 out of 12 got hypoxic. So there was much bigger instances of apneic episodes. When they were just laying there breathing normally as opposed to being stimulated by hypercapnia, that would cause desaturations for that.

So again, illustration that the laboratory measurements of ventilation in response to hypercapnia is not necessarily a good predictor of what the apneic episodes would be.

Albert's lab has done a lot of work looking at multiple drugs. This was a study that I did

Page 68

with him when I was on a sabbatical there, looking at remi/propofol interaction and ventilation with the function of CO2 for control, propofol decreasing the slope, remi-fentanyl causing a shift, and the two of them together causing a big decrease, both in slope and a right shift, the synergism between the drugs.

Then Albert set up a really nice system of looking at the two drugs together, with a response surface. So this is resting ventilation, so about 10 liters a minute when we don't have either drug. And then what the surface looks like as we increase propofol would be this way. And as we increase remifentanil, it would be this way.

Then as the two go together, we see a marked synergism between the two drugs, between the remi and the propofol, on resting ventilation, on the end tidal CO2 -- the scale is reversed, so the end tidal CO2 with the two drugs together. You've got up to 55, a little less synergism, interestingly there, a little flatter surface. Ventilation at a high level of CO2. A lot of synergism there. And

Page 69

on the bis, no synergism at all. The remi didn't have any effect on the bis.

Albert, do you remember the OAAS? Did we measure that? I think we just had the bis, just had the bis data for that, because the bis is not sensitive for the remifentanil.

So finally, as we think about respiratory measurements -- it's complex, I guess would be the answer to that, but I think it's useful when they're doing some of the trials to think about these three drives to ventilation, voluntary, behavioral, and metabolic.

We've got lots of changes between them, between -- I'm not going to go through them -- anaerobic exercise, which is low pH; it's probably just metabolic control; REM sleep -- non-REM sleep is probably just metabolic control, whereas REM sleep is more like a wakefulness drive, and there's not much metabolic control during REM sleep. Singing and talking is all voluntary.

Then looking at the drug effects, then you have to think about the particular kind of sedation

Page 70

protocol you're doing. Is there a painful stimulus? Is it an x-ray procedure, which there's no other stimuli?

So I think there's questions that we've been coming up with throughout the days here that I'll end with, what are the criteria for a respiratory adverse event? Are there any provocative tests that we would recommend doing: hypercapneic, hypoxic, Pcrit? Should these provocative tests be encouraged in the early, phase 1, clinical trials?

High-risk patients we've talked about. Should clinical trials be designed for AEs as noninferiority trials compared to what regimen? As Dan has pointed out, the current regimen is not necessarily a good comparison.

So I'll just finish with Rich's comment, that, "Breathing is truly a strange phenomenon of life, caught between the conscious and unconscious, and peculiarly sensitive to both." Which is what makes it so hard to design a clinical trial that looks at respiratory effects, which for a sedative, procedural study is the most important outcome.

Page 71

So I'll leave it open to your comments on these questions.

DR. RIKER: It was a nice review. Thanks. Should we move beyond ventilatory impairment, though, and talk about the respiratory events that could happen, aspiration with pneumonitis, with pneumonia, with neither, which half of the patients who aspirate will fall into? Does that precipitate airways disease? And I don't mean upper airway obstruction; I mean, small airways disease, wheezing, bronchospasm, blah blah blah?

DR. WARD: Yes. Those are important clinical outcomes. Again, those are going to be the rare events and clearly should be reported. Because I'm not sure if it's moving beyond this. Those are clearly in addition.

DR. RIKER: I just think of them as respiratory events but not related necessarily to ventilatory drives.

DR. WARD: Yes. Yes, yes. Clearly should be looking for and reporting non-respiratory drive-related events. I think what I'm hung up on, and

Page 72

I've been hearing it, too, is looking at just saturation is probably not an adequate measure.

So if we agree on that, then what are the things that we should be measuring? Should we be measuring things like thorax and abdomen excursion? Should we be trying to measure things more detailed than that?

I don't like the idea of measuring apnea by having an observer decide when there's a 20-second or 30-second apnea. And maybe Albert could comment on the thermistor/humidity sensor and its reliability and ease in use.

DR. DAHAN: In my opinion, mechanistic studies are almost impossible to perform clinically. Those are my problems, very complex studies. I'll show some examples later.

Real-life studies using the patient studies try to be as simple as possible, reliable data, easy data acquisition and reliable data acquisition. The more complex you make the studies, the complexer [sic] the -- also the increase in the possibility of mistakes and loss of

Page 73

data.
So it is a relevant issue, and your examples are exquisite examples of mechanistic studies to understand if you have a new drug, is it better than the previous drug. That's my opinion about it, and I'll talk a little bit on how I see how you should [inaudible]. I do agree outcome, of course, is most important. Yes.
DR. WARD: Jenifer?
DR. LIGHTDALE: I think drugs get a reputation. I am really fascinated by the concept. Some drugs, you'd say, they don't have many respiratory effects. Maybe we do need to do these provocative tests early on to really make sure we have the right reputation for the drug. People have a sense of, well, actually, you can push this drug, and it will do something. We need to know that.
DR. WARD: Yes. I think that's what dexmedetomidine is. The early studies including mine, in spite of what I showed you there, did have minimal respiratory -- I mean, if you didn't push

Page 74

2 mics per kilo over 2 minutes, it didn't have much for respiratory effects.
Suzie and then Dan.
DR. KARAN: With the sophisticated respiratory ventilation and monitoring that we're doing now clinically, I think there's still a signal out there in some of the clinical trials that we're not picking up, specifically use or peak pressures that are being captured. You interrogate SIM cards on CPAP machines, and people are on opioids, or even in the hospital or for certain ventilation.
So I think there's still a signal out there that maybe you don't have to be provocative about it, but you can still capture it. And we're not capturing these very sophisticated measures of ventilation out there. It's still a possibility.
DR. WARD: Dan?
DR. SESSLER: Do we have any kind of consensus that ventilation should be included in trials of sedation?
(Laughter.)

Page 75

DR. KARAN: Do you need a show of hands?
DR. SESSLER: So we won't specify the type because it might be context-dependent, or is sometimes transcutaneous, sometimes end tidal, sometimes bands around the chest, sometimes impedance. But there probably should be some measure of ventilation, because otherwise you miss a lot, especially in patients on oxygen if you're just looking at saturation.
DR. DAHAN: Personally, I would go for flow rather than ventilation.
DR. SESSLER: Okay. I would accept flow as a type of ventilation.
DR. DAHAN: Well, it's close to passage of air.
DR. SESSLER: I'm saying something besides saturation.
DR. WARD: Rick?
DR. RIKER: In the regulatory area, there's a move towards risk-based monitoring. So maybe we should apply a similar model here. So if we're looking at low-risk patients, maybe it's not a

Page 76

critical piece of it. But if we're certainly looking at these sleep apnea, obese patients who are at much higher risk for these kinds of events, then we need more sophisticated monitoring for those patients.
DR. SESSLER: But all of these proxies can induce ventilatory problems. I would say even if the patient's not at high risk initially --
DR. RIKER: The procedure may be.
DR. SESSLER: -- that [inaudible].
DR. DAHAN: I recently did a study in which I compared the effect of oxycodone in elderly and younger subjects. What I had not expected is, especially in the elderly, severe respiratory depression of their health [indiscernible}, patients above 65. Most of my patients are, in fact, elderly. But everybody's at a risk of some sort.
DR. MINER: That brings us to a good point, where you talk about where we should put sicker patients in the trials, I think they're a lot easier to study because they're much more sensitive

Page 77

to the adverse effects, especially ventilatory ones. The sicker the patient is, the higher the rate, the easier it is to compare agents and find a difference that's meaningful.

DR. WARD: You're correct, Albert, that some of these more sophisticated tools are difficult to use. But I look at our pulmonology colleagues doing sleep studies who have a pretty sophisticated array of instrumentation that they do hundreds and thousands of patients every night in sleep. And if you can do it in a polysomnographic laboratory, why couldn't you -- it may be provocative -- why shouldn't you be doing it in a clinical trial in a few hundred patients?

DR. DAHAN: You can. It's expensive. It requires a lot of teaching.

DR. KARAN: It's cheaper now. The ambulatory monitors are a little bit cheaper now. They keep evolving. Every month there's newer, more robust monitors out there that are in the ambulatory environment that are meant to be robust to the surrounding, and that are applied just by

Page 78

patients themselves without a lot of teaching. So it's evolving.

DR. WARD: Dan?

DR. SESSLER: The general rule for clinical trials is to use the best available methods to measure whatever you're trying to measure. And the fact that it's expensive or difficult only gets you so far. If you're doing a clinical trial, that's inherently a difficult and expensive process. You should use the best metrics you can.

DR. SEXTON: As an anesthesiologist, what are the things that let you know, if you're doing a case, before they start to desaturate? What tips you off?

DR. WARD: Well, I don't know. I'll leave it to the other anesthesiologists in the group. But to me, it's watching the patient and the signs and symptoms of airway obstruction.

DR. DAHAN: I have some slides that show irregular breathing.

DR. WARD: Irregular breathing.

DR. DAHAN: I can predict when somebody will

Page 79

have an apneic peak.

DR. SEXTON: Because it seems like that would be something you'd want to measure so you don't have to wait till they desaturate, and they have already have some options.

DR. WARD: And the abdomen and thorax measurements are the kinds of measurements that let you do that. So I think Dan's point was a good one, is that if you're doing clinical trials, you should be -- first we have said we need to measure ventilation, that just maybe saturation is not sufficient. And then it was brought up that if you're going to do a clinical trial, then you should be using the most sophisticated way to measure what you want to measure as you can. So that moves us into things like abdomen and thoracic bands.

Bill?

DR. CHAPPELL: I have a comment. I would imagine -- I could be wrong -- but I think that if you were developing a new chemical entity, that types of provocative and mechanistic physiologic

Page 80

studies with these very intensive measurements you described would probably be best situated maybe in early development.

DR. WARD: Phase 1.

DR. CHAPPELL: Then as you push the compound farther down into development, you're going to have to do larger scale trials, and you'll have to move to simpler designs and simpler data, as Albert was commenting. You need both, really, depending on the development stage of the project.

DR. WARD: Yes. But I think the preclinical trials in animals on respiratory effects don't correlate very well with the early clinical trials of respiratory effects. Opioids are an example in cats. They don't correlate very well with what happens in humans. So some of the preclinical data needs to be repeated, and physiological data needs to be repeated in humans.

So I'd think, paraphrasing Dan and Albert, and a few others, is it sort of the consensus here -- you've got a good point, Phil, that the level of sophistication of measurement can decrease

Page 81

as the size of the trial -- and as we move further along. But the concept would be that as we start out, we want to use better measurement devices and not just rely on a clinical trial that has saturation and clinical observation of respiratory rate. It would not be adequate for a clinical trial.

Is that kind of -- and Mark, you said you saw a lot of those kind of trials in the literature.

MR. WILLIAMS: That's right, yes.

DR. WARD: So I think we're recommending something that's different than the current things that are being done. So I think that's an important outcome for this.

DR. GREEN: And I agree with that. But we have to, in the end, recommend things that actually can be studied in -- these are not exactly real world because they'll be held to a higher standard of measurement as we're looking at this, but things we can apply in MRI in children, things we can apply to that as to how we measure those.

Page 82

Because most of these sedations we're doing in kids, and then we have to get through an IRB, and efficiencies in hospitals, and say we're going to do this, but you can only do two MRIs a day because of the level of monitoring; who are the funding to -- not just the funding for that patient, but the lost income for institutions who are asking to do these studies.

DR. WARD: Right. And I think that's a great point and couples with what Phil said, is that the studies have to be designed such that as you move further along in the drug development, your monitoring doesn't have to be as sophisticated as it was earlier on. I think that's an important concept.

DR. CHAPPELL: You mentioned sleep studies earlier, and that is exactly the paradigm for the development of hypnotic medications. As you probably know, the FDA typically requires sleep laboratory studies, but then you move beyond that and do outpatient studies where you're not really doing sleep EEG work.

Page 83

DR. WARD: Right. Early on, if you've got a new sedative that you're looking for approval for it as a sleep medicine, then it's very much analogous to the way you progress.

Bob?

DR. DWORKIN: The other analogy, it seems to me, [inaudible – off mic] can show that there's a pretty standard phase 2 design where you study that drug against positive comparators in recreational substance abusers, and see if they have great drug liking or that they would pay a lot for it on the street.

That's a kind of intense, small phase 2 design, costly, in highly selected patients. And you do that before you study your drug in the phase 3 trial in patients with low back pain or whatever. And that seems analogous to me, too, here, a targeted phase 2 trial in a group of subjects that you expect to be informative, but that doesn't necessarily generalize to the ultimate population.

Rigo, does that seem analogous to you, that

Page 84

kind of intense phase 2, get to know your drug?

DR. ROCA: Yes, I think so. And I think -- which is another comment I think you guys were talking about, that will also help you design your phase 3 because you can identify certain things that you would like to focus on. And at the same time, you may be able to identify certain things that you don't need to focus on. But yes, your example is very analogous.

DR. WARD: Any other comments on the respiratory?

(No response.)

DR. WARD: And sorry to bore you with all the physiology, but I thought we did have to have a little physiology lesson before we got a reality check on how difficult it is to actually make some of these physiological measurements if we want to move beyond simple saturation and clinical observation.

So let's take a break. I think we're just a little bit ahead of schedule. So let's come back at about 10:00, and we can keep the day moving and

Page 85

get through in time for us to all make our flights this afternoon. Thanks.

(Whereupon, at 9:320 a.m., a brief recess was taken.)

DR. WARD: So there's more to life than breathing. We have to eat and do some other things, so the complications are not just respiratory complications. There's also other organ system complications. So I thought we'd spend a few minutes on looking at some of the other possible complications, both as I went through what to measure and how to measure them. So we will start out with TJ Gan on GI adverse events.

Presentation – TJ Gan

DR. GAN: Good morning, everyone. First of all, I want to thank Denham for inviting me to this -- I think this event is great. And those of you yesterday and today, I think you don't get a lot of opportunity where you can have multi-specialty people having a common interest come together in a room to discuss something that we're all passionate about. And I'm sorry I missed the

Page 86

first round, so I'm delighted that I'm here.

I'm asked to talk about GI events, what and how to measure. I'm going to just really show you a few slides and just to set the scene. And again, you are not here to come and listen to me talk, but it's more about discussion and what we can get out of it. So I'm just going to quickly go through some slides, give you some perspective as I see it, and then have lots of discussion.

So as far as GI events are concerned, I am an anesthesiologist, so the GI events that I am accustomed to may be a little bit different from the GI events gastroenterologists are accustomed to, emergency medicine accustomed to.

I'm just going to put out a list of what hospital GI events could be in clinical trials based on adverse events reporting, looking at drugs, and again, thinking about GI events. And if you think about from the mouth all the way down, the esophagus, the stomach, small intestine, large intestine, all the way to the other end, you can probably pretty capture some of these events that

Page 87

happened.

This is sort of a list of what we typically commonly see in a clinical trial setting. But talking about adverse events, I listed adverse events, presumably, these are what people don't like. Does anyone here like vomiting? I hope not.

(Laughter.)

DR. GAN: You may laugh, but sometimes it's surprising that some people actually enjoy vomiting. And this was a description of somebody who actually derived pleasure from vomiting. So while we called it adverse, it may not be adverse to some people. So it's all relative.

So anyway, we're not going to go there. Let's talk a little bit about this. This is a slide that is not to be included in the public domain, I hope.

Anyway, so nausea and vomiting, I'm just going to concentrate on one that I think happens particularly frequently, and one that we sort of think that patients really don't like and hopefully can reduce it.

Page 88

This is one of them. It's one of the GI events, nausea and vomiting. It happens commonly. And this is actually a 10-center observational study, just look at across the board, what is insulin nausea and vomiting.

Again, these are people having general anesthesia, and I'll show you some data on people under sedation. But you can see here, a common practice is that we give a lot of prophylactics, antiemetics, right, either one, and probably more appropriately, two or three, depending on their risk factors.

But nevertheless, if you see that these are one antiemetics two or three, nausea, about 30 to 50 percent of people with nausea, vomiting about 10 to 30 percent, so it still happens commonly. And about a third of these patients say nausea and vomiting interfere with their functional recovery. And I guess this is probably more important, that these are people sort of bothered by nausea and vomiting.

Some people may say, "I'm a little bit

Page 89

queasy, one or two on a scale of zero to 10. I don't like it, but it's not going to kill me," for example. But this is where I think we need to pay attention, that these are people who say that it does interfere with my functional recovery.

Now, as far as the incidence of nausea and vomiting following sedation, this was a study that we reported a few years ago, looking at patients undergoing upper and lower GI endoscopy. And the intent of the study was just looking at the depth of sedation, whether it has any implication in terms of adverse events.

We recorded a number of adverse events to include respiratory, apnea, saturation, all the rest of it. And we found an incidence of about 5 to 10 percent sort of nausea, including all-comers. And vomiting is pretty rare, in the region of 2 or 3 percent. And these again are fit, healthy people undergoing routine endoscopy, and it comes in. They go out. Again, these are patients who use midazolam fentanyl. This is not propofol sedation.

Now, let's talk a little bit about nausea

Page 90

and vomiting. How do we measure it? And again, adverse events reporting and how we normally measure it when we look at efficacy is a little bit different. Right?

But yesterday, as you heard, there is some overlap between efficacy and adverse events reporting. But this is typically -- when we are to do a nausea and vomiting study, this is typically the outcome. If you have a phase 3 study, this is typically what the FDA people will tell you these are the outcomes.

So I want to look at nausea, which can be assessed using a variety of scales. Vomiting is another outcome. Incidence, typically, severity is a little bit difficult to quantify. How do you quantify severity of vomiting? Retching is typically lumped together with vomiting. It just means that you have nothing to -- there is no content that is expelled.

Often in clinical trials, the use of rescue antiemetic is used as a surrogate for nausea, because nausea is so subjective. And the thought

Page 91

is that, if nausea at a certain point is bad enough, then you ask for something. And that something that you ask for, an antiemetic, is a proxy for sort of bad nausea.

Then there are other things that are less commonly mandated by the FDA, but increasingly being asked to be collected by sponsors who want to do a phase 3 trial interference with daily activities as well as patient satisfaction, again recognizing the limitations of some of these scoring because patient satisfaction sometimes is not well validated in terms of who say their satisfaction.

In fact, people could be sick, and if a nurse is there, really be with them all the time and console them, they were happy even though they were sick, just like in other pain studies as well.

So just to quickly run through some of the scales here that have been used, this is typically called a verbal rating score, no nausea, 1N, and then 1 to 10 was nausea. And then for children, as Denham earlier brought up, this is called the BARF

Page 92

scale.

(Laughter.)

DR. GAN: It actually stands for Baxter Animated Retching Facial Scale. And hopefully, when you're a kid, you can pick up which are the faces that you would like to be associated with. So these are typically for children.

Then there is the VAS, Visual Analog Scale, which is basically one line without these descriptors. And then there is also the Likert scale that basically have descriptors such as slight, mild, moderate, and severe.

So there are a variety of scales that are being used and have been validated in many studies. As to whether how valid this is, as we talked about yesterday, is it linear? Probably not, but we don't really have a better scale than this to recommend to investigators.

If you look into the literature, I mean, these are pretty standard. And I would say that if somebody had asked me, if you wanted to do a nausea and vomiting study, what scale should I use, I

Page 93

would say that, for adults, it would be either VAS or VRS scale for nausea, vomiting, incidence of vomiting, and collect use of rescue antiemetics.

That's what I would recommend. But if you go to the literature and look at what people use, it's really all over the place. I mean, there are people who use this scale. How do you interpret that? I can see that -- I mean, it's all over the place. And so sometimes, it's very difficult to interpret.

Then there is people using the Likert scale. And then you've got a vomiting score, how many episodes within a certain time to indicate some sort of severity of vomiting, not as well validated. And there are also people using this sort of grade 1 to 5. And I was intrigued to find out, in fact, if you got vomiting, grade 5, that means death, which sounds a bit drastic. I haven't seen anyone sort of vomiting until death, but anyway, you can see that people use all sorts of scales.

This is one of the chemotherapy ones, again,

Page 94

from no nausea, severe nausea, and also look a little bit about daily interference of their activity. So you can see it's all over the place in terms of scoring is concerned.

Now, the other question that got raised yesterday is that collecting this incidence and subjective experience is one thing, but maybe we should look at how much does it bother the patients. I mean, some people may be bothered by it, but some people may not, not that I think people enjoy nausea, but as you can see earlier, people may enjoy these adverse events.

But the question is how frequently does it happen, how severe it is, and also how bothersome to that individual it is, to the patient. So a few years ago, we sort of validated this symptom distress score, which is a bondification from the Portenoy Pain Score.

This is as a result of doing some pain study. And part of it is opiate-related side effects. And I've put this up just to give you perhaps an idea that some of these could be

Page 95

potentially scoring in nausea and vomiting, looking at really three aspects of it. One is looking at the frequency, how frequent and how often does it happen, rare or very common, or how severe it is from mild to very severe, and how bothersome to them. Does it bother a lot or is it really just very little?

