ACTTION SCEPTER-III - Clinical Trials to Evaluate Patient-Centered Outcomes in MVPs in the Adult ICU

March 28, 2019

A Matter of Record
(301) 890-4188

ACTTION

SEDATION CONSORTIUM ON ENDPOINTS AND PROCEDURES FOR TREATMENT, EDUCATION, AND RESEARCH (SCEPTER-III)

Clinical Trials to Evaluate Patient-Centered Outcomes of Sedation in Mechanically Ventilated Patients in the Adult ICU

Thursday, March 28, 2019
8:03 a.m. to 5:04 p.m.

Westin Georgetown Hotel
Washington, DC

C O N T E N T S

C O N T E N T S (continued)

C O N T E N T S (continued)

P R O C E E D I N G S
(8:03 a.m.)
Welcome and Introductions

DR. DWORKIN: Good morning. I'm Bob Dworkin, and I'll give you a very, very few minutes introduction to what ACTTION is. ACTTION is a public-private partnership that was established by the FDA in 2010. They're not here. I don't see them.

The people who were incredibly instrumental in getting this going and continuing it were Bob Rappaport, who's now retired from the FDA, and currently Sharon Hertz, the director of the Division of Anesthesia, Analgesia, and Addiction Products, and Allison Lin, who I think will be here later today. And lots of other people from the FDA have been involved in supporting and helping out with ACTTION since 2010.

So what is ACTTION? It's a public-private partnership. The FDA, its notion of public-private partnerships is to get everybody working together to accomplish something. So the initial mission of

ACTTION was to -- and I'm not going to get this right. But the essence of the initial mission of ACTTION was to figure out how to accelerate the development of improved treatments for acute and chronic pain, improved being either better efficacy, or better safety, or both within a couple of years.

Within a couple of years, I think around the 2012-2013 time frame options, ACTTION's scope has expanded to include three additional therapeutic areas: sedation, addiction medicine, and peripheral neuropathy. So since about 2012, ACTTION has tried to figure out, as a public-private partnership, how do we accelerate the development of improved treatments and interventions across those four different therapeutic areas.

I haven't figured out how to kind of analyze this in a quantitative way, but my gut feeling is about 40 to 50 percent of ACTTION's activities now are pain, and the other three areas of addiction, medicine, sedation and peripheral neuropathy are

split equally with the remaining 50 to 60 percent.

In the FDA's view, a public-private partnership brings together all the relevant stakeholders, and ACTTION I think has been quite successful in involving individuals from professional societies. For example, Denham was the ASA's first representative to the ACTTION executive committee. Jim Eisenach, who many of you know, is the current ASA representative to the ACTTION executive committee.

We have participation of multiple professional societies, academic investigators from around the world, and patient advocacy groups; pharmaceutical and device companies provide support and of course government agencies, not only the FDA, but NIH, CDC, and occasionally DEA. And we do our very best to get international participation, specifically from the EMA, but also from other European initiatives.

The mission has remained the same, so with respect to the mission, I think our focus across the four different therapeutic areas of pain,

sedation, addiction, and peripheral neuropathy has really been clinical trials. How can we optimize and improve the design of randomized clinical trials across those different areas of medicine?

How can we optimize the design, the outcome measures used, the statistical approaches to analysis, and make sure that the data are interpreted correctly? That's not the entirety of what ACTTION has done, but the bulk has really focused on clinical trials and improving their design and execution and analysis.

The other thing I should say before I end is we've also done our very best -- and we think this is incredibly important -- to encourage the participation of junior investigators whenever possible and in any way possible. We provide support every year for a 4-day pain school that's held outside of Montreal, Canada, where 30 trainees, both basic and clinical, spend 4 days learning about how to do pain research.

We're doing the same thing this summer with a preclinical boot camp in Dallas, Texas. With

Page 9

respect to initiatives like this meeting, and publications, and systematic reviews, we do our best to get junior investigators involved. So if you have a junior colleague who would like to get involved in any of ACTTION's activities, please, please just have them shoot me an email, and we'll figure out a way to plug them into something that they would be interested in.

I think I've said everything. I'm looking at my notes, two other things. Just in terms of funding, ACTTION has been financially supported by a series of grants and contracts from FDA. We've had two contracts and two 5-year cooperative agreement research grants.

As I said, we also get support from pharmaceutical and device companies. We've had a little bit of philanthropy, not much philanthropy but some, and even less royalties. But the bulk of the funding is really industry support and FDA support, and we've just actually submitted another contract application to FDA.

Finally, this is the first time I'm saying

Page 10

this publicly, so it pleases me to be able to say that last week, ACTTION got its 100th publication accepted for publication, so we're really proud of the milestone.

(Applause.)

DR. DWORKIN: Thanks very much. We're really proud of the milestone of having published a hundred articles since ACTTION was launched by the FDA in 2010.

Before I sit down and shut up, any questions about ACTTION?

(No response.)

DR. DWORKIN: Okay. The only other thing to say is it's ACTTION with 2 T's, and our website is acttion.org, and there's a whole lot of information on the website. Thanks very much.

DR. WARD: Thanks, Bob.

A nice introduction to what ACTTION is and what we're trying to do. I got involved with it when I was his department chair when he came to me and said I want to put in this thing to the FDA to get some money. And I said, "That's a great idea

Page 11

if you want to get money." But then I got more involved with it when it expanded to sedation.

I think maybe we'll start out -- after the cocktails at dinner last night, it sounds like everybody knows everybody, but I don't think that's quite true. So maybe let's start with Rick Riker and go around and introduce yourselves. There is a list of all our participants and I guess any other comments that you want to make.

DR. RIKER: Rich Riker, clinical and neurocritical care at Maine Medical Center in Portland. What a tremendous group we have here. So thanks. I'm glad to be here for sure, from the same Maine Medical Center and our clinical pharmacists, and honored to be here.