So maybe they're using this sort of composite scoring that we can make a better assessment in terms of these qualitative experiences of these adverse events. In fact, we actually correlated with the opiate doses and found there's a good correlation in terms of these scoring, the symptom distress score. The more severe it is, the higher opiate doses is associated with higher symptom distress scoring.

So I'm going to stop there, and I don't know whether we then want to open the questions or we want to wait for the rest.

DR. WARD: Thank you. We'll take a few questions now. I would comment that, 100 years ago, people did die of vomiting making oceanic

Page 96

passes with seasickness. So it used to be possible, but hopefully nobody dies from seasickness today.

DR. GAN: Also, in the days when ether anesthesia was a regular anesthetic -- and again, I think everyone in the room may be too young for that, but ether anesthesia, you pretty much guarantee everyone gets sick. For 89 percent of patients, they were sick from ether anesthesia.

Any other questions, comments, thoughts? Yes?

DR. MINER: For the visual analog scores for nausea, I always try to use those, but I always seem to end up with binary data, where everybody says, yeah, I'm on one side or I'm on the other, and there'll be like one person at 50. I don't know [indiscernible] any more.

Does anyone else find that? I think it's been hard to get it distributed score.

DR. GAN: What I find is the Visual Analog Score is sometimes a bit difficult because patients are sometimes sedated. And when you present a line

Page 97

when you're sedated, a line may become two or may become 3, and you just couldn't quite focus. Usually, in those settings, I find that a verbal rating score seems to capture a better response, from zero to 10.

As to your point about whether people cluster it on the left or the right, I've looked at that. In fact, that's what we commonly thought. But if you actually look at all the studies, it's actually pretty well -- I mean, if a patient is awake enough to interpret it, it's actually not a bad score provided a patient actually understands it.

I think, sometimes, it's difficult to explain it to the patient. That's why, in the clinical trial setting, we try to explain to them before we actually do the procedure, so that they somewhat understand because, when you're trying to explain to them when you are sedated, groggy, I think it is impossible.

DR. WARD: It sounds like a somewhat similar issue that we had at the first meeting, looking at

Page 98

the number of validated sedation efficacy scales, the number of validated nausea scales. Is there any consensus or is it possible to reach a consensus on recommending one of those?

DR. GAN: Yes. I would say that if you look over the trials over the past 10 years or so, the scoring system would be either VAS or, more commonly, VRS for nausea. Vomiting would be incidence. And then the use of rescue antiemetic would be another one, which often sometimes serve as a proxy for severe nausea.

Those are pretty standard, accepted by the FDA. I don't think we have FDA colleagues here, but that would typically be supported by the -- we do have FDA colleagues. There you are.

What do you think, any thoughts on nausea and vomiting scale for doing puking studies?

(Laughter.)

DR. ROCA: I think you've hit it on the nose in the context that there's so many of them. And depending on the study, the patient population center, some may be more appropriate than others.

Page 99

But I think we probably all would like to have whatever scales to be used to have some data to show that it is a validated scale for that particular study.

DR. GAN: Right. And I think those I'd describe typically are the most commonly used, so the FDA doesn't raise questions about validity.

DR. DWORKIN: So in the pain world, it's been shown, going back a couple of decades, the 10-centimeter visual analog scales are problematic with older patients and patients with mild cognitive impairment. So given that older patients and patients with MCI can be in a clinical trial, we almost always prefer a zero to 10 scale rather than a 10-centimeter line. And perhaps especially when they're also sedated, that might push towards the numerical rating scale or the verbal scale.

DR. GAN: I totally agree. I think that was what my recommendation is, would be zero to 10, which would make it much easier. And while we have done both and, often, I get complaints from my coordinator or nurses saying that the patient can't

Page 100

do it. And when you ask them zero to 10, then usually they can. Absolutely.

DR. WARD: Maybe even the faces scale, even for adults, might be useful.

DR. GAN: Interesting, but that has rarely been tried. I think it's mainly in pediatrics.

DR. LIGHTDALE: It strikes me that you also need to think about prolonged nausea, so how long does the nausea last, and then the severity, and they agree are two different things. So if it's bad nausea and it's terrible right at the beginning, that could be a problem.

If it prolongs your recovery so you're sitting there for a while, that could be a problem. But then also not such severe nausea, just that first queasy face, but you're still feeling that two days later, that would be clinically relevant.

DR. GAN: No. I totally agree with you. And, in fact, that prolonged nausea, some studies actually collect that information. But as I mentioned earlier, usually the time to rescue would be taken as surrogate as the duration of nausea.

Page 101

The thinking is that, if the patient is bad enough, at some point they're going to ask for rescue antiemetics. It could be vomiting. But for those who didn't have vomiting, that is typically as a proxy for the duration of nausea. But some studies do collect the duration of nausea.

You're right. I think if people just keep getting sick -- and I happen to have a firsthand experience -- it's just the worst experience you can ever imagine.

Yes?

DR. RIKER: Did you find anything on surrogate assessment versus patient self-report? I mean, in pain, there's clearly a lot of data that says surrogates are horrible at estimating pain. Is there similar data in nausea and vomiting? I mean, vomiting is pretty easy.

DR. GAN: When you say surrogate, can you define which surrogate are you thinking of?

DR. RIKER: Well, a clinician, for instance, estimating the degree of nausea based on that.

DR. GAN: Yes. You can't. I don't think

Page 102

the clinician can assess in any way, shape, or form accurate in terms of patient's nausea.

DR. RIKER: All right. So outside the setting of young children, then, we would say the expectation should be patient self-report.

DR. GAN: It's for patient, yes. Absolutely.

Yes?

DR. GREEN: I'll just chime in again on the concept of interventions. There's going to be patient variability in how they rate their nausea, but yet the act of somebody needing to give an antiemetic is kind of an objective and tangible measure of the clinical importance of the symptoms.

DR. GAN: Right. There is an attempt to try to make it objective, but again, recognizing the fact that some people ask for rescue when nausea is 3, other people don't ask for rescue until nausea is 8, so there is some subjectivity as well.

DR. WARD: That's the same issue that we brought around, and they had to do a clinical intervention. What's the criteria for that

Page 103

intervention? And it can be procedure specific and nervousness specific. I mean, if you have a patient who's at high risk for aspirations, you may be treating prophylactic or nausea quicker than you would for somebody who [inaudible – off mic].

If the same issue keeps coming up when you use clinical intervention, then how do you define the provider, physiology of procedure indications so that [inaudible – off mic.]?

DR. GAN: By making trials also have some sort of definition or another criteria when you can administer rescue antiemetic. Maybe you have vomiting, or your nausea scale is above 3, or some sort of duration, at least somehow to control that variability.

Any other thoughts, questions? Thank you very much.

(Applause.)

DR. WARD: John is going to tell us about cardiovascular.

Presentation – John Berkenbosch

DR. BERKENBOSCH: I can't promise that this

Page 104

will be nearly as entertaining.

(Laughter.)

DR. BERKENBOSCH: So I'm going to just jump into this. A little bit, by way of background again, I think a lot of this is going to be stuff that we've heard already. Hopefully, what I've done is just put something into a package and just summarize some of the thoughts.

But behind respiratory, cardiovascular events are probably the second most commonly reported outcomes in procedural sedation, the most common being bradycardia and hypotension. And incidences range anywhere from 5 to 25 percent depending on what agent you're using, what definitions are used.

Less that I've seen in the literature, looking at hypertension, other than, say, ketamine literature, potentially may be significant in patients that have particular risk factors. I think that's something that we may want to think about considering.

Then other issues, the major events are

Page 105

obviously the things that are going to be pretty universally reported, cardiac arrests, the need for CPR, just really basically requiring either electrical or medical conversion. Those are some of the things that have been typically reported in the literature.

I'll throw out the question. We've sort of bounced around this already. But when does an event become an adverse event, and when do we want to report those things? And another question is, when does an adverse event become an adverse outcome, because I don't think that those are necessarily the same thing.

When we're designing trials, I think we need to be transparent up front about how we're going to make those definitions, and then if we choose one thing, then call it what it is.

A little bit about what's been reported and how things have been reported from a cardiovascular standpoint, early studies have been relatively simplistic in that they were predominantly event-based, hypotension, bradycardia, and again,

Page 106

we talked about hypertension or rhythm issues. A lot of definitions out there, a lot of them are either threshold based, and that could be a change from baseline. It could be a drop below a certain percentile for age, gender, whatever.

I think, as was brought up yesterday, one of the things about using time-based thresholds is that if you intervene before that time-based threshold is reached, patient's becoming bradycardic, you intervene by stimulating, giving an anticholinergic, whatever, you may not reach that threshold of time, but you've made an intervention.

So that's where this idea of shifting to intervention-based is becoming more popular, the idea there, I think, being to tie an event to a functional outcome or to a perceived risk, and then as we discussed before, some of the challenges associated with differing thresholds between different clinicians about when to intervene. And those are some of the challenges that will need to be addressed when designing clinical trials.

Page 107

I think, really, at the base of looking at cardiovascular adverse events, though, to me, the question is, what is the adequacy or is there still adequate cardiac output and/or tissue oxygen delivery? How do we assess that?

Clinically, we often use it with organ-system-based approaches. Clinically, right? We look at mental status. We look at urine output for the kidneys, things like that. Those are often masked during sedation because you can't assess them.

They're also masked somewhat because these are very short-lived procedures. So if a decrease in urine output during a 15-minute sedation for a fracture reduction, or lumbar puncture, or something like that, that's clearly not something that's going to be clinically meaningful.

There are issues related to access to the patient. And again, I'm thinking to my own practice as a pediatrician, about half of what we sedate for is MRIs, where you have very limited access to the patient. That may be different than

Page 108

what is sedated for in the adult world, but those issues are going to limit your ability to assess cardiac output.

So there is one study that I was able to come across. It was one that we did at our institution. I had not seen anything else that looked at, actually, near-infrared spectroscopy monitoring during sedation to see if we could correlate changes in vital signs with changes in tissue oxygenation, or at least NIRS monitoring as a surrogate for tissue oxygenation. Again, we can go into the benefits of or the limitations of NIRS monitoring.

But what we found -- it was a limited study; we only did 100 patients -- there was relatively poor correlation between desaturation events. We looked more at that than we did cardiovascular events, poor correlation between that and changes in cerebral oxygenation.

Again, I would argue it's probably more the somatic oxygenation with NIRS monitoring that's valuable than the cerebral, but at least that was

Page 109

one attempt to look at are there other surrogates of cardiac output and tissue oxygen delivery that we can look at that aren't necessarily available to us as clinicians when patients are sedated. Obviously, NIRS monitoring is going to have some limitations as to where you can use it as well.

I think another question we have to ask -- and again, this has been discussed round and round; I won't belabor the point -- is which events are clinically relevant. So we see these statements in numerous studies: Regimen X resulted in more hypotension, but the clinical relevance of this is questionable; Drug X resulted in significant cardiovascular changes compared to baseline, but these changes were not associated with adverse events.

I think we all probably believed that those are true statements all throughout there, and you can do with that what you will. The question, without actually measuring cardiac output or tissue oxygen delivery, can we really say that these are true? Probably, but what's the degree of

Page 110

confidence that we have in those things?

I'll also throw out the question of what do we do with the concept of cardiovascular parameters in patients who are awake versus asleep. All of us become a little bit more bradycardic, a little bit more hypopnic, a little bit more -- and blood pressures typically drop when we are asleep compared to when we are awake.

So if we are looking at normal blood pressures for an awake person, and comparing that, and using abnormal blood pressures based on awake parameters in persons who are sleeping and sedated, is that really legitimate? Again, I don't know the answer to that. I throw it out there as a question for us to ponder.

We've touched on this as well, but are all studies equal? I think we have to look at what's the purpose of this study, because that may alter some of the outcomes of interest. Early on, some of the phase 1/phase 2 types of studies -- really identifying the human dynamic profile of a drug is probably the most important thing.

Page 111

The FDA, people that were here yesterday clearly said these are events that we're mandated by law to report, hypotension, bradycardia, hypertension, dysrhythmias. They may not be interesting to us from a, do we think that this was an event that caused harm to the patient, but there are events that were mandated to report in those early studies.

Later on, perhaps some of the more interventional based types of things to get a better handle on, the clinician felt that, that blood pressure dropped to the 70s in a 5-year-old was actually clinically important and decided to intervene in whatever way, shape, or form he or she intervened. So perhaps looking at a graded approach to outcomes based on the phase and the type of study that we're doing is something to consider.

We've talked about the Quebec guidelines and the TROOPS tools. I think I'm not going to belabor those points, other than I think there are probably some things on those tools that could be added if

Page 112

we would like, cardiovascular interventions in particular on the TROOPS tool, I thought.

There could be some other things such as medications given for bradycardia, hypotension, and weaning the infusions, which is often something described in the literature as well for things like propofol or dexmedetomidine. Those are just minor little additional points.

Other questions, I think -- and again, I apologize. I'm going to give much more in the way of questions than provide recommendations for what we should do. But is the duration and the depth of the events something that's relevant? Is 5 minutes of hypotension or 5 minutes of bradycardia more relevant than 30 minutes? Quite possibly.

Is 5 minutes of a heart rate in the 40s versus 30 minutes of a heart rate in the upper 50s -- is there a difference between those? Again, without knowing what the impact is on tissue oxygen delivery, that may be something that's difficult for us to get a good handle on.

We've talked some about this with looking at

Page 113

enrolling high-risk patients in clinical trials, but I think that the role of comorbidities is something that's particularly important to cardiovascular events, the cardiomyopathy patient for whom small changes in blood pressure, small changes in preload may be very much more clinically relevant than in an otherwise healthy patient; the patient with single-ventricle physiology, which our group and others that do pediatric sedation are going to be seeing a lot more than on the adult side.

Patients with brain injury, where you're looking at cerebral perfusion and patients with vasculopathies, the Moyamoyas, blood pressure changes, they're exquisitely sensitive to those compared to patients that would be ASA1s and 2s, where relatively marked changes are probably of much less clinical relevance.

I think that we, at our institution, probably struggled somewhat with reporting during the procedure versus during the recovery. I think that recovery is done by people. I know that

Page 114

recover is done by, in most places, by people different than the proceduralist who may have different motivations behind what they report.

Vital sign frequency typically is something that is decreased substantially during the course of the recovery period compared to during the procedure, our standard, every 5 minutes during the procedure, more if we are finding derangements and we're intervening. During recovery, it's typically every 15 minutes.

So you've got a 15-minute period where there may be some events that are happening. There was lots of discussion about the benefits of having electronic medical records, stuff streaming from the monitors into your electronic medical record.

That isn't something that's available at our place for certain types of studies or for certain types of settings. Our MRI scanner or monitor simply won't do that, so we're reliant on what we measure clinically during the scan when we're sitting at the bedside or in the scanner room.

Then I think that there is a role for at

Page 115

least considering what the third party needs, maybe within the setting of sedation. So industry has different interests in reporting events than regulatory organizations such as the FDA do, versus what institutions are interested in, because they want to know are people doing this safely.

So differences between me, my partners, the hospitalists, the intensivists, the emergency medicine people, all of those may play into what our institution is looking at as far as are you on benchmark. And maybe that's not as relevant to clinical trials. But if you've got institutions where multiple different providers are engaged in a clinical trial, that may be something that's important.

So as final thoughts, I probably should have reworded this as final thoughts to look at rather than recommendations, per se, but I think the question then becomes what to assess. Obviously, the big events are something that we look at. Was there cardiac arrest? Was CPR performed? Was either medical or electronic conversion of the

Page 116

dysrhythmia -- did that need to occur? That's important.

I think there is certainly a role for looking at bradycardia, hyper or hypotension. They're probably appropriate, particularly for pilot studies or lower earlier phase studies 1 and 2, simply because we need to look at the pharmacodynamic properties of a drug before we get into more detailed trials about them. And that's going to give us more information about how to appropriately dose, et cetera.

Same thing can probably be said about new regimens or combinations of drugs rather than just a new drug. Maybe you've got two different drugs that have lots of experience, but we don't know what the two of them do together, event reporting such as that. But I think that those need to be thought about.

I think, if we want to look at those events, then the question becomes, is it a change from threshold or is it a change relative to normative values? And I would advocate for using

Page 117

age-appropriate norms versus a percentage change, particularly -- and Jerry talked about this earlier as well -- for me, in pediatrics, a lot of these kids come in agitated. Their heart rates are sky high. A 30 percent drop from a heart rate of 170 to 130 or 120, big deal as opposed to if a child goes below a certain threshold, that's probably more physiologically relevant.

Intervention-based, I think there's lots of increasingly good reasons to think about using intervention-based approaches, but the question becomes, what's the trigger or rationale for intervening or not intervening for an event?

Ron, you talked about the kid who becomes hypotensive during dexmedetomidine, and we just sort of go, "Yeah. They're probably fine," or bradycardia, and we just kind of watched them. They may a little bit below normal, but they otherwise look fine.

So it may be worth looking at an intervention-based if we're looking at those outcomes to think about tying not just

Page 118

intervention, but maybe tying some sort of rationale questions into the intervention.

Did you do it because you were anxious? Did you do it because of how long it occurred, that sort of thing. I think those are questions worth asking. And I think, as we talked about, since studies are somewhat situation-dependent, maybe the outcomes may be somewhat situation-dependent as well.

Then I throw out there that may be some way of assessing cardiac output in certain settings. And again, you're going to be limited to scenarios in which some of that stuff can be done, but maybe NIRS monitoring, looking at two sites, cerebral and somatic sensors, maybe bioimpedence monitoring in certain settings, that may be something worth looking at and tying it to some of the vital sign changes to get an idea of whether these events are clinically as relevant or irrelevant as we would like to think that they are.

I don't know if you want to open up time for questions now or save it until later, again, not

Page 119

very much, solid recommendation, but just some thoughts and questions that I think we need to ponder as a group.

DR. GAN: So I've got two questions. One, you had mentioned about urine output monitoring. Now, we all know that urine output, especially around the perioperative period, can be difficult to interpret, that sometimes it's not uncommon you get oliguria, not so much because of any events or anything. It's just patient response to surgical stimulus.

How useful it is do you think that collecting urine output other than if there's no urine at all?

DR. BERKENBOSCH: I think my point was that it's probably completely unhelpful, because we're not going to be able to get good data, partly because the time frame that you're looking at is so short. We're certainly not going to put urinary catheters in these kids because that's just an unacceptable risk.

DR. GAN: But even in adults --

Page 120

DR. BERKENBOSCH: I'm sorry?

DR. GAN: Even in adults.

DR. BERKENBOSCH: But what's the frequency of a normal adult voiding? It certainly is much longer. The time between that is going to be much longer than the procedure itself. I think you also need to consider drugs like dexmedetomidine where you actually have a stimulation of urine output with dexmedetomidine. So I don't think it's particularly helpful.

As a surrogate for cardiac output, which we use in the ICU all the time. If you've got adequate urine output, it probably means that you've got adequate renal perfusion. I think, in this setting, it's probably unhelpful.

DR. GAN: The other question is, you had mentioned a last slide about potentially using a cardiac output monitor. And I'd like to get to maybe feedback from FDA colleagues. Maybe Dr. Roca can comment on it.

These are sort of relatively new monitors. And what is FDA's view if you are doing a clinical

Page 121

trial that you require more continuous cardiac output monitoring? Would that be acceptable as part of the obviously not standard monitoring, but additional monitoring to get additional data? How valid these are? Any thoughts on that?

DR. ROCA: I think a general statement would be that we are definitely flexible from the standpoint of somebody would like to try a new type of monitoring and new type of information. I think it's safe to say that we are very much interested in getting data, getting information. And even if it's a new monitor, we'd be interested in it.

Consequently, with that, though, it would be important for whoever's proposing it to give us information to make us feel that it is a reasonable way of monitoring. If they have information that would compare to other methods of monitoring, then we can get an idea of a comparison. That would also be useful.

But the bottom line is, yes, we're always open, flexible, but it needs to be supported with additional information and more data.

Page 122

DR. BERKENBOSCH: I guess one of the reasons to put that up -- there was also, I think, similar to what you presented, Denham, was with looking at alternative respiratory monitoring. It gives you information that we currently don't have available with the monitoring that we have.

Whether it helps us or not, I guess since there is such limited data in the procedural sedation environment, it may or may not be something that's helpful, but it may be something worth going through the process of to evaluate the usefulness in the procedural sedation environment.

DR. MINER: I think I've been putting tissue monitors on my sedation study just from about the last year or two. And I think one thing I've noticed is that there's tons of variation and changes that we're not picking up just from vital signs, so the big question is --

DR. BERKENBOSCH: Yes.

DR. MINER: -- is this because we weren't really picking up a lot of important stuff before or are we just getting a lot of background noise?

Page 123

And then which one of those are which? I think that it's got a lot of sorting out to do. I think it has a lot of potential.

DR. BERKENBOSCH: Yes. And we found the same thing with the study that we did with NIRS monitoring that we found -- cerebral desaturations were not common, but often were not associated with changes in respiratory rate, pulse oximetry values, and so on.

So they may be giving us different information. I'm not sure that they're necessarily correlated, but there is certainly different information. Yes?

DR. RIKER: Many of the non-invasive cardiac output monitors have clinical scenarios where they're known to be non-reliable, primarily spontaneously breathing patients with variable title volumes, and venous return, et cetera. I wonder, if we are going to go into this area, if we should describe situations, or monitors, or types of patients where maybe these are not felt to be reliable and cut that off at the pass, so to speak.

Page 124

DR. BERKENBOSCH: That's a good point.

DR. WARD: Thank you. Jim? Sorry.

Presentation – James Miner

DR. MINER: Jim Miner from Hennepin County Medical Center. I've met most of you now. I first started doing sedation research in 1999. I was actually hired by a head injury researcher who wanted to move all of her work from the lab out to a clinical setting. And she said -- I'm brand new, right out of residency -- she's like, "Hey, I got a great project for you. Why don't you develop some more precise and accurate measurement tools of level of consciousness, because we don't really have anything that's going to work for my head injury research."

So we used a sedation model because they started normal, became abnormal, and back. We thought, what a perfect place to develop. This will be a slam-dunk. We'll develop a bunch of objective measures of level of consciousness.

So we did head injury research, and it turned out I'm still working on that.

Page 125

(Laughter.)
DR. MINER: So that's a pretty hard thing to do, and we talked about it a lot at the last meeting. Neurological adverse events, you can't talk about without talking about level of sedation. But we're talking about adverse events now more than just measuring that level of sedation, which can be really hard to do. We'll talk a little bit about awareness and recall, paradoxical responses, and then permanent complications last.
For the level of sedation, there's a lot of different scales out there. I think a lot of people have been settling on the OAAS scale, which is pretty universal. It tracks really well with a lot of just the definitions of level of sedation. If you look at it, it really fits with general anesthesia, deep sedation, moderate sedation, and the levels that work.
I think it's got some good external validity, and we all accept that these are levels that probably mean something in our practice and are associated with increasing risk.

Page 126

There's a lot of different scales out there, and every couple years, someone will come up with a new one and name it after their institution or friend. But they all essentially do the same thing.
Whoever's measuring it can say whatever they want. It's not an objective measure. You can try bringing other people in to measure, and they almost never agree. I've tried to study moderate versus deep sedation so many times, and the confounder is people just call it what they want to call it based on how they wanted the patient sedated for the procedure. So it gets really, really hard to compare just because it's so vague.
But as we're talking about what can be the adverse event, I think we can all pretty much agree that if you're trying to do procedural sedation and you end up with a generally anesthetized patient, that's an adverse event.
When we first started looking at this, the numbers that were coming back from our studies weren't as high as I knew just from my clinical

Page 127

practice.
People weren't reporting it well, so having an outside observer measure it, I thought, was really important, because understanding that you're not supposed to do general anesthesia during procedural sedation, people are reluctant to self-report that this person indeed was not responsive to any form of pain unless someone else was doing it for them and not telling them what they're measuring.
So they'll try to wake them up, and harder for you, or wait a few seconds until they're not generally anesthetized and not report it. But I think this is a more concrete outcome. Usually, we can get in there.
Deeper level implant is much harder just because eye opening falls in those scales a lot, and that's a hard one to use, because if we really use eye opening as a hard scale, people close their eyes all the time when they're being sedated. It has nothing to do with them being more deeply sedated. We've all seen very, very deeply sedated

Page 128

patients whose eyes really aren't closed.
So this one gets a little harder to use, but I think it's important because when we're talking about what's clinically useful out in clinical practice from the research we're going to report when we study things, if an agent consistently brought people deeper than they intended to bring them, at least from training purposes, people who use the agent probably need to be trained for both levels of sedation, and it comes up for credentialing purposes.
For example, a drug like propofol, where you can get moderate sedation with propofol, but you're going to end up with deep sometimes. And if a person isn't good at managing deep sedation, then it all ends up being a problem. So I think it's a valuable thing for us to try to record and report.
Missed target level is much harder. That'd be an outcome I'd love to be able to say, where you wanted to hit a target and you couldn't because the agent was too variable. Again, that's a harder one to measure, but I think could be a very useful

Page 129

measure as we come up with more tight ways to measure it.

Objective monitoring, EEG and Bis scores still work pretty well. They're not great from clinical practice, but from study purposes, we've had some luck. This monitor I think has been pretty useful from research purposes in terms of giving an objective measure, especially for predicting who's in that recall or not from the sedation.

I think it's got some value to it. I've kind of moved away from using it because it can be time consuming and mess up the trial from that way. But I think there's some value there and trying to come up with a non-scale, something that you can't change based on not being blinded to a drug or figuring out what the person looks like.

Capnography can give us a little clue into people's level of sedation. There's more variation. Just we haven't been able to come up with a really good way to do it. So I think that's still just something that's out there. Maybe

Page 130

someday, we'll be able to look at the variation in breathing and predict people's level of consciousness from it.

I think it would be very useful and a lot of us can do it clinically. We just can't report it in a way that's useful for anyone else to pick up or we could write it down on a paper that a reviewer would accept.

Biomarkers have tried to look at -- specifically I've looked at salivary amylase and cortisol and serum catechols, and it's not ready for really big use. I think what I've seen is that when you've sedated a person, people who have an adverse respiratory event have huge catecholamine surges, but people who get a pain medicine or don't, and then get a procedure done, the changes are too small for us, or the variation is so high that we can't pick up a change in the size trials that I'm capable of doing because it's a lot more expensive.

So really, again, it's not really ready for us to use. But I think there's some potential

Page 131

there if we can come up with good ways. Salivary amylase is a rough measure of synthetic stimulation. It goes up very, very quickly when somebody is in distress or pain. It could be useful for us in the future. It just adds $200 per patient when you're doing research, which limits our ability to do the trial.

The same with salivary cortisol, it goes up the next day if people had a stressful response, the problem being adverse respiratory events seem to stimulate that response much, much more so than pain or being awake. So there's a confounder there.

What's stressful to a person? Probably under-perfusion or not respirating is much more stressful than some pain you don't recall to the point where it's heard to use biomarkers that we currently have identified.

Other things you can use that have a lot of inherent sense to them but are hard to pull into the adverse event territory are the duration of procedure and difficulty for the provider. If

Page 132

you're under-sedated or over-sedated, the duration will expand or contract relative to a procedure.

Also, recovery time tends to get longer when people are over-sedated or there was a complication from it. As with the patient satisfaction, under-sedated patients tend to be dissatisfied with the event unless they have a ton of opioid on board when they were awake.

For awareness and recall, recall is not that hard to study, but it's not great. There's not a lot of great tools out there that we all accept. I've tried a number of them. Probably a lot of you have tried a number of different recall tools. You get mixed results.

Probably the best way is to just ask the person do you remember anything and was it bad or was it not bad. My experience has been that if you're using an opioid, say, just give somebody alfentanil or something, they'll recall the event but say, "Well, it really wasn't that bad," which is confounding, because in propofol, they won't remember at all. And if they remember at all,

Page 133

they'll say it was horrible if you don't give them an opioid, too.

But it's useful, and it makes a lot of inherent sense. If somebody remembers something and it bothers them, it probably was bad and, if they remember something and it doesn't bother them. It just gets very imprecise as you're trying to measure it for trying to do studies.

For non-procedural things, I've tried a lot of different cues, and I think a bunch of us have tried different cues. Reading words to people and having to repeat it back is really valuable for objectively assessing their level of consciousness because whether or not they can hear you and repeat a word back is a pretty good measure of how sedated they are.

But on the recall end, even when we do this experiment just to wake people who haven't been sedated at all, they miss words kind of randomly in a hard-to-predict fashion. So it's not a very good cue of what they're remembering and what they're not. Although it's been used a lot, I think most

Page 134

of the evidence we have for retrograde amnesia comes from word studies like that, and they've been okay.

We've tried some pictures. I was having a little bit more luck with picture cubes now, giving people flashcards with pictures on them once a minute and having them describe the picture to us. They use the description to describe their level of consciousness, and then show them the cards afterwards, and say which one of these did you see during the procedure, and have them point it out. That's been a little bit more accurate when we do it in a non-sedated control than it is in patients.

So there's work to be done in recall. I think it's valuable information, knowing exactly what points people have and don't have recall. Coming up with tools that'll give it to us on ends will help us define what people are remembering because incidental memories occur. People can put together memories that didn't really happen. And we've all seen it. When you ask people, do you remember stuff, they're like, oh yeah, I remember

Page 135

the whole procedure really, really well. What they're really remembering is somebody finishing, putting a bandage on at the end of a fracture reduction, or cleaning up and they heard some clanging and it made them upset because they were slightly sedated.

Then they're fully alert three minutes later, and they form that into a well-formed memory, where they filled in all the gaps with gory details that they were worried about before it happened. And you get this confabulatory, very clear memory from the patient, but you start to pick up clues that this is not actually something that happened, and they just kind of put it together in the confusion of waking up, which makes recall as an adverse event very difficult.

Again, stress response, same thing, biomarkers for recall, this is a little bit harder to do. When we've looked at this, we can't show a difference between when somebody appears alert but can't recall pain to where somebody did not appear alert when they had something painful happen, the

Page 136

distress biomarkers, at least of catechols, cortisol, and amylase. It doesn't seem to show anything relative to the magnitude of response received from hypotension or hypoxia.

Other things, excitatory movements, muscular rigidity, usually as long as it's well defined in a study, is a good adverse event. I think most people would agree it can get in the way of a lot of procedures you're sedating for. Myoclonus is the same way.

Generalized seizures, we don't see a ton of and. Usually by the time somebody's having a generalized seizure from procedural sedation, they've got something else going on that's already got them well into the realm of adverse events.

Paradoxical responses to sedation and unpleasant recovery actions, there are some great tools out there to measure these. I think we all agree that these are adverse events that we want to avoid. I think the key is encouraging investigators to all use similar tools that we can compare our data from study to study. But these

Page 137

are clear adverse events I think most of us can agree on.

Lastly, on permanent complications, I've heard a lot discussions about is there cognitive deficits or any permanent neurological injury from sedation, from people who get sedated multiple times, or for any of the adverse events that occur.

It's really hard to measure because we've been talking about very subtle changes by definition. And picking up these subtle changes can be pretty hard. I don't think we have precise enough tools maybe to pick up the level of deficit we're picking up unless someone gets sedated regularly for a long period of time, which are the people there's been reports that maybe we're seeing some cognitive deficit there.

I think, getting towards what John was talking about a few minutes ago, maybe we should be looking at something else like cerebral perfusion or something that would cause that problem. It's an easier step that we're closer to achieving. This is definitely something we need to think of a

Page 138

good way to measure it or come up with a surrogate at least to measure risk for it so we can look at it going forward.

Death, obviously that's the adverse event that's hard to skirt around and usually has a pretty clear definition, but would be the ultimate neurological complication of procedural sedation.

All right. Thanks. I had a much longer version of this and decided that you didn't need to hear that, so I shortened this up last night after dinner.

(Laughter.)

(Applause.)

DR. WARD: Questions? Yes.

DR. TOBIN: John, thank you. As an anesthesiologist, we worry about post-operative delirium after major surgery and, after some of the most complex surgery, it can be 30 to 50 percent. But even after hip surgery is done under regional anesthesia with what we would call moderate sedation, the incidence of delirium is 9 to 15 percent.

Page 139

Should we be looking at delirium in kids or adults? And in fact, much of the delirium is hypoactive. It's not hyperactive. They just don't respond with their environment appropriately. And I know that's really hard to assess in a kid, but should we be looking at the acute post-operative cognitive deficits, not just the permanent ones?

DR. MINER: I think that's a great idea. Some of that gets lumped into when we measure recovery time now, assuming that the person is truly measuring getting back to baseline. That's the only surrogate for that that I've seen consistently reported. But it seems to me that would be a really important thing to know about an agent.

There's a question. Yes?

DR. IRWIN: Yes. I was just going to comment on the use of processed EEG, because I don't think it's a good monitor of consciousness, actually, because there's a lot of intra-individual variability.

DR. MINER: Sure.

Page 140

DR. IRWIN: And it's very drug specific. So if you're doing a trial comparing one drug against another, at sort of equi-sedation levels, you would actually likely find, even if they were genetically identical, a different level of, say, best, or sed line, or whatever.

DR. MINER: I completely agree. I think it's all over the place. It's been challenging. It's not a good measure, though.

Well, thanks, everybody.

DR. WARD: And David? There he is. Okay.

Presentation – David Gozal

DR. GOZAL: Great to be here. I am coming from Jerusalem. As a matter of fact, it looks to me that it is much easier to talk about respicardiac adverse event as anything. Saying that, I think there is not much to say about liver, or kidney, or massive adverse events. So I put it into questions to the group, and I hope that it will be an interactive session.

Does it really matter if there is a muscle, or kidney, or liver dysfunction? And if it does,

Page 141

what do we need to measure concerning these kind of adverse events?

If there is a organ dysfunction, does it really change our sedation plan or strategy? Are there really some limitations concerning our sedation plan? Anyone in the group?

DR. LIGHTDALE: I mean, David, I could talk mostly thinking liver. Right? So let's just say you have a terrific sedative, but it makes your AST/ALT go up into 400, 800 range, and I don't know how long that's going to last, yes, I would say that matters.

The same for kidneys, you wouldn't want to cause kidney damage, even if the sedative has no cardiovascular effects.

DR. GOZAL: Yes?

DR. IRWIN: Yes. There's also the effect on PK/PD of the organ dysfunction itself in terms of the duration of oxygen of the drug or the sensitivity of the individual to the pharmacodynamic effects. Is that what you're thinking of?

Page 142

DR. GOZAL: But in practical, does it really change --

DR. IRWIN: It would change your dosing, I guess, and maybe change your choice of drug for that reason also.

DR. DAHAN: Well, take for instance a child that needs repetitive radiation therapy.

DR. GOZAL: Okay.

DR. DAHAN: And you sedated with ketamine. You see an increase of its liver function. I would change my anesthetic. I would use propofol the next time.

DR. GOZAL: So why not to start --

DR. DAHAN: Ketamine?

DR. GOZAL: Yes, from propofol to propofol.

DR. DAHAN: Instead of propofol, because I used ketamine quite a lot for sedation, I would talk a little bit more. This is my preferred way of treating especially small children for sedation rather than propofol.

DR. GOZAL: So I'll give you some ideas. Concerning the [indiscernible], I don't know if

Page 143

anybody here asked me to sedate an ALS patient. When I was statistician conducting a study concerning ALS patients, one underwent bone marrow aspiration. And we were concerned about how this patient could breathe, and could we send them directly after the procedure home? Any suggestions?

DR. LITMAN: Well, we know that some anesthetics in some patients can cause muscle damage. And we usually measure that with post-operative CK, creatine kinase. That seems to be a reliable marker.

DR. GOZAL: Actually, we had a serial for about 80 to 90 patients with ALS, and we did propofol and alfentanil, and they were on spontaneous respiration, lying on their left side, and they did very well, actually.

What do we need to measure? Does it really matter? I put here some muscle strength tests. We can discuss it. Liver, we should look after some neuropsychiatric science concerning hepatic encephalopathy. Propofol is said to be a better

Page 144

drug than anything else in patient with liver disease.

Kidney, instead of measuring potassium and creatinine, we should measure the GFR. That is much more [indiscernible] on these kind of patients. Thank you.

DR. WARD: I guess a question is, [inaudible – off mic] case 1 toxicology studies [inaudible – off mic], and is that sufficient, or do we need to be recommending anything more than that?

Although I think the point of understanding of pharmacokinetics or the dysfunction is a part of [inaudible – off mic]. I wanted to include this, realizing there wasn't a [inaudible] up there. But I didn't want to be criticize when we send a paper [inaudible].

DR. GOZAL: That is exactly the point, that you send a paper, you talk about cardiac or respiratory dysfunction, but never all adverse events when it may be neurologic or nausea or vomiting, but nothing concerning kidney or liver.

Page 145

DR. WARD: Looking at [inaudible – off mic]. So that was a kidney effect. There are other kidney effects [inaudible – off mic].

DR. GOZAL: Thank you.

DR. WARD: I'm trying to be complete.

So a little different topic that I asked Dr. Dahan to do was sometimes we can do, interventions from respiratory, nausea and vomiting. We've heard about pharmacological interventions to prevent adverse events.

Not thinking about a clinical trial for a new sedative, but thinking about a clinical trial if you have a new compound that would prevent adverse events, how would you design a clinical trial to do that? Thank you.

Presentation – Albert Dahan

DR. DAHAN: Thank you. Thank you, Denham, for inviting me. It's a great pleasure to be here.

Well, I've changed the topic a little bit, to be honest. And I will be talking a little bit on the trials that we've been performing on the effect of especially opioids on ventilatory

Page 146

control.

Most of the studies were on drugs that were used in chronic pain because most of the studies that I currently do are pain studies, and I consider respiratory effects side effects, not a primary effect of the drugs that we've been using.

So I have a couple of topics. I talk about intervention that might cause less adverse events aimed at respiratory depression. I will talk a little bit on what Denham just said, a co-medication that might prevent or reduce the probability of adverse events.

I will talk about what I think, and what Bob already spoke about, an especially interesting concept in especially phase 1/phase 2 trials, the development of utility function in which you look at benefit as well as the risk of a side effect.

Then at the end, because I understand that this is the most important part of this meeting, what should we measure in trials of sedation, what would I measure in my trials of sedation. I will talk a little bit on monitoring online, and I talk

Page 147

a lot in my own institution about education in sedation.

We have two large programs in which we educate both non-anesthesiologists as well as nurses. In our institute, non-anesthesiologists apply mild sedation, and people that we call sedationists, which are nurse anesthetists with a three-year additional training to give sedation, they apply deep sedation in our institute, but all supervised by us. So we do not have nurse anesthetists that work on their own.

So the first question, are there drugs that cause less respiratory depression? That's not an easy answer to give. There has been recent publications on so-called biased ligands. These are opioids that affect the new receptor differently from classical opioids.

After activation of the opioid, you have two pathways, at least in animals they've shown the beta-arrestin pathway, which induces respiratory depression and the G proteins that induce analgesia. And there are specific opioids being

Page 148

developed aimed at G protein and not at the beta-arrestin. However, I've seen various human studies that were extremely disappointing in this respect, but I might be mistaken.

Another new development is a drug that I'm currently testing, which is a mu agonist and a delta agonist in which the claim is that the delta agonist antagonizes the respiratory depressant effect of the mu agonist. I haven't done the study as yet. I'm developing the study, and it might be interesting.

I have been involved in studies on opioids that also act on non-opioid receptors. And I give you a result here of a study that we recently finished. This is on a drug called tapentadol. You have it in the U.S. It's a drug used for chronic pain. I'm currently working with the company to develop it for perioperative use in IV formulation. The reason is that people claim that it causes much less respiratory depression.

Well, I've done a whole series of experiments in a large group of volunteers, and

Page 149

this is the result. This is a parameter that Denham already showed, isohypercapnic ventilation at 55 millimeters of mercury. This is the placebo. These are crossover trials, so every subject was tested 4 times placebo.

The opioids, we're talking about and oxycodone. And this is an equianalgesic dose because we also did pain testing analgesia, nociceptive testing. And as you can see, in terms of depression of isohypercapnic ventilation, the effect is quite large.

I'm not really sure how this extrapolates to the clinical setting, but in my hands, we showed that, for the first time, this drug produces an equianalgesic dose as compared to oxycodone, less respiratory depression.

What it means, I do not know, but it's an interesting concept that, if you want to use an opioid for chronic pain, and you have a patient prone to respiratory depression, an obese patient, a patient with pulmonary disease, this might be an attractive alternative.

Page 150

Now, another way that I am looking at sedation is the use of ultra-short-acting opioids. In my country, labor pain is now being addressed apart from using epidurals with PCA remifentanil. I'm not sure if this already hits the United States. It's an epidemic in my country. The main reason is that many of my colleagues, anesthesiologists, refuse to give women in labor an epidural outside of the daily hours, the daytime hours, I have to say.

So they came up with this solution, remifentanil PCA. I did a very large number of trials to test the safety of this approach. Just to give you one example -- this was published a couple years ago -- this is a single-bolus PCA dose of remifentanil. And as you can see, this minute ventilation goes down.

If you add oxygen, so if these patients would have been on oxygen, you can see two things. By the way, this is respiratory rate and this is saturation. First of all, you miss the occurrence of saturation.

Page 151

This was a patient without any respiratory activity for more than 3 minutes. We put her on mask ventilations, or this, artificial ventilation because I did not want to have a patient in a study that stopped breathing for over 3 minutes.

I recently replicated a similar study post-operatively and had some of my patients that were without any respiratory activity on oxygen for over 5 minutes. To me, it's totally surprising, no effect on saturation.

So this shows you how dangerous it is to just look on desats in your studies. I don't even use it as an outcome parameter anymore at all. And I hope that, in -- did it hit the United States, remifentanil PCA, for labor pain?

DR. KOCHMAN: I know that some people have experimented with it at Penn. One person wants to use it a lot more than others. I think most of us have a fear of this.

DR. DAHAN: Well, we've done already over 50,000 patients with it. I'm not really much in favor of it because there are side effects, and we

Page 152

have had several resuscitations, because it's a potent opioid, and flushing is an issue.

DR. KOCHMAN: What dose do you use?

DR. GAN: I use a bolus dose of 30 micrograms with a lock-out time of 5 minutes.

DR. GAN: And do you put any continuous infusion? No?

DR. DAHAN: No. That's all they get.

So now, about co-medication to prevent or reverse adverse events, again, respiratory depression. There are several drugs being developed for this purpose. There are several ampakines, although I haven't seen any phase 3 trials with that. Physostigmine is a possibility, talk about it later a little bit. Tianeptine is an antidepressant. It is non-specific anti-depressant. It reduces the amount of serotonin in the brain. But the turn-over of serotonin at the synapse increases.

I am working on that drug. So far, the results are limited; disappointing, I would say. Animal studies do show large effects on

Page 153

respiration. Again, in humans, maybe it's a dosing issue. I'm not really sure yet. We're still working on it.

Serotonin agonist ketamine, I am in favor -- I use ketamine quite a lot in my practice, both perioperatively and for sedation. It's not a drug that actually stimulates breathing in the sense that your minute ventilation increases, but it reduces the variability that you see -- and I will give you some examples of that -- so that breathing becomes more stable, and the probability of an apneic event decreases significantly.

So this is really important. It's not a respiratory stimulant like, for instance, drugs that I show here below, drugs like doxapram, GAL021, and other experimental drugs. This is an example of this GAL021 drug.

This is an alfentanil-induced respiratory depression at isohypercapnia. In orange is the placebo response. You see severe respiratory depression by over 50 percent. And adding this drug, GAL021, increases respiration.

Page 154

It does so by blocking potassium channels at the carotid body, mimicking hypoxic event. The body is not hypoxic, but the carotid bodies think the body is hypoxic. Sadly, the company that developed that went bankrupt recently.

I find that, of all drugs that I've tested so far, the most promising one, doxapram in my hands is without large effects. The reason for that is when you stimulate the carotid bodies to increase ventilation, you need to be aware that the actual depression is not at the carotid bodies, but it's within the brain. So there is a large ceiling effect. You need to overcome a real hurdle.

If respiratory depression is severe, so if it's more than 50 percent, also GAL has a very limited effect, and you have to refer to drugs like ampakines. So there are some promising data, but I have seen little development further than phase 1 or phase 2 trials so far.

Now, the group of Denham Ward and Dr. Karan did a recent study using physostigmine on the apnea/hypopnea index on sedation. And they

Page 155

observed some interesting observations. One of them was the reduction in apnea-hypopnea index.

But what surprised me from that study, and I liked it very much, is that, in oxygen, the severity of apnea-hypopnea index increased, while after the application of physostigmine, the events decreased significantly. Maybe Suzie can talk about it a little bit later.

Now, like I said before, I always perform analgesia experiments together with respiratory depression experiments. Why is that? Because I'm very much interested in the whole picture of it.

If you have a certain amount of analgesia, how much respiratory depression is connected to that? And I call that a utility function. In fact, a utility function is just a probability of analgesia minus the probability of an adverse event.

An adverse event can be anything. In my studies, I use respiratory depression. And this is an example that we published in Pain I think a year ago, in which we looked at a utility function of

Page 156

fentanyl. And what we did, we made a distinction between subjects that had a high analgesic potency for fentanyl and the ones that did not.

What was quite surprising to me, something that I really had not expected, was that the ones that had a high analgesic potency had a very positive utility function, meaning that the probability of respiratory depression was very limited.

It was really very limited because this is a very high utility of almost 1, indicative that over this dose range, the drug in this group of subjects is predominantly an analgesic with little respiratory depression. Of course, at very high dose, you will induce respiratory depression and the utility function reverse to zero.

On the other hand, subjects that had a low analgesic potency, low analgesic efficacy from fentanyl probably due to genetic issues or other issues, they had a very negative utility function, indicative that these are the subjects that will in fact be the ones that will have a high probability

Page 157

of respiratory depression.
Now, we know this. You know when you have a patient that does not respond to your analgesic treatment, you keep treating them with higher dosages. Stop. There are many case reports in the literature that show that you should not treat your patients like that. At one point, you have to stop and revert to another therapy.
Why do I think the utility function is extremely interesting also in the development of drugs? Because you can compare drugs. For instance, here, we again have the fentanyl function. This is the total population, so there's a negative effect, as expected from the last slide as well, and compared to, for instance, morphine that has a predominantly positive effect.
So to treat for instance pain patients with either morphine or fentanyl, I would choose morphine in the sense that the probability of an analgesic effect exceeds the probability of a respiratory event.
Now, what should we measure? Like I said

Page 158

before, you can make it as complex as you wish. You can go for purely ventilation, hard to measure. You can go for respire rate. We discussed about it. I measure flow at the mouth. It can be CO2. It can be temperature. It can be humidity. You can do experimental tests like hypercapnic responses -- Denham showed you that -- hypoxic arousal responses, showed all of that.
In our trials of new medication being delivered, administered to patients, we keep it extremely simple. I nowadays look at the entitled CO2 pattern -- I'm very much interested in a ventilatory pattern, and I will show you later -- and your current apneic events. I'll give you some examples of that.
For instance, this is a study that was recently performed. It was recently accepted for publication in which we looked at oxycodone. If you look at the literature quite well, especially the forensic literature, you will see that patients that die from oxycodone, a lot of them, over 50 percent had alcohol, ethanol, in their blood.

Page 159

So we wanted to have a look at the interaction between ethanol and oxycodone on respiratory events, and we measured apneic events. In my lab, an apneic event is the cessation of flow at the mouth for over 1 minute, 1 minute or more, I have to say. And we looked at 2 populations.
I talked about it a little bit already, although that was another study. In this study, we looked at the interaction between two levels of alcohol, 0.5 grams per liter and 1 gram per liter.
This alcohol was delivered intravenously, and you can compare it to 3 to 5 beers per hour in this population. And believe me, patients like this take their oxy. I have many examples of patients that take their oxycodone with a very nice glass of whiskey.
So it's not academic at all. Look at the literature. And what we showed, there was a tremendous explosion of apneic events in the elderly population at the highest concentrations of alcohol. So this is a relevant and non-academic observation that predicts the occurrence of a

Page 160

deleterious, fatal events at home.
Now, I talked about breathing variability quite a lot. And this is an example of a patient on morphine. And this patient in the end developed apnea and gasping, but we could predict it already minutes, at least 20 minutes, before it actually happened because his breathing became extremely irregular.
This is something that Denham showed us also with dexmed. There is at one point cyclic breathing, episodes of normal breathing, followed by a period of normal breathing. And if you look at this, and you know how to look at these patterns, and you quantify these patterns, you're able to predict the occurrence of apnea and gasping in your patient. I use this map quite a lot.
By the way, in our previous slide, all these apneic events were preceded by -- we quantified it by periods of irregular breathing.
Well, there is a new measure on the market marketed by Medtronic. It's called the Integrated Pulmonary Index. This is an index that takes into

Page 161

account four measures: heart rate, oxygen saturation, respiratory rate, and entitled PCO2. It uses a specific algorithm. The algorithm is available and comes up with one number.

This is a patient in the ICU where I measured these four measures on the IPI device. And there are certain indications that there might have been going on something, a sudden increase in heart rate here, a large increase in heart rate.

Respiratory rate at one point was low. Entitled CO2 was low, which is an indication that the patient is hyponeic or something else. Maybe some obstruction is going on. And the IPI gave us values below, let's say, 7 and certain indications.

But the problem I have with these kind of measures, I have no idea what it means. The patient did well. I saw this the next day. I withdrew data from the system. I asked the nurses what happened. They said, "Well, nothing happened. What do you think happened?" I said, "I don't know." So it's really, really difficult to link the things that we measure in our patients with

Page 162

actual outcome.

Okay. I think this is it. Any questions or other items that I might have missed?

DR. GAN: Question -- Albert, that's a wonderful summary. I have two questions. One, you mentioned about ketamine making the respiratory pattern more regular.

Do you hassle to hypothesize what mechanism is that through?

DR. DAHAN: Well, I use ketamine for sedation quite a lot. I do not combine it with propofol because in my sedation, I don't want a patient to be totally unconscious. I want them to be comfortable.

So I combine ketamine with remifentanil. It's my preferred way of sedation of patients. And believe me, by doing that, because these two drugs act synergistically on pain relief, on sedation -- ketamine by itself is a sedative -- patients love it.

Last Monday, I did three sedations with the system. There were ablation procedures. And one

Page 163

of the things that I like very much of adding ketamine is that if you give for instance propofol/remifentanil, there are lots of expiratory pauses.

The patient exhales and there is no breathing activity for some time. All of a sudden, he inhales or she inhales and, when the catheter is positioned at the spot where the surgeon or the cardiologist wants to ablate, there's small movement.

I've had repetitive discussions with the cardiologists on this, and we decided to move on to a new way of dealing with that. How it changes the respiratory rhythm, difficult people - claim -- because in Rett syndrome, R-E-T-T, which is a complex pediatric syndrome -- you might know all about it; I am not a pediatric anesthesiologist, but I've read a lot about it -- ketamine also helps to make breathing more regular because these children also have a breathing problem.

I think it works for the AMPA system in

Page 164

which there is more tonic activity from specific areas in the brain, like the pre-Botzinger complex, but this is very complex, and breathing becomes more regular in the sense, like I said before, the probability of irregularity decreases. I really think that.

DR. GAN: Just a follow-up question, in terms of dosing, do you give boluses of fusion?

DR. DAHAN: No. To be honest --

DR. GAN: Do you combine with benzodiazepines?

DR. DAHAN: No.

DR. GAN: Any side effects of the --

DR. DAHAN: I am from the country of TCI, so everything we do is computer controlled, so I know exactly which range, it's around 100 nanograms per mL. I target that. And might go up and down a little bit, depending on blood pressure, but with remifentanil, blood pressure goes down a little bit; with ketamine, it goes up. It's quite stable.

Side effects, yes. Side effects do occur. some subjects have nasty dreams. We talk about it.

Page 165

This is a matter of coaching. I coach them through it. In the end, there's a minority of patients that dislike it, I do agree. Yes. There are side effects.

DR. GAN: Do you find that by adding a bit of benzo, that --

DR. DAHAN: I never do that. I don't think it's necessary. Some of my colleagues use clonidine, which is an alpha 2 agonist, to dampen the side effects. I dislike midazolam, long acting, interaction with all kinds of effects. Why do that? Talk to your patient.

Another issue is I use esketamine. I'm not using the racemic ketamine. This is relevant in the sense that the equianalgesic, equianesthetic doses produce less salivation. I like that. With racemic ketamine that I used years ago, there was lots of production of saliva. You know what I mean? The mouth might be full of saliva that interacts with breathing as well. But this is my personal opinion. It's not evidence based.

Yes, Suzie?

Page 166

DR. KARAN: Respiratory arrhythmias that you said that you observed before you have the apneas, are you measuring any duty cycles or any --

DR. DAHAN: I'm measuring expiratory time, to be honest. I'm looking at the variability in expiratory time. I have a set-up that allows me to do that. So when the variability is increased by a factor of 2, I know something is going to happen.

So it's not that breathing becomes slower. I do not care because I've seen definitions of a serious adverse event when breathing is lower than 8 breaths per minute. Measure your breath while you're asleep. Do you have an adverse event? I don't think so.

DR. KARAN: In that study that you did, though, your patients were awake with their eyes open. Right?

DR. DAHAN: They were awake. They had apneic events for over a minute while awake. We measured the EEG, correct, yes. And I am not the first one that shows that. Gordon Drummond from the U.K. in Edinburgh, he found the same thing. So

Page 167

this is not new.

In my opinion, when you give potent opioids to patients, they all become OSA patients, obstructive sleep apnea patients, especially in their sleep. But the fact that they obstruct awake -- no. I'm not saying obstruct wide awake, become apneic awake. It was a big discussion with the referee [indiscernible] also. Yes, he was right. Okay.

DR. GAN: Sorry, just one quick question. Do you believe that, for kappa opioids, it reduces respiratory depression?

DR. DAHAN: The drug that I just showed you, the mu delta drug, also is a kappa agonist. There was a big discussion on the mechanism of the reduction of respiratory depression. The company claims that it's the delta agonist that reverses the respiratory effect.

I am not really happy with kappa agonists because they cause dysphoria. They claim that it's the new agonism that counteracts the dysphoria. o we're dealing here with very complex molecules

Page 168

that, in the end, I have no idea how they work. And if you would take out one component, what would make it do, I have no idea. Okay. Thank you.

DR. WARD: Everybody else come back up, speakers, and get some summary before we have lunch, Jenifer, I think, around the topic of [indiscernible] measure.

Q&A and Panel Discussion

DR. LIGHTDALE: Does anybody have a provocative question to start us off? Okay. I'll try. So it just strikes me that each one of you have talked a lot about transient phenomena. And we talk about relevance and we talk about transience.

How should we define transience? Should that be something we really focus on, too? I'm actually thinking about yours, Albert, and then --

DR. DAHAN: Transient phenomena. Well, I remember a paper from Dan Sessler, who measured hypoxemia post-operatively, and many of the patients were hypoxic, actually. Only when the nurse came around and they took 1 or 2 breaths, I

Page 169

guess, saturation went up.
So some of these events are not at all transient and might occur. I do not know the clinical effect of these long-term hypoxemic events. I do know that I'm quite afraid, because I didn't speak about the fact that, in my study, I had to ventilate one or two of the patients or stimulate them to take a deep breath.
So even a short apneic event is significant if it is not restored to normal breathing. But in normal sedation, there's always somebody there that doesn't look at the patient and looks at the pulse oximeter or the respiratory rate monitor.
DR. WARD: I think in a clinical trial, you could define the length of time for the transient event when it becomes important. But I think it was Suzanne yesterday that kind of separated the measurement instrument from the interpretation instrument. And I think that's an important concept.
You could be measuring these things continuously, and can report continuously, and then

Page 170

decide in this particular relevance that none of these were long enough to really get to the point where they're really considered adverse events, but not to try to kind of pre-select and ignore a transient event. If you can measure it, you should record it.
DR. MINER: I think one of the problems we run into is we're not really measuring adverse events most of the time. We're measuring events that probably we all agree might be associated with an adverse event that's so rare that we can't really study it.
So then we have to decide, well, what's the amount of time that matters where we're not just over-calling it? Because if we over-call it, when we first started using entitled CO2, we over-called changes. We were finding, like, 40 percent of patients had changed. And that is just meaningless. We can't compare anything.
So then we have to say, well, what matters? And I think there's just a lot of work to be done yet as to which one of these adverse events

Page 171

actually is going to be a risk factor for something that matters, that we see 1 in 10,000 or 1,000 times.
DR. WARD: Yes. And I think that's absolutely the right point. Almost everything that we're measuring is a biomarker, and the literature is full of biomarkers that turned out to be either no good or bad.
I think we've got some that are at least no good. I don't know if I got any examples of biomarkers are bad, other than they maybe create situations that had too many interventions. But yeah, I think, for clinical trial, we've got a lot of biomarker data across all the organ systems fixed.
I thought nausea was not the one that counted, but TJ came up with one in which nausea actually could be good.
DR. BERKENBOSCH: The degree of transience is also somewhat dependent upon the frequency and method with which you are monitoring. And I'm thinking of things like respiratory rate and oxygen

Page 172

saturation, heart rate.
Those are measured continuously, so you can see transient versus things are a little bit longer lasting, whereas with blood pressure, we're measuring every 5 minutes most of the time, every 15 to 30 minutes during recovery. And once they've finished the recovery, we assume that they're fine, and they go back to getting measured every hour to 4, whatever if they were admitted if they were inpatients. If they're not, they go home and we never know.
So it becomes challenging to, I think, know about the degree of protraction in some of these events.
DR. GAN: But there are technologies now that you can measure non-invasively continuous blood pressure. The question is, should one adopt that sort of technology in a clinical trial setting where you at least can collect the data.
DR. LIGHTDALE: Yes. I'm just going to finish my own thought, though, and then I'll call on you, Hannah. Actually -- I'm sorry -- I'm

Page 173

curious about Dan. But it strikes me now we're starting to get repeated measures within each patient, and we may need some statistical workaround in how you get at that.

Then also, to the biomarker discussion, obviously, a lot of stats right now on how to start to wade through just too much data. So I don't know. Maybe some statistical thought on how to design studies gets important.

Hannah?

DR. WUNSCH: Yes. Actually, this follows on that. In critical care, there's a lot of interest in trying to predict who's going to fail extubation. It's in early phase, but looking at sort of patterns and variability in respiratory status prior to extubation.

There's been a little bit of mention of that. I'm not sure it's the answer to what we measure right now, but it may be something for the group to think about in terms of what to recommend for what should we be looking at.

Is it about some of these patterns, as

Page 174

people have mentioned, rather than any specific transient event that we're ultimately, maybe 10 years down the line, going to be most interested in, in terms of what is a harbinger of a true bad event, which I think is what we're really trying to get at.

DR. LIGHTDALE: Yeah, Dan?

DR. SESSLER: Albert was kind enough to mention our paper on Ward hypoxemia. I think there were two lessons that came out of that. What we found was that hypoxemia was common, severe, and persistent. So first lesson might be that if you don't look, you won't see things.

That gets back to my plea for continuous monitoring. It's easy to do now. It should be the standard of care. But the more serious problem, which John just mentioned, is that we don't know what this means. And everything we're talking about here is just a marker for the severe events that we're concerned about, and we simply don't know what that link is. And that really limits our ability to come up with any convincing

Page 175

recommendations for what constitutes an important outcome.

DR. LIGHTDALE: Suzie?

DR. KARAN: On the lines of the continuous monitoring and patterns, does anybody on the panel or in the room have any experience with pulse arterial tonometry?

DR. LIGHTDALE: With what?

DR. KARAN: Pulse arterial tonometry. It's with the sleep that you're using as a surrogate measure of respiratory arrhythmias. Okay.

DR. GAN: How does that work?

DR. KARAN: I'm not actually sure. I'm curious. And somebody who has a brand of a device to measure respiratory arrhythmias during sleep has told me that it definitely correlates with some degree of respiratory arrhythmias during sleep. I don't have a very good understanding of it yet, so maybe this would be a good group to run it by.

I think it's got some relevance in people who are not beta-blocked, and measures some kind of sympathetic tone or response in response to a

Page 176

respiratory arousal.

DR. GAN: Well, there are a number of monitors as has been -- again, this recent ASA, for at least 3 sympathetic monitors in one fashion or another, that potentially might be useful in looking at a certain degree of sympathetic activity, which probably when one was part of breathing came into that algorithm.

DR. WARD: It also mentioned looking at some bio stress markers for sedation. And I recall a dexmedetomidine study that I did years ago, where we actually majored a lot of bio stress markers, particularly catecholamines, epi and norepi.

We didn't get those results back for days afterwards. So we had discharged all our subjects 12 hours after they received the compound. And when we eventually got the epi and norepi back, we discovered that, at time of discharge, it was unmeasurable in every single subject, that they had no measureable levels of either epi or norepi at the time of discharge.

DR. GAN: One of the questions maybe I could

Page 177

ask Albert -- and you know that we did respiratory arrest project with Frank [indiscernible]. And what obviously that is, is sort of extreme when somebody has a respiratory problem and then had a respiratory arrest.

Do you think we are at a state where devices can sort of help to try to prevent, or mediate, sort of predict some of the people who are at risk for this catastrophic event, maybe hours or maybe days before that?

DR. DAHAN: Of course, I'm not sure. But imagine what you're saying means that you need to have a device, if there is a device, in each bed in your hospital. But it is possible. I know, in the U.S., there are hospitals that now are buying systems for every bed.

We still do not know if we can predict this. What I'm doing are small, very well controlled, high-dose opioid studies in otherwise healthy young and elderly volunteers. It might well be that in patients, there are so many co-factors, that this predictability is reduced tremendously by whatever

Page 178

other underlying disease they have, drugs they use like beta blockade, et cetera, et cetera.

But it's worthwhile to look at that. And Dan already showed that it's possible to measure for hours. The longest period that I've been measuring now continuously is 24 hours in our ICU and in the ward.

So it's possible maintenance for days, as these were patients that were not post-operative. These were patients in the hospital. And what we showed, as I remember it well, that the probability of a cardiorespiratory collapse, required resuscitation increased with both the use of opioids and benzo.

DR. GAN: Perhaps those people who are on opioid and PCA post-op should all have ketamine infusion.

(Laughter.)

DR. DAHAN: I can tell you, that's something that we're doing right now.

(Laughter.)

DR. DAHAN: So it's absolutely true. I use

Page 179

a lot of ketamine also in the ward, but it's really low dose, and I'm not really sure if this was still -- it's just to induce opioid sparing.

DR. WEISS: When you're saying low dose, like 1 to 2?

DR. DAHAN: 0.1 milligram per kilogram.

DR. WEISS: 0.1, so very low dose.

DR. DAHAN: Very low.

DR. WARD: That's not the racemic. That's the --

DR. DAHAN: Esketamine. So you have to double or triple it with -- I'm currently doing a study to compare the analgesic effects of the racemic and the ketamine. We know the comparison for anesthesia, but not for pain relief. I'm doing that right now.

I was surprised how low the potency is of racemic ketamine compared to esketamine. Esketamine is really a more potent drug than the racemic mixture. So I will tell you, in the future what dose you need to use.

AUDIENCE MEMBER. To which patient do you

Page 180

give ketamine?

DR. DAHAN: I use it for large abdominal surgeries, patients who do not have an epidural. They get either IV PCA morphine or sufentanil oral sublingual tablets currently with or without ketamine low dose. So we have this system from ExceleraRx [ph], which allows you to give a little bit of sufentanil under the tongue.

DR. WARD: A lot of the conversation of both natures has been about our inability to come up with a biomarker that really predicts outcome. And we don't have any type data, but I think we go back to the midazolam story and look at the study from Peter Bailey group that I showed.

We know that midazolam and opioids create much more frequent problems of respiratory arrest than were expected pre-approval. And if you looked at Peter's study, which was after approval, then there were very few apneic episodes separately. There was none with midazolam alone. There was a few with fentanyl alone, and there was almost 100 percent with the two together.

Page 181

If you looked at that data and you said -- ahead of time, you would say, okay, if you use midazolam alone, it should be okay, but if you use it with an opioid, you're going to probably have more true adverse outcomes, coming back to you? What's an event? What's an adverse event? What's an adverse outcome?

So I think we do have some limited data further on the respiratory system that more frequent apneic events in a phase 2/3 clinical trial is probably going to be predictive of more adverse outcomes when used clinically. But it's not a tight connection.

DR. LIGHTDALE: That, I think, again gets back to, you're capturing small little things that don't see very relevant, but perhaps they're predictive.

DR. WARD: Right, but perhaps there might be.

DR. LIGHTDALE: Yes.

DR. WARD: And I think we're limited by -- we've got these clues, but we're limited by

Page 182

not having a very tight connection there. And part of the limitation comes from the rarity of true adverse outcomes.

So I think there's a real need, that if we could devise some studies there, that would make that connection much tighter. What should we be measuring in a clinical trial, that we have more confidence in correlating with adverse outcomes when used clinically.

DR. LIGHTDALE: Anybody else? Comments? Yes, Aaron?

DR. CONWAY: So from what I read from the first report of the efficacy things, that we came up with recommended measures for the different domains of efficacy, and we didn't really get into like what would be clinically important differences between the groups within those measures.

So I'm just wondering, it seems to me that we would measure ventilation, oxygenation, nausea, those sorts of things. Are we going to decide what would be a clinically important difference in ventilation between the groups from today?

Page 183

DR. SESSLER: Decide on what basis?

DR. CONWAY: Would it be like a physiologically significant difference or a safety difference?

DR. WARD: Yes. I think that's kind of the next step, what I was just saying. Peter defined apneic episodes in a certain way. We have no idea if defining those differently would have made a tighter correlation or not. I mean, I think that goes along with defining the event and the severity of the event, how does that predict an adverse outcome.

DR. LIGHTDALE: What is the barrier to simply collecting continuous data in every sedation and dumping it?

DR. WARD: Right. And I think for clinical trials, that's exactly what we need to be doing. Just looking at the pulse oximeter and looking at clinically how the patient is breathing is not adequate, and recording the pulse oximeter every 5 minutes and the blood pressure every 5 minutes is not going to be an adequate way to get enough data

Page 184

that we can finally look at this.

DR. MINER: I think, too, is differentiating background noise from important -- if we have a 6-second apneic spell, that's pretty important if the person got a hypoxic from it or aspirated during it. But otherwise, it's just a pause in breathing. It could be nothing. So I think we end up with a lot of background noise when we collect. I agree we should collect continuously. I just think sorting through it is difficult.

DR. LIGHTDALE: But doable, yes.

DR. SESSLER: We need a very large cohort here to make sense of this because if you're looking for rare events, you need to study lots and lots of people, far more than will ever be in a clinical trial.

So what we need is continuous monitoring in a fairly sophisticated way that is monitoring lots of different parameters continuously in tens of thousands of patients. And until we get those data, we won't be able to link these markers with the events that we care about. But doing a quick

Page 185

sample-size calculation, you get numbers between 10,000 and 20,000.

DR. WARD: Yes. That even sounds small.

I think what we'll do is break for lunch. If we could get back at 1:00. I want to make the discussion this afternoon. We'll do a 2-hour, 1:00 to 3:00, and make sure we're through at 3:00. There will be sandwiches and things available to take with you as you go to your airplanes.

I want to focus on what are the recommendations that we can -- maybe "consensus" is not the right word, but maybe have some level of agreement on that I can start putting in the paper.

So I've been putting together a list of some possible recommendations. If you have some, would you please write them down and give them to me during lunch? And I will put them in the slide that's got the list of things, that we can see if we can get some agreement on. Thank you.

(Whereupon, at 11:50 a.m., a lunch recess was taken.)

Page 186

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

(1:01 p.m.)

Group Discussion

DR. WARD: Okay. In order to keep everything running on time, what I had laid out on the schedule for this afternoon was always a little bit tentative because I didn't know where we would be at this point in the meeting and what we would still have to discuss.

So we're really going to, as I mentioned earlier, combine the last couple of sessions, and we'll be through by 3:00. Val is going to tell the hotel staff that people will be needing taxicabs at 3:00, so she'll arrange for having extra cabs out there for those of you who need to catch a cab to the airport.

Once we get through, I'm going to back and get the transcripts of this in a couple of weeks, and see if I can't put some stuff together, maybe with a consensus or recommendations like we did on SCEPTER I. And those of you who would like to participate as authors, let me know, or I'll send

Page 187

it around to everybody. And if you want to participate as an author, let me know, and if you don't, that's okay, too.

But please help me out getting the paper written. It'll probably come around as a true draft and not as a final, final, final draft. So don't expect it to be perfect the first time through, and I'm looking forward to a lot of input.

It's always difficult with a big group. Having gone through the big group last time, it's too big to do a round robin, but it's a little difficult when I get all the input in at the same time. So use track changes on it, and I'll see if I can -- so don't feel bad if you send this in part on modification, and then when you get the next draft, it's not there. I didn't do it on purpose. It just kind of fell through the cracks. So put it again, and we'll work our way through it.

Since part 1 was published in Anesthesia and Analgesia, we'll probably send it there again. I've had a nice discussion about where to send these papers because it is multidisciplinary, and

Page 188

maybe the anesthesia journal is not necessarily the best place for it for everybody to be reading. I don't think it's a New England Journal kind of paper, but maybe PLOS or one of the more general kind of journals, so we can talk about that.

I think this one -- because Anesthesia and Analgesia did publish part 1, I'll probably send it in there for part 2. Because they let me call the first one on efficacy part 1, it sort of gave them a little bit of a commitment.

(Laughter.)

DR. WARD: Okay guys, here's part 2. Remember, you told me you wanted me to call that part 1. I'm going to call this one part 2, and there we go. It should be in the next month or so, to try to get early drafts of that out.

I want to thank you all for the participation in both SCEPTER I and SCEPTER II. We'll think about what we're going to be doing next for SCEPTER. This is part of Bob's action, so I have to work with Bob on what action we'd like to see done in this area.

Page 189

Please send me or Bob or Val any comments you have on the organization of this and the other meeting, things that we could do better and do differently in organizing. It's a different kind of meeting; it's not a CME meeting where we go and listen to people talk. So this has been a great group, great networking, and good participation.

In combining this, what we thought we would do is I put together some points that we might be able to agree upon, and then Steve and Mark are going to talk about some issues that they have. I thought we would get this panel up here with Albert, TJ, John, Jim, and David, up here as kind of leaders of the discussion along with Steve and Mark.

So let me go through this one first, and then when Steve and Mark come up, we'll all become part of the panel, but I think it's a panel of the whole.

So things that we can agree upon, points we can agree upon. What I think I want to do with this is go through them, not necessarily, unless

Page 190

somebody starts jumping up and down, spend a lot of time on them, think about them, and then we'll go back to ones -- make a note of ones that you think are great or ones that you think are horrible. And then in the panel discussion -- we're going to have two hours now to go through all these -- we can back to them after Steve and Mark give their presentation.

So I've got some points that a couple people gave me. These are from my notes that I've taken. So the disclaimer is all the errors are mine. This is how I've kind of interpreted it, and certainly it needs to be more of an agreement among the group.

I think the scope, from what I've been hearing, covers all the phases, 1 to 4. It may make the paper writing a little difficult, because things that might be happening, things that particularly Albert and I talked about, are more phase 1 kinds of things, and some of the other things more phase 4, but that'll come out in the paper. But I think I heard things that would cover

Page 191

all phases of clinical trials.

Physiological profile of new compounds should be well-characterized in phase 1 clinical trials, particularly the respiratory, cardiovascular effects, even though we don't know that it's going to necessarily have a strong connection to outcomes.

The MEDRA/TROOPS system should be used to classify adverse events. Data should be collected in ways that facilitate pooling for meta-analysis. For many outcomes, both severity and length of time data should be collected, and we heard that as diverse as saturation and nausea, particularly for the length of the time.

The relationship or lack of relationship between events, adverse events, and adverse clinical outcomes should be defined as much as possible in a clinical trial.

When clinical intervention is used as an outcome, it should be procedure specific. It's been brought up several times that putting nasal cannula on an endoscopy is not the same as putting

Page 192

a nasal cannula on in an MRI, and have well-defined criteria for the intervention.

Particularly in a clinical trial, this has come up several times, that if an intervention is the outcome, then the definition of what it is that triggers that intervention needs to be very well defined.

It's useful to consider the measurement instrument and the relevance instrument separately, and I think Susan was the one that gave that concept out yesterday, and I like that idea. We have a lot of things that we measure whether or not it's going to be relevant. Will it be procedure, patient, drug specific?

If you do the relevance instrument too soon, you may not measure things that need to be measured. So you can measure things, and then decide whether or not they're going to be relevant: saturation, nausea, variety of things.

The cause of the AE needs to be separated by drug or drugs, the disease, the procedure, the provider, particularly as we're doing sicker and

Page 193

sicker patients. It may be more the patient and the disease than the drug itself, and that's important to distinguish.

Decrease sophistication in the measurement techniques can be used as the compound progresses through phases 1 to 4. Expenses and things that you would -- instruments in an early clinical trial are needed, but may not be needed as the clinical trial gets to larger and larger populations.

High-risk populations need to be included in phase 2/3 clinical trials. Non-standard monitors that is cardiac output may be useful in clinical trials, but the accuracy should be defined if they're going to be used in that trial.

Major ventilation is needed in phase 3 clinical trials. Clinical observation and saturation only is not adequate for a clinical trial. In some studies, longer follow-ups, as weeks to months of adverse events is appropriate. We talked about that for neurological in particular.

Recognizing it's limitations. Patient

Page 194

satisfaction should be measured in late phase clinical trials, and TJ gave us a symptom distress score that might be a useful measure for that.

Miss target level sedation, either too light or too deep could be reported as a useful outcome measure, and that's the overlap between efficacy and adverse outcomes. And I didn't put that up there, but it also kind of speaks to, if you're going to look at adverse outcomes in a clinical trial, you have to make sure that the target sedation was the same. Probably comes out more often in pain studies, as Albert mentioned.

If you haven't achieved the same level of pain reduction, then you can't compare the respiratory effects, because the dose-response curves are going to be similar.

Failed, aborted change in agents are useful outcomes to measure. Stress biomarkers are not yet useful as measurements. Future work is needed to define the major post-procedural neurological deficits. Incidents of post-procedural delirium should be recorded.

Page 195

Patient recall may be useful, but will need to be defined in relation to expectations. Unexpected, painful, unpleasant recall. In seeing patients before surgery, I've had a lot of patients tell me, "Oh, the last time I had anesthesia, I remember the whole thing."

"Oh, what did you have?" "Oh, I had a colonoscopy." You know, well maybe you weren't anesthetized for it. So the expectations of the patient as far as recall is a very important piece of it.

The difficulty relating the occurrence of -- we talked a lot about this -- for lack of a better word, of biomarkers, desaturations, et cetera, in phase 3 trials to rare SAEs mandates postmarketing studies and data collection.

So those are kind of the points that I came up with. I don't know if we agree upon them, but I would like to spend the next time thinking about some of these and some of the ones that Steve and Mark will present, because those will form the basis of the paper as I write it.

Page 196

So Steve and Mark, and if I could have the rest of the panel members, Albert, TJ, John, Jim, and David.

DR. GREEN: So this will, no doubt, be a little controversial as we're trying to put it all together, and I'm sure there won't be full agreement on a number of things, but at least we can debate some ideas.

What we're going to do is go through our -- just kind of use the TROOPS concept as a way to organize it by organ system as we do our discussion. We have slides.

DR. ROBACK: Thank you, Rob. I'd really like to -- Steve and I would like to thank Drs. Ward and Dworkin for the opportunity to participate in this really, really fun event. It's been really extremely interesting. And as we look at the breadth of ideas and experience, and perspective from this group, it really brings home the need for this kind of a multidisciplinary approach.

Steve and I both have done emergency

Page 197

department studies, largely of ketamine, and investigating adverse events with various different ketamine and ketamine adjunct combinations.

That's when we kind of started thinking about adverse events. I was part of Maala's consensus panel on the Quebec guidelines, and that's when my thinking about adverse events began to evolve, and from the threshold duration to the intervention type model.

Maala presented her work to Joe Cravero and to John Berkenbosch in the Society for Pediatric Sedation Group. And we had further discussions, some quite spirited, about this concept. And after the thoughts yesterday and today, for me this is kind of been a déjà vu all over again type of a discussion, which I've really enjoyed.

So as a quick summary, the goal of this tracking and reporting outcomes of procedural sedation tool was to have a standardized approach, something that was very practical, intended for everyday use. More of the surveillance, safety surveillance, phase 4 type of trials, but also can

Page 198

be used for greater sensitivity and specificity in phase 1 and 2 trials, interested in adverse events, interventions performed, and then outcomes.

Again, this should be for any tool that is proposed, really needs to be for all patients in all settings, as has been discussed.

We talked about severity, and it was important to rank it. We have our sentinel events in red. Those are life-threatening. They'll warrant immediate reporting within sedation care systems and the highest level of peer review. Intermediate outcomes and interventions were serious enough to endanger patients, if not promptly managed, or reflect sub-optimal sedation quality or patient experience.

The thought about our discussions in the last couple of days about the outcomes for sedation and efficacy, I think it's really important that we emphasize how do we measure sub-optimal sedation quality, and how do we really take into account the very important considerations for patient experience.

Page 199

These intermediate outcomes and interventions then represent things that should be reported in a timely fashion and periodic peer scrutiny.

Then the final piece in green are the minor outcomes and interventions, which we'll go over, but this is really more for peer research or for certain specific clinical situations.

This is the primary tool. You can see we have them broken into organ systems in the first column, airway breathing, circulation, neuro, and then the sedation quality and patient experience, much like what was presented earlier today.

We have the severity of the outcomes, intermediate, and sentinel, the suspected cause of the outcomes and interventions. Then each one is related to the organ system of note. So we thought we would go through these and get people's thoughts about our current proposed approach.

DR. GREEN: Of course it's worth nothing that this is geared for clinical use, and what we're talking now about is research. So this can

Page 200

guide the discussion, but obviously there will be differences.

DR. ROBACK: So under airway and breathing, the point was made earlier that they're definitely distinctly different, and I think that's a very important point. But the things that we felt, our group felt, were important, were patients that had problems -- apnea, respiratory depression, laryngospasm -- which resulted in them receiving positive-pressure ventilation, unplanned reversal of opioid or benzodiazepine, tracheal intubation, neuromuscular blockade, and then of course, pulmonary aspiration.

Yes, please?

DR. RIKER: How do you view interventions like requiring oxygen post-procedure that you didn't require before, or need for bronchodilator therapy, or things like that?

DR. ROBACK: So those are obviously important things to consider, and we've included them in the second part of this, which would be the more in-depth for research part of the tool.

Page 201

DR. WARD: What about the bronchodilators, intermediate or minor? I mean, sometimes somebody really gets a bad asthmatic episode. That seems to be more intermediate kind of intervention.

DR. ROBACK: So I think that that's a really good point. And I believe we have it under minor, but if you're needing to do it aggressively, it's probably going to impact their length of stay, are they going to now require hospital admission, or are they going to require intensive care. So it would be captured under those, which were more of our intermediate outcomes.

Yes?

DR. PRATIK: Mark, where does airway obstruction, which requires oral airway, does that come under minor? Where does that fit in?

DR. ROBACK: That's also in the research tool. Under minor, we would have the visual or the tactile stimulation or breathing cues. Here we have under respiratory then, airway positioning, suctioning, capnography. Then here we have your oral or nasal airway under intermediate.

Page 202

DR. GAN: If you can, sorry, go back to that first slide, and I'm just trying to in my mind distinguish between tracheal intubation and neuromuscular blockade. I assume you have both up there to signify that one may be different from the other? What are you thinking about?

DR. GREEN: Yes. Certainly, they overlap, and I guess there would be a scenario in which the patient could be intubated without neuromuscular blockade, or another scenario in which neuromuscular blockade is done, but then the intubation doesn't ultimately occur. So we just wanted to cover both of those bases.

DR. GAN: So neuromuscular blockade was happening without intubation? How could that be possible?

(Crosstalk.)

DR. BERKENBOSCH: It breaks. They start ventilating, because the succs will wear off fairly quickly, and you may not end up progressing to intubation. Probably I would bet that that's a very unlikely scenario. But most of the time that

Page 203

you chemically paralyze somebody, you're going to be intubating them.

DR. GAN: I'm just trying to see why there were two separate --

DR. ROBACK: I think that one of the criticisms were that we were a little too coarse with these, and we needed to have some greater definition of what occurred. Tracheal intubation is a physical placing of the tube in the airway, and then there's also a pharmacologic response, and we wanted to acknowledge both.

DR. WARD: You didn't include any superglottic airways there, which I would think -- in an anesthesiologist's hands, if I'm having trouble, I'm going to put an LMA in a lot sooner than I do tracheal intubation. So somewhere in between positive-pressure ventilation and tracheal intubation, I'm going to put a superglottic airway in.

DR. ROBACK: Right. So we thought that positive-pressure ventilation would be acknowledged in that, and then in the research tool have the

Page 204

LMA.

DR. GREEN: Actually, LMA is in the footnote under positive-pressure ventilation.

DR. ROBACK: Oh, that's right. No, that makes sense.

DR. WARD: I might put an LMA in without positive-pressure ventilation.

DR. ROBACK: Rob, can you bring up the -- okay, and then go over airway -- under positive pressure to -- right there, the next one over, positive-pressure ventilation. Yes.

So there it defines positive-pressure ventilation. It includes BiPAP, CPAP, and LMA.

DR. WARD: Like I said, I would most likely put LMA in without positive-pressure ventilation.

DR. ROBACK: Oh, without.

DR. WARD: Because if I just have some airway obstruction, they're trying to breathe, I'm not going to do positive-pressure ventilation.

DR. ROBACK: Right, right, no, I see.

DR. WARD: I'm just going to do an LMA, and they'll start breathing.

Page 205

DR. ROBACK: Yes.
DR. WARD: Continue breathing.
DR. GAN: Anything that would fit into that middle column there, LMA, tracheal intubation, maybe common aspiration.
DR. KARAN: The same way I wouldn't consider CPAP or BiPAP necessary positive-pressure ventilation. It's just support. It'd be in a different category, and maybe that would just be kind of in the same area as an oral airway or a nasal airway.
DR. GREEN: Yes. I mean this is not a final version, so we're wide open to suggestions.
DR. CARLSON: So are you looking at intermediate as sort of precursor events, significant precursor events, and sentinel events as actually having patient harm? Is there any correlation in that, or is this harm issue, true harm, or a precursor of where you prevent harm in this sort of taxonomy at all?
DR. ROBACK: Rob, can you go back to the other one, please? I'll just bring it up again.

Page 206

So there's our formal definition. The sentinel event is life-threatening, whereas the intermediate one serious enough to endanger if not promptly managed?
DR. CARLSON: So in harm taxonomy, you might have temporary harm, but it would be fleeting in the sense of -- intubation, you don't expect to be intubated as harm. But you're saying within a relative period of time, things are normal. Where the sentinel events would imply, for the most part, pretty significant patient harm.
DR. GREEN: Well, not necessarily harm. I mean, let's say a patient with laryngospasm was intubated, but then extubated 8 hours later, they don't have any permanent harm.
DR. CARLSON: Well, that's temporary harm. You intubate somebody, you don't intubate I think in harm -- maybe safety would be harm, but it would be temporary harm, and probably intermediate temporary harm by most taxonomies.
DR. GREEN: Yes, okay. So this isn't stratified by the level of harm, but yet the level

Page 207

of worry such that quality improvement systems should pay attention to it.
DR. BHATT: I just had a question on your next slide about the three sentinel outcomes. I realize that you've got the outcomes and the interventions there, and most of them are interventions, but you do have a couple of outcomes like pulmonary aspiration and neurologic deficit.
So it looks to me like you guys were looking for a home for those serious events, but it doesn't -- you're not looking -- it doesn't map well to a suspected etiology because it would be a suspected etiology for a grave intervention.
So is there a benefit?
DR. GREEN: Yeah. At first, we tried to construct this entirely out of interventions.
DR. BHATT: Yes.
DR. GREEN: But yet, then as a failsafe there's certain things that obviously are going to need scrutiny even. You can imagine a very rare scenario where someone just didn't intervene, and this is to make sure that all of those are

Page 208

captured.
DR. GAN: In a way, they almost need to be, in my mind, separated because, for example, one intervention following pulmonary aspiration is intubation.
DR. GREEN: In that case, two boxes would be checked there.
DR. ROBACK: We went round and round on this model, and we're soliciting a better approach or a better way. Because like you said, essentially everything in these columns, the two middle ones, is something that was done. Pulmonary aspiration and death are things that occurred.
DR. GREEN: I think it's worth pointing out that the original version of this was about three times more complicated, and we realized something complicated like that would never really be used or done. So we did our best to trim it down as much as we could.
DR. ROBACK: We always were considering the context of it's embedded in the electronic health record or in whatever your usual documentation is

Page 209

for the sedation event, and it's just part of what you do. It should take less than a minute.

DR. BERKENBOSCH: I actually kind of like the idea of keeping them both in the same column, even though one is an intervention, one is an event or an occurrence, because I think you -- we can argue about which one should be in the intermediate versus sentinel.

But the idea here is that you're capturing data that I think we all believe is probably really important to capture in a way that is simple. And when you get to the point of entering it into the database and analyzing it, you can tease out which were the events and which were the interventions very easily, just by the way you code the electronic database. So I don't have a problem with that.

I think it's more -- I think we're getting hung up in the definitions a little bit and almost getting a little bit lost in the forest for the trees kind of deal, where what we really want to do is get a tool that's going to capture the events of

Page 210

interest, and do it in a way that's simplistic, doable at the bedside so that if you are doing a clinical research project in multiple different areas, you've got the universally acceptable tool that is easily doable by the principal investigator, as well as the other sedation providers who may not be as invested in the research, but still collecting data that is going to be valuable.

So from that standpoint, I think a simple tool is something that's incredibly important, which is similar to what we've experienced with the Pediatric Sedation Research Consortium. You have to decide how much granularity do you want versus how doable is it, and doable and collecting the important outcomes is probably more important.

DR. ROBACK: I think we've all tried to do work where we wanted every bit of information as possible, because clearly it's all important, but at some point, you're just not going to get it, because people won't take the time to fill it in.

DR. BERKENBOSCH: Right.

Page 211

DR. RIKER: How do you currently within this tool look at apnea, a respiratory rate of 4, or a heart rate of 40, or a blood pressure of 72, when no intervention is performed versus if it is? Because one of the things we wrestled with is these isolated vital signs that may or may not be clinically significant, and you may have already solved the problem.

DR. ROBACK: Well, that is the crux of it, isn't it? I mean you've hit the nail on the head. Rob, if you could bring up the other one again, and if you could go over to apnea on the far right corner column.

We really considered that in order for apnea to be meaningful, it would have to be enough respiratory depression and hesitation in ventilatory effort that you needed to do something.

DR. WARD: And they're checking apnea without having necessarily checked any of the other boxes, the way that's defined. You could just --

DR. BHATT: That's respiratory depression.

DR. WARD: For respiratory depression.

Page 212

DR. ROBACK: Well, but it does say, without positive-pressure ventilation. So you'd have at least that to be considered apnea.

DR. WARD: Just without positive-pressure ventilation so --

DR. ROBACK: No. It responds to brief --

(Crosstalk.)

DR. WARD: I wouldn't necessarily have anything else checked. I could check respiratory depression.

DR. GREEN: Correct. So in this case with --

DR. ROBACK: Oh, that's on respiratory depression; it's not on apnea.

Rob, can you go up to apnea? I was confused there. Right. So apnea requires that you have positive-pressure ventilation checked, whereas respiratory depression does not.

DR. GAN: Like the previous slide, if that's going to be standalone, I think all those asterisks or definitions probably need to be a bit more defined. And also, if you go to a previous slide,

Page 213

this paradoxical response, do we know what -- I mean is that clear enough for people to say what?
DR. ROBACK: So that's on the green one.
Rob, if you could go back to the --
DR. GREEN: Yes, there are definitions that you hover to see.
DR. ROBACK: Yes. The paradoxical response is on the research tool, right?
DR. GREEN: Yeah.
DR. ROBACK: Go down. There you go. So if you could hover over "paradoxical response." Right there. Thanks, Steve. So that's how we defined it, unanticipated restlessness or agitation in response to sedatives.
DR. GREEN: So if this were embedded in your electronic medical record, you could quickly hover over and reference definitions.
DR. GAN: How is that different from the one below?
DR. ROBACK: Unpleasant recovery, agitation? Well this is in response to the -- in a sedative drug, what you expect to occur as the drug was

Page 214

given, whereas the unpleasant recovery action is during recovery.
Rob, could you down one, please?
DR. BERKENBOSCH: The paradoxical could occur both during induction and during recovery.
DR. ROBACK: Yes, I mean most likely it's going to incur during induction.
DR. BERKENBOSCH: There's probably some overlap.
DR. ROBACK: Yes.
DR. BERKENBOSCH: I only wonder if -- there's a certain -- I was just thinking of the paradoxical co-responses to benzos. Are they really unexpected? They're certainly not in -- you would hope that most people don't, but you know that there's going to be a certain percentage of kids, particularly, in my experience, adolescents that -- again, I don't want to get into big --
DR. ROBACK: No, I tell every patient.
DR. BERKENBOSCH: -- wonder if that's -- if that terminology is going to prevent people from checking it if there's a paradoxical response when

Page 215

they said, you know, I know there's a certain incidence of it. This kid did. Or this patient -- I'm a pediatrician and we always to talk to kids. Sorry.
DR. GREEN: The things in the etiology only apply if there's something in the intermediate or sentinel column. So in this case, the paradoxical response would only apply if there were patient family dissatisfied or hospitalization or escalation of care. If it were less than that, it wouldn't be tracked on this.
DR. ROBACK: So you would expect that if they had a paradoxical response that someone would be dissatisfied and that sub-optimal sedation quality would be realized.
DR. GREEN: If it were of sufficient magnitude. Mike, you've had your hand up the longest.
DR. IRWIN: Go back to the airway a bit, because maybe you should have -- upper airway obstruction should be in there, rather than even -- maybe it's much more common, say, than

Page 216

laryngospasm. And simple maneuvers such as obstruction --
DR. ROBACK: Rob, can you go up?
DR. IRWIN: -- or pharyngeal or nasal airways. That type of thing.
DR. ROBACK: So we have apnea, respiratory depression, and laryngospasm.
DR. IRWIN: No upper. I mean, one of the commonest things I think with the upper airway obstruction actually, because of loss of muscle tone. And then usually we'd use a oropharyngeal or nasal airway.
AUDIENCE MEMBER: That's the function of the more minor events, though.
DR. ROBACK: So we have those in the secondary -- this is the primary and then the secondary part, there's airway obstruction. An airway obstruction that really was significant, that required you to perform positive pressure, you would capture it here as far as the intervention goes.
DR. GREEN: So if all you needed was a chin

Page 217

lift or something like that, it wouldn't be considered an intermediate outcome.

DR. IRWIN: But you wouldn't be able to put it in the suspected etiology. There's nothing to take for that.

DR. ROBACK: So you're proposing that we put partial airway obstruction here.

DR. IRWIN: Well, yes, to me that's probably the commonest airway incident whenever you --

DR. ROBACK: But I guess we put it in the secondary tool because it's one of those things that falls into the category that this is just good sedation care. You gave the guy a jaw thrust. I mean, that's what we do. And we're trying to really think about outcomes and interventions that are out of the ordinary or are more meaningful. We would capture -- it's still important to identify partial airway obstruction and the jaw thrust you provided, but that would be in the secondary tool.

DR. CARLSON: As suspected, etiology, it's not as -- I mean, it does seem in the sense that apnea and obstruction sort of fit together for

Page 218

better checking. We did positive-pressure ventilation, check mark, obstruction. Or is that just -- you think you capture it -- you'll get to the secondary part. Because apnea is most often sedation obstruction.

DR. ROBACK: Right. I mean there --

DR. CARLSON: Or does it fit as a subcategory within apnea?

DR. ROBACK: Rob, can you switch to the other slide, please?

DR. WARD: I was going to say, the laryngospasm is often misdiagnosed I think.

DR. ROBACK: That's a good point.

DR. WARD: A lot of times I find residents call something laryngospasm when it's really not laryngospasm.

DR. CARLSON: I find that most of the time it's simple obstruction, and laryngospasm is the term of the day when you're talking about ketamine particularly.

DR. WARD: Right. And you see it in kids. You almost never see it in adults, and this is a

Page 219

form not just for kids, right? So I think airway obstruction is a much more general term than laryngospasm would be.

DR. ROBACK: So I think our committee felt like if it was partial airway obstruction, that it should be managed by one of these interventions here. Although you didn't mention using the LMA without positive pressure as one way of doing it that we hadn't considered.

DR. KARAN: I wonder also differentiating between invasive and non-invasive ventilation, in which case LMA and tracheal intubation would be in one category, and applying a mask with just some CPAP or BiPAP would be in another category.

If I'm using an LMA for sedation, and it's unplanned, I would think that's a little bit more severe -- not in the ED but for sedation -- than just applying a mask and apply some BiPAP or CPAP for it.

DR. BERKENBOSCH: So that begs the question to me, and maybe a show of hands, how many people would call placing an LMA invasive ventilation?

Page 220

(Show of hands.)

DR. BERKENBOSCH: How many would not?

(Show of hands.)

DR. KARAN: Invasive ventilation in the setting of it was unplanned for your sedations, as a rescue technique. I mean placing an LMA for general anesthesia I don't think is so invasive because that's planned, but if you weren't planning to do that for your sedation --

DR. BERKENBOSCH: My point being, I think if you state -- I guess if you clearly define the terms upfront, stating invasive/non-invasive ventilation, you'll capture the -- I think they're different interpretations of what the placement of an LMA would be, and that would be something that would be potentially ambiguous if it's not clearly defined upfront.

DR. ROBACK: Well, we defined it as part of a subset of positive-pressure ventilation, and maybe that needs to be broadened.

DR. BERKENBOSCH: Maybe.

DR. WARD: So I want to come back to the

Page 221

clinical trials question, because I think you've put a lot of work in a great QI tool. I particularly like that it's connected to MedDRA, because I understand all these terminologies map to MedDRA terminology, which is important for the FDA.

This is a question for the group whether it's a show of hands or whatever. How useful is this tool if you were -- and I'm assuming we're talking about a late phase 3 or phase 4 study -- if you were just putting a late phase 3/phase 4 study together, and you'd have to have some sort of sedation score, some efficacy score, maybe OAAS or some of the other scores we've had, and the only other data you collected was this TROOPS data, would that be adequate for a late phase 3/phase 4 multi-institutional -- because I think one of the advantages of this is the ability to collect the data online with a website in a multi-institutional.

Is this a useful tool for that?

DR. DAHAN: Well, what I wonder is suppose this happens once during the case, and you tick it

Page 222

off? But now suppose it happens recurrently? How do you cope with that, when you're doing your -- so you changed your -- for instance, you change your propofol dosing, but still the patient has an apneic event. You need to fill in several forms now.

DR. ROBACK: Well, we do have the -- at the bottom there's the "other," and I guess we could specifically note that if you have recurrence of any of these adverse events, you could record it in the "other" category. It's an important point, because clearly if you're doing bag-valve mask for 5 seconds or 10 seconds or 30 seconds, that's one thing. But if you're doing it repeatedly over the course of 10 minutes, that's a much different encounter.

DR. DAHAN: For clinical purposes, this is great, but for research purposes, it's very restricted in the sense that things might happen, things might change, especially patients that do extremely poorly.

DR. GREEN: Yes, I think if there were

Page 223

recurrent apnea that were clinically important, at some point you would give a reversal agent if applicable, or you would convert it to tracheal intubation.

I think just one concept from a research angle, obviously sentinel outcomes are going to be rare enough that they can never be your primary outcome. But what you could say -- say it was a trial of propofol versus another propofol-like drug with a large trial, your main safety outcome could these airway breathing things. Either a composite endpoint of positive-pressure ventilation, reversal agent, tracheal intubation, neuromuscular blockade, or pulmonary aspiration, comparing that between the frequency of those.

DR. ROBACK: Frank?

DR. DEXTER: Yes, I'm struggling with that, because let's suppose that different people are going to intervene at a different level of saturation. Then unless it is specified in the study what hypotension you're going to treat and how you're going to treat it, what level of low

Page 224

saturation you're going to treat, or hypoventilation you're going to treat, I don't follow necessarily that that would be sufficient. As long as it's designated, sure, and prespecified. Then you have a randomized trial because now it's between the groups.

Similarly, in terms of being sufficient, I think that the continuous measure of vital signs is a logical endpoint in addition to the intervention, so you can understand why did the clinician do the interventions that they did. What information data did they have?

DR. GREEN: Yes, again this is intended to be clinically oriented for what people are doing in actual practice. For research, it would be great to have an endpoint like this, but underneath it to have all that data, to know when they intervene. I don't think you're going to get people to agree on that they're going to give positive pressure at any specific number.

DR. DEXTER: So I would say that even in regular [indiscernible] practice, if I were to do,

Page 225

for example, is have a large amount of anesthesia information management system data like we have, we see these data all the time, and it isn't possible, because we can do study after study we've done showing that people will treat based on different vital signs. Some people treat and others will treat, and we don't know that that's better or worse; it's different.

If we were to have a similar scale for general anesthesia or something like that, we would still say what does that mean necessarily? You combine it with the vital signs. Then you can understand, at least describe quantitatively what the drug is doing.

DR. GREEN: But yet it's more than vital signs that determine when you're going to intervene. It's looking at the patient and multiple number of factors. I think Maala was first.

DR. BHATT: I just wanted to comment about that. I think that it's true, people will intervene for different things. But I think if

Page 226

you're looking at intermediate or sentinel outcomes, I think that people will be more uniform in terms of when they intervene.

We had six different centers with probably over 200 providers providing us -- not 200 providers. That's probably an exaggeration. A lot of providers. It made no sense just providing sedation and still our rate of serious interventions or significant interventions, as we deemed them, was very low. It was 1 percent.

DR. KARAN: Were they all ED providers?

DR. BHATT: They're all ED physicians, and these are healthy people, so it's a different population.

I also wanted to make just one comment on -- and so I think if you're looking at more into variability, I think that that comes in with the minor outcomes, because people will react very differently towards more minor things. In fact, you need the actual data to decide how important those interventions were. But I do think that if you are acting on these serious things, I think

Page 227

that you can -- I think things balance out in the end. People act more uniformly for these significant things.

But one comment I wanted to make, Steve, about what you're saying about making a composite outcome. It will be different in different populations, but if you were to power an ED study based on those amount of outcomes, you will need a million children because we don't -- the frequency of these events is very, very [indiscernible].

So I would say that, I think in a healthy population of children that you're sedating, you are going to need the other part of your tool as your primary.

DR. GREEN: Yes, very true. When I said that statement, I was thinking of an adult propofol style thing. But for ketamine and for other things, that wouldn't work.

Jenifer, you've had your hand up.

DR. LIGHTDALE: I'm going to use the dreaded aviation analogy. I'm actually coming around to the concept that maybe it's okay to go with the

Page 228

fact that this is going to be a perception of whether or not an adverse event happened. That's okay. You know, we should get that. If people think something bad happened, just why do they think it happened?

As long as we have a black box, those actually are collecting [indiscernible]. I think the clinical trials have to collect that data, and they have to -- now you get say, okay, people have thought something was happening, and we can see that it was, and we can see how they kept the plane flying. You know?

DR. GOZAL: We had three things in the tools. First of all, the kind of provider. Second, the type of procedure. And maybe also there is a class.

DR. GREEN: Yes, this is just the adverse event outcomes part of your sedation record. It's not going to be all the pre-sedation assessment, all the other information that would still be part of your record.

DR. ROBACK: Keira had the same comment to

Page 229

us as we were preparing this. And you know, the World SIVA tool has all that information in the beginning, and that would have to continue to be, because like you said, how sick the patient is, their age, the procedure, what drugs you're using, all those things matter.

DR. WARD: So I guess I'll get back to the question that I kind of want to get a concept from -- I think it's a great QI tool that you've got started. It has very attractive features on being online, Web-based, if I'm designing a multicenter clinical trial. Clearly, as you just pointed out, I've got to get some demographic data collected for the clinical trial that would have to get that this doesn't go with.

But given that I get the demographic data, but I probably don't have the black box -- the data if I'm doing a multicenter clinical trial -- that's altogether different.

Would it be acceptable for this to be the sufficient measure of outcomes in a multicenter, late phase clinical trial? Because it's got some

Page 230

attractive features, but it was designed as a QI database, not a clinical trial.

DR. GREEN: I think it covers major outcomes, but as Maala pointed out, you couldn't make this your primary endpoint for a lot of sedatives that don't have high enough frequency of these things.

AUDIENCE MEMBER: But one thing I just want to point out. Isn't that a word missing under the patient, stress, family. What? Is it satisfaction or --

DR. ROBACK: Yes, that is satisfaction. That's an old version, I'm sorry. I think the ones that were distributed to you -- but again that's satisfaction of the family.

DR. WARD: So if we include the research, not just the intermediate and sentinel outcomes -- you've got another column there, right, with the research outcomes, too.

DR. ROBACK: Right.

DR. WARD: So using this tool with all three columns, so you've got more pieces of data that

Page 231

maybe you could go ahead and show that. So we got the minor outcome/intervention. So the tool includes that.

Is this now a sufficient tool? Because we're talking about writing this paper, and I'm going to make a statement here that we should use MedDRA. I think that's kind of mandated. But then I'm going to say that there's this thing out there that is very useful because it's online, multi-institutional. And again, I'm going to say something about it's a great QI tool, and it is or is not a sufficient tool for clinical trials.

DR. SESSLER: Well, I think many of us had said that you probably should have continuous monitoring for key things. You should have some sort of black box. But as usual, it's context dependent. There is a role for pragmatic trials. Pragmatic trials are usually extremely large but simple, and you make up for the simplicity with very, very large numbers.

But in a typical clinical trial, I don't see much excuse for not continuously recording key

Page 232

things, key measures of at least ventilation.

DR. CARLSON: I think as you write this up, we have a pragmatic tool that may have some weaknesses, but it can be tweaked to fit a need. The idea of that black box is a pie in the sky for what we're trying to look at, which is thousands of patients and safety.

That black box, individual institutions, and physiology. But what do we do when we're actually looking at these outcomes, which are rare events, we do need a tool of which we can apply the multi-institutional to thousands of studies, and that will have some gaps and the weaknesses in the sense of the continuous vital signs.

Maybe it's both, but do we not move forward with that until we have a data repository and equipment to catch all this data and do this. We're all in various states of electronic capture. Most of us have it in our ORs, but actually most of us don't have it in the places that reduce sedation. It's increasing, but we don't have it in the place we're doing sedation in many places.

Page 233

DR. DEXTER: I'll try one more time. I don't understand how it's sufficient. If you were to say that you specified in a study how people are going to use the vital sign data and respond, no problem. I totally understand that not necessarily having all the vital sign data for all the thousands of patients.

But I know that as a statistically oriented reviewer, if I were to have a study where they say essentially, we have a multicenter and there are two particular drugs and one drug has more airway interventions, without knowing how people are using the drugs or what it is that they're seeing, and whether or not it's homogeneous -- you've got to show that there's homogeneity of treatment among the centers.

If you don't know whether or not it's due to the drug or whether it's due to different practices among the centers, I don't understand -- what would you do at the end of your study if have heterogeneity among the centers? I think that in that circumstance, you'd be left with a failed

Page 234

trial.

But of course there's going to be heterogeneity among centers, potentially if it's a new drug. If it's an old, established drug, I can understand why it would be homogeneous. But like when propofol first came out, I know that as a resident, people would say, give more, give more to stop patients from moving like say. That's my concern in terms of it being a new drug.

DR. WARD: If the demographic data and the use data -- dose, things like that -- were collected in addition -- this is the outcome tool -- you've clearly got to have a demographic, site-specific stratification, demographic data, dose data, in addition to this.

DR. BHATT: So what you're trying to get at is that the different techniques, sedation techniques that are different, at each of these different sites with a new drug?

DR. DEXTER: They're going to be entirely different how people are going to see certain airway -- they're going to be doing different

Page 235

procedures with the drug, to some extent.

DR. BHATT: So we can adjust for that in the analysis.

DR. DEXTER: Not if you don't have -- that's the thing. You can adjust it for the analysis, provided you have -- it's not just demographic data. You have to have the continuous sensor information to know what the person in the operating room, or the procedural location was responding to.

DR. ROBACK: But is it practical or possible to say, all right, if the patient has apnea for 3 seconds, or their pulse ox goes down to 78, that's when you can do positive pressure. I mean, clinicians are going to do what they're going to do. And I guess that's the challenge. And finding a way to -- in instances where you have the ability to capture all of the continuous physiologic data, and then coupling that to the interventions that were performed --

DR. DEXTER: If that's what clinicians are going to do, that's why you have continuous sensor

Page 236

data to study how the clinicians respond in the real world to the drug, so to speak.

DR. RIKER: So this is designed to get rid of those insignificant events that nobody cares about, and to focus on the clinically significant interventions that everyone would agree are significant. And I think what everybody's saying here is by itself, this isn't enough for what we're talking about, and we need both.

We need all of this -- you know, how often was there an apnea of X-duration, maybe with or without any intervention, or an airway obstruction and you did a jaw thrust and you're done?

People are saying they want that data included, but they really want this as well, so that we can sort through, when was it just a heart rate of 48 and nobody cared, and when did somebody give atropine or do something clinically significant? We need both.

DR. SESSLER: I couldn't agree more. You need both. This is a perfect quality instrument. But if you're doing research, you have a higher bar

Page 237

to meet. Research has to be done in a way that you would not do clinical practice, and in a way that you don't evaluate quality. You really need to measure the things that you think matter.

I'm also worried about sensitivity, because I think when you presented this yesterday, you said that adverse events were actually quite rare. You could possibly end up with false negatives if you don't have a very large trial here, in a way that you would not with continuous data, where you would see things.

DR. GREEN: Yes. I don't think any here is disagreeing with the optimal nature of having this continuous data if you have it available. The question is large trials --

DR. SESSLER: Well, I don't think if you have it available is an excuse, frankly. I appreciate that you don't often have electronic records set up in the places where you're doing sedation. But if you're doing research, you have a different bar to meet. It's easy enough now to plug a USB disk into almost anything, download it,

Page 238

hook up a computer. This is not technically complicated. So whether or not it's available, if you're doing research, you have to do high-quality work, and that includes continuous monitoring.

DR. GAN: How do you define this blood pressure change?

DR. GREEN: Can you bring up the other one please, Rob? And then go under hypertension.

DR. WARD: And then while he's doing that, when you're doing continuous -- when you say continuous morning, that's in order to connect to the intervention, correct, Dan?

So if you're doing continuous monitoring, you can back and say, okay, that was just intervention, and now I can see why the intervention occurred. Is that what you're referring to?

DR. SESSLER: Well, that's one reason. The other is to detect all the events that happened that did not provoke interventions, which I suspect are 10 times as common.

DR. WARD: Right.

Page 239

DR. SESSLER: And that's exactly what I mean by sensitivity, is that if you're looking for anything that's relatively rare, you need a very large sample size. If you're looking for things that are common, you don't.

If you have 10 times as many desaturations as interventions, you can detect potentially important differences between two different treatments with continuous monitoring in a way that you never would just by tracking clinician interventions. Recognizing that clinical interventions are important; they're measuring something different. I'm absolutely not saying ignore that. I'm saying you need to do both.

DR. GAN: So you expect people to know what 2 standard deviation from normal -- it is?

DR. ROBACK: Well, we put this because that's the American Heart definition for children. Again, what we care about is did someone think the blood pressure was low enough for them to receive bolus IV, isotonic crystallite? Did they think that the heart rate was low enough that they

Page 240

started chest compressions?

I mean those are the things -- and again, we're not going to get the continuous data, the drop where there's no intervention performed. But I guess we assume that if their heart rate went down and nobody did anything, and they woke up and they were fine, that it didn't much matter. But it does matter if you're comparing two drugs in a phase 2 or 3 trial.

Yes, Maala?

DR. BHATT: Can I ask you a question just on the continuous data monitoring? If you collect all this information and you're looking at this as separate from the patient, how do you separate new events from monitoring errors or pickup errors in the signal.

DR. ROBACK: I'll leave that to our experts.

DR. DEXTER: I think that at this point, whether it's the Cleveland Clinic, we've been doing it since the 1980s. I mean seriously at this point, we're very good in terms of the data filtering and stuff like that. The software is

Page 241

written. But as part of clinical trials, when I say experience doing that, I haven't found that -- at least pharmaceutical companies have wanted us to filter, because one of the whole challenges has been is whether or not that rare aberrant event, it was actually an artifact or not an artifact.

My experience has been that oftentimes the clinicians cannot necessarily determine whether or not it was an artifact or not. Sometimes it's completely obvious in terms of you just change the transducer and you can make a note or something like that. But there have been several situations where people thought it was an artifact, and it wasn't an artifact, with new devices.

DR. ROBACK: Jenifer?

DR. LIGHTDALE: I'm also going to say you still need the human person to say whether or not something went wrong. And I'm actually going to come back to Joe's comment yesterday about IV infiltration, that he has put that as something that he's noticed everybody complains about it.

Page 242

The truth is we wouldn't know if there was constantly IV infiltrations unless we let people write that down, like drug X that keeps sclerosing veins, or whatever, means that we better pick that up. So we're going to need the humans also to keep track of what's happening.

DR. ROBACK: And that could easily be answered in the second part. Yes, David?

DR. GOZAL: What do you mean by sedation complication?

DR. GREEN: So that would be the patient satisfaction. So if a provider's dissatisfied, or the patient or family are dissatisfied due to a sedation complication. So it's just having an etiology to go with the items over on the left.

DR. ROBACK: It could be any of the things listed above or in the next piece.

John?

DR. BERKENBOSCH: I was just going to say under cardiovascular, I think there's probably a couple things I would add. One would be intervention for dysrhythmia, whether it's medical

Page 243

or electrical conversion. It's going to be a rare event, but I think it's probably worth putting into the tool.

DR. ROBACK: Intermediate or sentinel?

DR. BERKENBOSCH: That would probably be sentinel. You could argue SBT, that's hemodynamically stable and gets a dose, but then it might not be sentinel. But I also think it's not uncommon for those of us who use dexmedetomidine a lot to give a little bit of atropine or glycopyrrolate for bradycardia, and I don't see that as something that's being captured. But yet, that's an intervention that's probably worth capturing. So medical intervention for bradycardia. We've got the fluids for hypotension, but --

(Crosstalk.)

DR. ROBACK: Vasoactive drugs?

DR. BERKENBOSCH: I'm not -- well, do you want to -- I'm not sure a lot of people would think of atropine or glycopyrrolate as a vasoactive drug.

DR. GREEN: Is atropine for dex a common

Page 244

thing? I thought it caused hypertension. I don't use the drug myself.

DR. BERKENBOSCH: We give it occasionally.

DR. GREEN: Is it necessary?

DR. BERKENBOSCH: Or I give a dose of ketamine sometimes.

DR. KARAN: I use dex a lot and I've never used atropine once.

AUDIENCE MEMBER: Yeah, me too, I use it all the time.

AUDIENCE MEMBER: I never use it.

DR. GREEN: I think that's why we didn't include it because that was the feedback we received.

DR. ROBACK: So if you put it in the secondary piece as an important thing –

(Crosstalk.)

DR. WARD: So I want to come back to clinical trials, though, because this is a clinical -- and we can help you design the tool, which I think is a great thing to do, but it's probably not this conference's agenda.

Page 245

So before you kind of leave this, is there any other -- Susan?
DR. KARAN: I have a practical question. Again, understanding the importance of the continuous monitoring, I'm just wondering, let's just say we're evaluating a drug that we're going to be using for mild sedation, and it's not clinically relevant to do continuous monitoring, then would you propose for a study like this that you would do it in kind of a blind way to capture it, but allow the clinical course to happen so you're actually capturing what an interventionist or what a proceduralist would be doing when you're giving the sedation drug?
Would it bias the study? If you're doing continuous monitoring, it's now right now a standard of care to do continuous monitoring for a mild or moderate sedative.
DR. SESSLER: Okay. Continuous monitoring is absolutely not the standard of care, and I'm not proposing it as a standard of care. I am proposing it for most studies, because we're doing] we're

Page 246

doing mild to --
DR. KARAN: Has it been continuous collecting versus continuous monitoring?
DR. DEXTER: I would like to see that before a drug is going to be routinely used, that in the ideal world, the clinician frankly doesn't even see all the data that's being collected. The clinician just sees the data at the rate at which they're actually going to see in routine use. And then, what you can learn is how -- when this drug is being used by clinicians around the country or around the world, what are they actually going to do in response?
So for sake of argument, with one drug, if they are not going to notice, for example, that there is multiple episodes of low saturation that are transient -- I'm making that up or something like that -- you want to know whether or not they would notice it and respond to it.
DR. KARAN: So just for semantics, it's being continuously collected, but maybe not during the course of the sedation being continuously

Page 247

monitored.
DR. SESSLER: Yes. It may not be available to clinicians. I believe we often blind our monitoring.
DR. KOCHMAN: This is not an unusual circumstance. I was on a DSMB, and I can't get into the details of it, but that's the way that this DSMB for an anesthesia device worked. We had all of the data collected. The clinicians made whatever response they made, blind to the data, and then we adjunctively went back and crossed the events with the data. Then we also looked at the data to see if there were corresponding events.
So that's not unusual. I think it's a reasonable model.
DR. SESSLER: We often do our studies that way.
DR. CARLSON: I mean, I agree with all this and this is the best way of doing it, we probably should go that way. But this is a huge infrastructure cost of how we get this. So you're talking about isolate OR suites where you're doing

Page 248

this, where the data is available, that infrastructure is there --
DR. SESSLER: No, no --
DR. CARLSON: -- but we're talking about sedation units where maybe I'm naïve in the sense of collecting this data in places that we're doing these sedations.
AUDIENCE MEMBER: Research is research.
DR. SESSLER: I disagree. I'm sorry, I disagree. Collecting continuous data is now relatively easy. It's at the laptop. You hook it up. You can move it from -- you can put the whole thing on a card with a special set of monitoring.
DR. WARD: So can we agree that for a phase 3 trial, this would not be an appropriate tool to recommend for use for a phase 3 trial, unless there is also continuous data associated with it. Is that the consensus that I'm hearing?
DR. RIKER: Make us vote.
DR. WARD: May we vote?
DR. GAN: But we have to also define a little bit what you said, Dan. When you say

Page 249

continuous, what do you actually mean by continuous? It could mean that you set a blood pressure every 2 minutes or 3 minutes, right, in the standard way, and then you collect all these data. You may not [indiscernible] versus, in extreme, putting a line [indiscernible] to do a continuous.

DR. WARD: Let's define data, whether it's continuous. This is a tool that doesn't have any associated data with it. I don't have a vital signs record with it.

DR. GAN: We need to define what continuous means.

DR. WARD: You can have a whole discussion on what continuous means. Right now, I kind of want to focus on -- this is a tool by itself, without any -- with the demographic data. You've clearly got to have the patient level demographic data with it. But you don't have any measurements of the data, continuous or discontinuous, or whatever the sample rate happens to be, which these clinical interventions outcomes are based on.

Page 250

So I'd like to get a vote. Would this be an appropriate sufficient tool in that situation for a phase 3 trial?

DR. RIKER: Without continuous data.

DR. WARD: Without the physiological data -- not talking about the sample rate. Maybe it's continuous, maybe it's once every minute, maybe once every five minutes. But this is a tool that by itself doesn't have any physiological data connected to it. It only has -- it's an intervention outcome collection tool. Great for QI.

DR. SESSLER: This is like measuring the frequency of car crashes with two different cars without recording the speed.

DR. WARD: It is. That's not a bad --

DR. PRATIK: Middle ground, if you cannot get to this discussion about continuous, having thresholds over here without calling them adverse events, you get caught up with thresholds because you start calling them whether they're adverse of not.

Page 251

But if you're just capturing data, even if it is dichotomous, it will be a step up beyond just having the tool, but still not getting to point where we are in the state where we have the continuous thing.

So in large RCTs with many centers, it's probably going to be difficult at investigator-initiated study to have --

DR. WARD: No, I understand what you're saying, but for what we have up there, vital sign abnormalities are defined as changes greater than 25 percent of baseline and 2 standard deviations. I think as a PI in that study, unless I had that data to look at, I wouldn't trust that out in the hundreds of places that there was enough consistency of the investigator to have checked that hypotension, exactly having met that vital signs criteria.

DR. PRATIK: But the thresholds can be made easier when you're not trying to call them adverse events. So if you just say anything less than 90 systolic, I want you to report. Okay, and you're

Page 252

not calling them adverse events. You're just saying two groups --

DR. WARD: I'm using this tool.

DR. PRATIK: Yes.

DR. WARD: I'm not redesigning the tool. I'm using this. If we redesign the tool, then we're back to collecting data for a clinical trial. I don't see anybody saying that this is a sufficient tool --

DR. PRATIK: In its present form.

DR. WARD: -- in its present form for a phase 3 trial. Phase 4, postmarketing survey, now you've got -- the drug is in use, you want to do some postmarketing. You put out there, use this tool and we're going to put it into a central repository, does anybody think this is a sufficient tool for that?

DR. RIKER: Phase 4?

DR. WARD: Phase 4. For phase 4.

(Hands raised.)

DR. WARD: Possibly? So a little bit possibly sufficient.

Page 253

Maala?
DR. DAHAN: Still too static in my opinion.
DR. WARD: So no consensus that this would necessarily be a monitoring phase 4, postmarketing survey.
DR. DEXTER: With hypotension listed as hypotension or relative to an unspecified standard, which you can't know for different patients, then -- because knowing a baseline depends on the frequency of monitoring by definition. Without specifying the frequency of monitoring, then it would not be interpretable in terms of the hypotension.
It might be good for screening purposes, for quality purposes. But as part of a phase 4 trial, it wouldn't be interpretable at the end. If you were to say, for example, a mean blood pressure of 65 or something like that, when monitored every 2 minutes, then I think that that would be quite reasonable in that circumstance.
DR. ROBACK: Maala?
DR. BHATT: So I just had a question for

Page 254

Mark and Steve I guess, given the compensation, people seem uncomfortable with the brevity of this tool, which is I think what makes it powerful for a clinical setting for sure, and perhaps in the phase 4.
But what do you think about going back to the Quebec guidelines if people want more documentation, you know it won't -- we still I think are left with no continuous data, but you are documenting a lot more in terms of hard numbers, like for oxygen saturation and such. I don't know what the room feels -- or what you guys feel like in terms of the evolution or if that's still insufficient.
(No response.)
DR. BHATT: Well, I just wonder like, because it sounds like people are uncomfortable with the brevity of the tool. Correct?
DR. ROBACK: Right. But I think we always come back to the threshold and duration and what does it mean? And no one can agree on what is important.

Page 255

To use the car wreck analogy, well, I don't know the speed, but I know there's a scratch on the bumper and the airbag didn't deploy, so I don't think that's a big deal, versus the one where it's a total problem, and three people died. But to get to that, the physiologic numbers, that's a different level of scrutiny, and is that required for a phase 4 trial?
DR. WARD: Well, that's where the question -- I mean, I think we've already decided it is required for a phase 3 trial. But what I'm sort of hearing that there's not much enthusiasm for this as a tool for a phase 4.
Yes?
DR. ROCA: I just want to make a quick comment. As you consider whether it would be useful for phase 4, keep in mind that sometimes a phase 4, depending on what you're looking at, and depending on how circumscribed the phase 3 trials were that garnered approval, phase 4 may actually may be opening to a whole completely new patient population, where you really don't have information

Page 256

in the phase 3 trials that would apply to those.
So you may actually have to consider a phase 4 trial to be needing as much data as the original phase 3.
DR. WARD: Right. Okay. So I think you've got a great QI tool.
(Laughter.)
DR. WARD: But I don't hear a consensus at all that I'm going to put in the paper that it's a tool that would be favorably considered for a clinical trial. Is that fair enough?
DR. GREEN: And I just challenge everyone to come up with something better.
(Laughter.)
DR. CARLSON: Well, I think there's issues that we do -- this may not be it, but there's nothing else now. Doesn't mean that you go with this, but that challenge is in the paper to say that it is paramount for thoughtful people to put something together that works.
DR. WARD: And this has a big advantage as you collect the data, of having a database out

Page 257

there that sort of says how many corners have had an automobile accident? We don't know much about it, but we know that street corner has an issue for it.

So I'm not saying it's not useful data in a clinical trial, but designing a clinical trial in which this is sufficient as the primary outcome measure, doesn't seem like it's useful for that.

DR. GREEN: Yes. You know, I think in clinical practice and in research, there is an incredible diversity of these outcomes and what I would consider an unacceptable diversity that prevents us from doing any meaningful comparisons. The whole reason why we took on this project in the first place is to try and come up with something that was better than total chaos.

DR. WARD: Oh, absolutely.

DR. GREEN: So admittedly, not everybody is going to agree on something like that. So this is just a humble attempt.

DR. WARD: And this may be a platform, because it is going to be a Web-based platform with

Page 258

a database behind it, it may be a platform that will have TROOP 2 that will be designed as a clinical trial platform, because I think it has a lot of advantages for that.

DR. GAN: Let's say, for example, if you want to recommend this, let's say you recommend for a clinical trial, don't you then have to look at internal consistency, how to people rate it and that sort of thing?

DR. WARD: Yes --

DR. GAN: -- to be acceptable?

DR. WARD: You'd have to go back and do that.

DR. GAN: Yes.

DR. WARD: And for that you need the physiological data to validate for that.

Maybe we could swap places and let me in the last 45 minutes, put you guys on the panel.

Can I have my slides up again? And I'm going to force some votes here, and then some comments from the panel. It was the consensus.

So these are statements, and don't take them

Page 259

too literally. But if the statement in some form appeared in the paper, I want to know whether it would be a yes or a no. You would like that statement to be in the paper, or you would not like that statement to be in the paper. Again, not those exact words.

DR. DEXTER: When it says, "for many outcomes for severity and length of time data should be collected," the thing would be, as Dan has talked about, it's really sort of an integral. In other words, because when you say severity, the problem is that if you use the -- let's take pulse oximetry. If you use the nadir measurement, that's very subject to artifact.

DR. WARD: So you're down on this one here.

DR. DEXTER: Yes.

DR. WARD: We're going to go through them in sequence, but that's okay. We'll take it out of sequence. We'll start with this statement.

DR. DEXTER: I think that it isn't the severity, because inherently, that's the nadir which is very poor statistical properties, and it's

Page 260

highly sensitive to the sampling rate, and that will differ among pulse oximeters, for example, based upon the averaging.

In the same way, the length of time data will also differ markedly among different monitors, even connected to the same patient, whereas, if you use, just like Dan had said in his presentation, a method, which is an integral type method which combines the two, you have the raw data, that takes that into account.

So it just would be a bit more circumspect in the wording there.

DR. WARD: Like I said, don't take the wording too literally. I've just tried to write down -- the concept includes what Dan was talking about, that there are two aspects to the data, is the severity and how long you desaturate. A two-second desaturation to 90 is not the same as a 30-minute desaturation to 60.

Dan?

DR. SESSLER: Nowhere here do you say "continuous monitoring." You imply it, but you

Page 261

don't actually say it.
DR. WARD: I don't actually say that, and that could be an oversight --
AUDIENCE MEMBER: You won't forget.
DR. WARD: Yes, that may be an oversight. I wrote it down or it may be intentional. I'll have to decide which it's going to be. You have to define continuous though, Dan.
DR. GAN: How do you even define it?
DR. WARD: Yes, that's why it isn't there. Defining continuous is a difficult situation.
DR. DEXTER: Defining continuous, I don't perceive it as being particularly challenging because it depends on the time course of physiological change when you're talking about measuring something in nanoseconds or longer periods of time. So you look at the time course of events, and sample at a rate that's comparable to what is going on.
DR. WARD: But it is sampled, following continuous --
DR. DEXTER: Concerning that pulse oximeter,

Page 262

what you're getting is averaged over a certain number of seconds anyway, and so forth. There's never a reason to be sampling at a rate that exceeds the average end period, so to speak, of your device.
DR KARAN: Blood pressure measurement, that's not arterial?
DR. SESSLER: I would measure whatever you've got. I'm not proposing putting an arterial line into everybody or even using a non-invasive continuous monitor. But if you get a blood pressure every 5 minutes, fine. If somebody hits the stat button and gets an extra reading, record it.
DR. GAN: Most people would not be --
DR. WARD: Yeah. You're saying collect all the data available.
DR. SESSLER: And maybe I should use the ASA word, which is "continual."
DR. WARD: Continual, yeah. Continuous gets to be, for people like Frank and I especially, gets to be mathematical kind of terminology.

Page 263

Breath-to-breath data, for example, is by definition discontinuous because it's only measured on a breath-to-breath-to-breath.
DR. DEXTER: From the perspective of a early -- let's say phase 1/3 trials, something like that, 5-minute increments is a very long period of time, I mean, if you're giving the bolus of a drug or something like that. So initially, you're going to do it more frequently once you've learned the nature of responsiveness and the timing.
DR. WARD: So all those caveats would be incorporated as I write the paper. So take the statement as a -- so let's start at the beginning. I want to hear a vote here.
The scope of our recommendations include phase 1 to 4 clinical trials? Yes?
(Show of hands.)
DR. WARD: Anybody say no?
DR. SESSLER: What are the recommendations?
DR. WARD: We haven't gotten there yet.
(Laughter.)
DR. WARD: But we are excluding phase 1

Page 264

trials.
DR. SESSLER: Shouldn't that be the last question?
DR. WARD: Yeah, it might be.
The physiologic profile of a new compound -- some of these are motherhood and apple pie, I understand -- a new compound, well characterized in the phase 1 clinical trials. And that gets to some of the things that Albert and I talked about, 2 response curves, those kinds of things.
DR. SESSLER: Do you want to be specific about dose response?
DR. WARD: Yes.
DR. SESSLER: Because it's not in there.
DR. WARD: Yes, would include a dose response.
DR. DWORKIN: Maybe that should be phase 1 and 2.
DR. WARD: Phase 1? Bob, would you mind taking some notes on this for me as to which ones kind of say yes, and which ones have more

Page 265

modification? So phase 1 and phase 2.
This one I think we've already discussed. MedDRA is probably useful to classify, but TROOPS is not a sufficient classification system.
Everybody say yes? Hands up?
(Show of hands.)
DR. WARD: Okay. Data should be collected in ways that facilitate pooling from meta-analysis? Yes?
DR. SESSLER: What does that mean?
DR. WARD: I think we heard some of that from our statistical discussion yesterday.
DR. DEXTER: I think that I kind of agree with Dan's question. I think if you collect data as part of phase 1 through phase 4 studies, you're collecting all the data. So I'm not quite following what, as part of clinical studies, that means. If you're collecting all the data, how is that different than facilitating pooling for meta-analysis?
DR. LIGHTDALE: Sedation studies are routinely confused -- I mean, outcomes are

Page 266

different across sedation studies, and makes it very difficult [inaudible -- off mic].
DR. SESSLER: Right. Okay. So what you do when you do a meta-analysis is you try to find a bunch of studies that are generally similar in terms of interventions and in terms of outcomes. And the problem is the outcomes here more than anything.
DR. GREEN: So in order to do that, you have to have some form of tool to standardize it. So this is just one. Come up with another one.
(Laughter.)
DR. WARD: Making patient-level data available would be one way to facilitate meta-analysis, like the FDA has.
DR. SESSLER: Well, there are patient-level meta-analyses, but they're not usually going back to continuous data. That's pretty difficult.
DR. WARD: Okay. So we should ignore --
DR. SESSLER: I think I would not put that in there because it's not the least bit clear, at least to me, what that means.

Page 267

DR. WARD: Okay.
DR. GAN: But your intention is to collect similar outcomes so that you can make comparisons, right? I think that is what that meant.
DR. WARD: That's what I meant by that, right.
DR. GAN: Motherhood and apple pie, sure. Who wouldn't agree to that?
DR. SESSLER: I think it's meaningless.
(Laughter.)
DR. GAN: The second one is meaningless --
DR. DEXTER: If the point is that you should use standardized measures of efficacy and standardized measures of adverse events.
DR. GAN: That's exactly --
DR. DEXTER: My recommendation would be to say that.
DR. GAN: Exactly. That's exactly what it says.
(Crosstalk.)
DR. DEXTER: Just say that.
DR. DWORKIN: [Inaudible - off mic].

Page 268

DR. SESSLER: A general rule of practice is say what you mean.
DR. WARD: I sat down, and in 15 minutes with my handwritten notes over the meeting, so don't pick on my ruining too severely. I'm lucky that there's no misspellings there.
DR. GAN: Frank articulated what you just said. It's the same thing, what we are talking about.
DR. DEXTER: But I think if you say what I said, then you don't have to mention the pooling or the meta-analysis. We should use endpoints of efficacy that are standardized.
DR. WARD: Right. That's what I'm saying.
DR. DEXTER: Even if you're not going to do meta-analysis, then you should use measures of adverse events that are standardized, even if we don't do meta-analysis.
DR. WARD: Okay.
DR. SESSLER: Say what you mean.
DR. WARD: Well, I will in the paper, Dan. Give me a break. In the 15 minutes, I scribbled my

Page 269

notes to put these down. That's why I'm putting these up here, for you guys to help clarify for me to say what I mean. And your names are going to be on the paper, so you guys get to say what you mean, too.

All right. So we talked about this one already, with the caveats that we talked about. I think we've heard that a lot, that you have to classify both severity and length.

DR. GOZAL: Maybe we should "precise" after "intervention."

DR. WARD: Yeah, intervention to correct it?

DR. GOZAL: Yeah.

DR. WARD: Those are both things that would be collected in order to decide when an intervention was made, would have to be based on both. And we've been talking about that quite a bit. The relationship or lack thereof between adverse events and adverse clinical outcomes should be defined as much as possible.

These kind of papers have a lot of motherhood and apple pie kind of statements in it.

Page 270

And that's not bad because these statements kind of are made to each other, but they don't end up in a paper like we have put together, or ACTTION has put together in several other areas.

That's one thing we've struggled with, what is the relationship between events and adverse/severe. Will they go on to be a severe adverse event that we rescued from; will they go on to be a severe adverse outcome? I think clinical trials, we've talked about that. If they have some sort of statement in there about that possible relationship, depending on what they're measuring, they could be useful.

Everybody okay with that? Jenifer?

DR. LIGHTDALE: Are you saying define a priori as much as possible?

DR. WARD: Yeah, as much as possible. Yeah, a priori.

DR. LIGHTDALE: Could you as a secondary comment, that also studies should be prepared to use data to further define, re-define?

DR. WARD: Right.

Page 271

If clinical intervention is used as an outcome, it should be procedure specific in a well-defined criteria for intervention initiation. That's Dan's example of the nasal prongs, endoscopy versus nasal prongs in an MRI, that if you're going to have an intervention that you're collecting, you can't just say we collected data when nasal prongs were put on to the patient if your study is heterogeneous in the kinds of situations that it was in.

We're okay with that? Again, these are concept statements. These won't be sentences in the paper; might be.

A clinical intervention scale would be useful, but we need to be shown they're reliable and valid.

DR. DEXTER: I don't think you need to have "valid" because it's very unlikely that people -- I mean, I know it's like apple pie. It sounds great, except it's not like people are going to be doing clinical interventions likely that are invalid. I think validity is pretty clear, but reliability

Page 272

will be the challenge.

DR. WARD: It's easy to show face validity on any kind of clinical intervention scale, but --

DR. DEXTER: Most likely.

DR. WARD: -- you probably can't go -- I agree -- much beyond that. But reliability is very important. And as I know, there really is no clinical intervention scale that has been looked at for its reliability. We have to go back to the physiological data and see if the intervention scale matched -- what you've already talked about -- both directions for that.

DR. BERKENBOSCH: I think one of the problems that you're going to have partly goes back to the preceding statement as well, about having well-defined criteria through intervention initiation. I think, as Mark and Steve have stated, and certain is true in our practice as well, clinicians aren't necessarily going to stick to rigid criteria for intervention.

So it might be worth adding as an addition to that statement above, have well-defined criteria

Page 273

for intervention, or describe the events relating to the intervention that was performed, because people aren't going to stick to criteria all the time. But there has to be at least some further characterization of the events surrounding the data or surrounding the intervention.

DR. WARD: Right.

DR. BERKENBOSCH: To give a little bit more of an idea of why the intervention occurred.

DR. WARD: The next one is not too much -- there's a lot of overlap between these. This was a concept that I liked when it came up. It's useful to consider the measuring instrument and the relevance instrument separately.

DR. DEXTER: What is relevance instrument?

DR. WARD: Well, it sort of goes back -- again, not to overuse Dan's example, putting nasal prongs on a patient could be something that you're measuring, or additional supplemental oxygen could be something that you're measuring. It's not particularly relevant in some procedural sedations, and very relevant in other

Page 274

procedural sedations.

So understanding that you might be maybe measuring a bunch of things, but then you've got to make a decision about what you measured is a -- the desaturation for 5 seconds to 90 percent, you measured it because you've got a lot of data coming in. Then you can make a decision that your screening tool, it then decides whether that measurement becomes a relevant outcome. So separating the concept of your measurement tool and your relevance tool.

DR. GAN: So more like context sensitive.

DR. WARD: Context sensitive, procedure sensitive, patient sensitive.

DR. DWORKIN: Don't you want a stronger word than "useful"? Isn't it important or imperative? Useful is a pretty weak word, given what you just said.

DR. DEXTER: Since you have the previous statement that a clinical intervention scale would need to be reliable, I'm wondering, because of exactly the point of useful, whether or not the

Page 275

second point is taking in account the third.

DR. WARD: Or the third taking in account the second.

DR. DEXTER: Yeah. The other issue is when you say "measurement instrument," it's not clear whether in instrument you mean like a instrument like a scale, or you mean an instrument like a physical device, like the pulse oximeter.

DR. WARD: Could be other one, or both. It includes both was the concept I think, as a general idea.

DR. DEXTER: Doesn't the second one incorporate the third, and you can just kind of not worry about the third?

DR. WARD: The third's a concept, and the second is a particular device, a scale, 1 to 5. No. I think the second is incorporated by the third as a general concept.

DR. DEXTER: So for example, how about instead of getting a word, a "clinical intervention scale," just say a "clinical intervention," or "measurement of the clinical intervention."

Page 276

DR. WARD: Okay. I'm not wordsmithing these. These are just concepts. We'll argue about the wordsmithing once I send the paper around. But I'm trying to find out what concepts; am I going to set somebody off if I include the concept in it or if I don't include the concept in it.

The cause of an AE should be separated by a drug, disease, procedure, provider. That's the context, and that's come up.

Mark?

DR. WEISS: I would consider also putting the word "site" as well, because I do believe -- it may not be politically correct, but I do believe that the site where we do these procedures directly influences, sometimes, either the rate of which we have an adverse effect, or our ability to respond to the adverse effect. So I would include the word "site" in that as well.

DR. WARD: Yeah. No, I agree. I think doing something out in the boonies, you have a lower instance of intervention than doing it down in the operating room when you've got lots of help

Page 277

around. So, yes, site is good. Thank you.
Other? Rather than call for votes in every one of them, I'm just going to kind of go through these, and let's hear objections or agreement.
DR. WEISS: I'm surprised that this went through without someone picking at it. It got through.
DR. WARD: I've been known to wear people out, too.
Where are we? Decreasing major sophistication can be used as the compound progresses from phase 1 to 3. We've talked about that, although we're sort of coming into the consensus that even phase 4 trials probably require a higher level of measurement than would normally be done clinically. There's no reason to say we can just collect the usual clinical data in certainly in a phase 3 and a phase 4, but I probably don't need to put all the RespiTrace and everything else as we get more and more patients as the study's progressing.
DR. KARAN: If your phase 4 trial involves

Page 278

OSA patients, or your phase 3 trial didn't, then you might want to use something else --
DR. WARD: That was a point you made, that if we expand the phase 4 to patient population, then you may have to back up. But as a general statement -- good point.
High-risk populations need to be included in phase 2/3 clinical trials. Is there general agreement on that?
DR. WUNSCH: I'm not sure I completely agree with that. I just worry that for certainly a phase 2 trial to be enrolling high-risk patients still feels to me quite odd. And even a first phase 3 trial I feel like, generally -- and there might be many exceptions to this, but that generally that wouldn't be one I would recommend.
DR. SESSLER: I agree. I wouldn't mandate high-risk patients in a phase 2 trial because a phase 2 trial may be based on 10 phase 1 patients, and you know almost nothing about the drug. It's prudent to start with a low-risk population.
DR. WARD: But a phase 3 trial?

Page 279

DR. SESSLER: By the time you get to phase 3, and you're trying to make a broad statement, and if it's successful, the drug will get approved, and it will go out and be used in many contexts, not just the one you studied, then it's probably appropriate to include a broad range.
DR. WARD: Because I think, as we all, clinical trials in sedation, efficacy gets shown pretty quickly. If something's not going to progress beyond a phase 1 trial, it really doesn't work. So we're really talking a phase 3 now as opposed to other compounds, where we're still not sure what the efficacy is in a larger population. We're only more worried about are there going to be too many adverse events that show up.
This is a different kind of climate. We're all aware of this clinical trial paradigm in a diabetic or in a hypertensive drug in which you don't expect a lot of adverse events. And for diabetic drugs, we've worked out ways to look at them, but we're still not sure how effective it's going to be.

Page 280

When we moved dexmedetomidine from a phase 1 to a phase 2, we were pretty sure that it worked as a sedative. We weren't sure whether its cardiovascular effects were going to prevent it from reaching the market. So that's why I meant perhaps including the at-risk population a little earlier than I would if I was still worried about efficacy. But I take your point. We'll just say phase 3.
DR. SESSLER: Right. I'm just saying don't mandate it.
DR. WARD: Yeah. Well, in this paper, we're not mandating anything. This group thinks that, is what the sentence kind of starts with for that.
Non-standard monitors, something like cardiac output, may be useful in clinical trials, but their accuracy should be defined if we're going to use them in a clinical trial. The newer infrared spectroscopy is another. It's easier to throw in a non-standard monitor; you know, let's throw this in and measure this, too. But unless you have defined its accuracy, it probably won't be

Page 281

useful.
DR. WUNSCH: Should there be something to differentiate between non-standard monitors that are non-invasive versus invasive in terms of how much we would suggest people consider using them? I think there's a huge risk profile that goes as soon as there's any invasive monitoring that gets added.
DR. WARD: What does everybody think? I think I'm more worried about accuracy rather than invasiveness. I think getting inaccurate data out there in a clinical trial has a negative purpose. I think invasiveness adds another ethical issue on top of it, but whatever the group --
DR. SESSLER: I agree. The issue doesn't seem to be invasive or not; it seems to be accuracy. And the burden is on the investigator to justify any methods used, and the reviewers will look closely at anything that's not standard.
DR. BERKENBOSCH: And you would appropriately describe the limitations of whatever non-standard monitoring you choose to incorporate

Page 282

into your trial.
DR. SESSLER: Exactly. This is just generic methods. We used some methods, and it's validated as follows, and here are the limitations. And then if they're real limitations, you put a paragraph in the discussion describing them.
DR. WARD: Future work is needed to define major post-procedural neurological deficits. We didn't get very far, how long afterwards we need to measure them, what we're looking at. It's pretty hard to make any kind of statement related to that. The incidence of post-procedural delirium should be recorded. If you're doing a clinical trial, that's one of the things that you should have in it.
AUDIENCE MEMBER: Do we need to say about instrument or --
DR. WARD: Well, we've got some instruments available. I think in the first meeting, we spent a lot of time discussing instruments because we had a lot of instruments to go through, and we came down to a few. I don't think we have as many instruments, and probably the data to validate and

Page 283

be able to select one versus another. That's my impression, but I'll take other views. Ron's shaking his head.
DR. LITMAN: No, we don't. Jerry has, as he talked about, the only published delirium score, but very few people use it.
DR. WARD: But it could be used in a clinical trial.
DR. LITMAN: When it's easy [indiscernible], he could use.
DR. WARD: Yeah.
DR. CONWAY: Is it really a safety outcome, then? And it's in the rate/goal [ph]. Like rate/goal is specifically in an efficacy paper. I'm not sure that delirium is a safety outcome.
DR. WARD: I'm sorry. What?
AUDIENCE MEMBER: It could be insofar as it might delay the discharge home.
DR. WARD: Yeah.
DR. IRWIN: I think emergence delirium is definitely a safety issue. I actually saw a case where a woman was placed on delirium, got on to the

Page 284

roof of the hospital, and jumped off the roof of the hospital. So that's pretty -- pretty serious
(Laughter.)
DR. IRWIN: This was in Los Angeles, actually.
DR. PANDHARIPANDE: I don't think we are restricted to just emergence delirium. There's delirium even in the post-procedure, where you emerge, where you may have high [indiscernible] delirium after that. And there are instruments. Apart from the emergence delirium scale, there's the Cornell scale. I mean, there are now instruments that you asses, not in the first few seconds after emergence, but in the period after that, in the recovery room.
DR. WARD: And I'm not saying every single phase 3 would have to have a major post-op delirium, but in the process, post-op delirium is one of the things that you'd want to look at for sedation clinical trials.
DR. SEXTON: And with respect to the FDA, the FDA would expect that to be reported in an

Page 285

adverse event.
DR. WARD: Right.
Patient recall may be useful but will need to be defined in relation to expectations. Unexpected?
DR. RIKER: This is a morass. In the ICU, unpleasant recall occurs more frequently in the deeply sedated patient than it does in the lightly sedated patient. And the issue is they're so deeply sedated, they establish no real memories, and all they have are delusions and hallucinations.
So that's in the ICU, and I don't know how long that has to be present to be even applicable to a shorter-term procedure. But we saw yesterday that with ketamine, emergence delirium, or hallucinations is 10 percent. So those patients might recall having been assaulted when that never happened. So this gets really tough.
DR. WARD: Yeah. And maybe it's a morass that we don't want have into.
DR. RIKER: Or it needs to be a more sophisticated assessment.

Page 286

DR. CARLSON: I agree with that, but I think in the sense of a lot of things, we need to measure it, and then decide later how to apply it, because not measuring -- I think recall and sedation is an important part of all the things we're talking about.
DR. SESSLER: It seems like an important outcome. If a patient remembers getting assaulted, it doesn't matter whether it happened or not. That's an adverse event. It's something we definitely need to know about because it is more likely with some drugs than others.
DR. WARD: But I think what you're also saying is that probably simply asking the patient what do you remember is not sophisticated enough to be included in this kind of clinical trial.
DR. RIKER: And as we talked about more in the first meeting, there are many procedures that recall is okay, and then others where it's completely unacceptable.
DR. WARD: And there are several papers that look at patients' expectations of recall. In fact,

Page 287

your paper, Frank, looked whether the patient has been told they're going to have recall versus not told whether they're going to have recall makes a difference, whether it's unpleasant or unexpected or not.
So it's a morass, but I think if you're doing a clinical trial, you can define some of the parameters that you want to look at. I guess what I'm saying there is that it needs to be defined in relation to the expectations. You just can't ask did you remember anything, and if they say yes, that's an adverse outcome. It's not that simple.
Finally, the difficulty in relaying the occurrence of biomarkers in phase 3 clinical trials to rare SAE mandates postmarketing studies and data collection. I mean, that gets back to the midazolam story originally. We don't really know what a CO2 response curve for example, or even a saturation during a sedation, means as far as how often you're going to get a respiratory arrest when you're using it in thousands of patients.
Any other ones people want to add?

Page 288

DR. GAN: Is it important to assess when the patient gets back to baseline? As we are now more patient centered -- again, this is not something we have in the past collected, but I just wonder whether it's something that we want to find out.
DR. DEXTER: I was kind of thinking of that as efficacy of the sedative agent. I'm not sure that it's safe -- I mean, I'm not sure. Is that safety or is that efficacy, in terms of the time course of the response?
MR. WILLIAMS: I think from the first meeting, we thought about that as well. It came under "efficiency" as well. We were thinking about as from the aspects of recovery, recovery scales, sedation.
DR. DEXTER: Definitely, if it were to be -- I know we haven't covered efficiency. If it ever were to be covered, then definitely it would be part of that.
DR. GAN: Well, it's part of both. I mean, say for example, if somebody is still sick the following day, for example, or still drowsy

Page 289

following after --
DR. WARD: That would be I think a neurological outcome --
DR. GAN: Right.
DR. WARD: -- but that's not kind of returning to -- I mean, that's different than kind of when are you ready to be discharged.
DR. GAN: Not so much discharged; when you get back to what your senses -- because sometimes people, they say I had this procedure, but it takes me two or three days to get back to normal.
DR. WARD: If you ever used Innovar -- I'm old enough to --
DR. GAN: Oh, yeah.
DR. WARD: -- have used droperidol and fentanyl when I was a resident, and you will have several days before patients say that they're back to normal.
So yeah, point taken. I thought you meant more you're ready for discharge.
DR. GAN: No. This one is more longer term.
DR. WARD: And I think that came up when we

Page 290

talked about longer term follow-up, the 1 week to 1 month kind of follow-up for neurological changes.
Jenifer, then Hannah.
DR. LIGHTDALE: Just very quickly, I think certainly safety, if you're not ready to drive, we really need to know when you can drive a car. Yes, I think there's a safety issue.
DR. WARD: Now I understand. Separating it from when you can discharge them from the recovery room issue. Hannah?
DR. WUNSCH: I'm just wondering whether we consider expanding the last one to not cover only biomarkers but interactions with other medications, because certainly the dental data was pretty compelling in terms of the addition of multiple medications together. I think the midazolam story is somewhat the same, that it might be that interaction that is not going to be picked up.
Group Discussion: SCEPTER Next Steps
DR. WARD: Drug interactions may not get picked up. But I think potential drug interactions should be investigated in phase 3 because you don't

Page 291

use -- if you have a sedation drug, it's probably not going to be used in isolation. So if you didn't do a phase 3 trial at the beginning of the midazolam, making sure you combine it with an opioid -- that's why I think it's going to be used, so good point.
Other final wrap-up comments from the panel?
DR. ROBACK: As I'm thinking more about continuous physiologic data and our recommendation that this is really required, I think we need to add that further work is really needed to really understand or to define what these numbers mean, because right now, there is really no consensus, and there's no -- that this number means one thing, and that number means another thing. And we're going to have to develop those as we go forward to make this data really be meaningful.
DR. WARD: Well, I think what Dan's point has been is that you can then connect it to the clinical intervention. The physiological data by itself probably is not going to be meaningful unless you connect it to the clinical intervention.

Page 292

But vice versa, the clinical intervention may not be meaningful if you can't connect it to physiological data. So that's a good point to be able to make there.
Other comments from the panel or the audience?
(No response.)
Adjournment
DR. WARD: Well, I think that wraps up the meeting. I hope you've enjoyed it as much as I have. I want to thank Bob for sponsoring the meeting through ACTTION, and the ACTTION FDA public-private partnership.
Please send me your comments. We have copies of your slides, and we'll start work on the paper. If you want to meet, you all have opportunities to be co-authors on the paper, so the first draft will be that. It will be a very early draft that I will send around to everyone. Thank you.
(Whereupon, at 2:47 p.m., the meeting was concluded.)

B

C

D

E

F

G

H

M

N

O

P

S

Z