DR. FRASER: Gilles Fraser from the same, Maine Medical Center. I'm a clinical pharmacist and honored to be here.

DR. WARD: David?

DR. GOZAL: David Gozal. I'm from Jerusalem, Israel. I run the sedation service at Hadassah University Hospital.

Page 12

DR. SESSLER: Dan Sessler, Cleveland Clinic. I'm a trialist.

DR. FLOOD: Pamela Flood. I'm from Stanford. I do anesthesia and pain medicine, and I'm also a grateful former ICU patient.

DR. SHAFER: Steve Shafer from Stanford University.

DR. VAN CLIEF: I'm Martha Van Clief. I'm at the Food and Drug Administration.

DR. BAZINI: Alla Bazini, also FDA.

DR. EGAN: Talmage Egan from Salt Lake City, University of Utah.

DR. BALAS: Michele Balas from The Ohio State University, College of Nursing.

DR. DEVLIN: John Devlin. I'm a critical care pharmacist from Northeastern and Tufts Medical Center.

DR. ABSALOM: Good morning. I'm Tony Absalom. I'm an anesthesiologist from Groningen in the Netherlands.

DR. MAZE: Mervyn Maze, UCSF, anesthesiologist.

Page 13

DR. SUN: Lena Sun, pediatric anesthesiologist and SmartTots. I'm at Columbia University.

DR. EGEROD: Good morning. Ingrid Egerod, I'm a professor of nursing at the University of Copenhagen.

DR. BROWN: David Brown, and I'm here representing ICU patients, and I'm a recovering academic.

(Laughter.)

DR. AITKEN: Leanne Aitken. I'm a professor of critical care at City University in London, and I do also have an appointment still in Australia at Griffith University.

DR. NEEDHAM: I'm Dale Needham. I'm a professor of pulmonary critical care at Johns Hopkins and then outcomes research and work in the medical intensive care unit.

DR. COLANTUONI: Elizabeth Colantuoni, biostatistician at Johns Hopkins.

DR. DEXTER: Frank Dexter, University of Iowa. I do economic studies, managerial

Page 14

epidemiology studies.

DR. COURSIN: I'm Doug Coursin. I'm an internist/ anesthesiologist/intensivist at the University of Wisconsin. I'm looking forward to the polar vortex leaving town so I can get my kayak and water. Thank you.

(Laughter.)

DR. TUNG: Avery Tung, anesthesiologist/intensivist from University of Chicago.

DR. SPIES: Claudia Spies, anesthesiologist and intensivist from Berlin.

DR. BURRY: Lisa Burry, ICU pharmacists at University of Toronto and Mount Sinai.

DR. SKROBIK: My name is Yoanna Skrobik. I'm from Montreal. I'm an intensivist and recently a pharmacology degree.

DR. SHEHABI: Good morning. I'm Yahya Shehabi from Monash University. I'm a critical care physician and an intensivist, and I'm sorry I missed the dinner last night.

DR. DWORKIN: Bob Dworkin.

Page 15

DR. TANG: Wing Yu Tang, Pfizer. I'm the health economics and outcomes research lead for our targeted hospital grants.

DR. PANDHARIPANDE: Pratik Pandharipande, anesthesia and critical care from Vanderbilt University Medical Center.

DR. HOPKINS: Mona Hopkins, professor of psychology and neuroscience at Brigham Young University and an outcomes researcher at Intermountain Medical Center.

DR. GIRARD: Tim Girard. I'm an intensivist at the University of Pittsburgh.

DR. KRESS: JP Kress. I'm pulmonary and critical care at the University of Chicago.

DR. URMAN: Rich Urman, anesthesiologist, Brigham and Women's Hospital in Boston.

Presentation - Denham Ward

DR. WARD: Great. Thank you.

Just as a little introduction, what we're going to try to do in the next couple of days, we've got a great group of people with a variety of

Page 16

interests, outcomes, statistics, critical care, pharmacology, and at least three continents that I heard. So I think we've got a group that should give us an interesting discussion.

We all know about how a new compound makes it to be used in our intensive care units, from discovery of the compound, through FDA approval, and post-clinical trials. What we're interested in this meeting is the phase 1 to 3 clinical trials. The past meetings have discussed that aspect of it.

As mostly phase 3, but as JP actually wrote it in a prospective, there seems to be a little lack of high-quality phase 1 and 2 trials occasionally before we end up with a phase 3 clinical trial. It's not just a new compound. I think we're also discussing possible devices with possible protocols, anything that would change our practice in the ICU; what's the evidence that we need to generate in order to change that practice so we all believe it?

There are a lot of perspectives to this. What we want to try to do at this meeting is take

Page 17

as many of the perspectives as we can. Obviously, a clinical trial design is just at the early end of this, and you still need good clinical practices to collect the data, and you have to have the right outcome measures. But it's different whether you're sitting at the FDA, you're a practicing physician, you're in pharma, or even more importantly, you're a patient in the public and what's your interest in the right kind of treatments when you're unfortunately a patient in the ICU.

So SCEPTER, as Bob alluded to, has been a sub-consortium in ACTTION. If you ever need an acronym developed, I know who you need to go to. In these days, it's very upon important -- as we'll see, most of the ICU clinical trials have acronyms, and if you get stumped, please email Bob. He will definitely easily come up with an acronym for you. Bob came up with this acronym for us, Sedation Consortium on Endpoints and Procedures for Treatment, Education, and Research.

We've done a little bit already. This is

Page 18

the third SCEPTER meeting. We've published two systematic reviews on procedural sedation. For procedural sedation, we decided to divide it up into a pediatric and adult for the systematic reviews. We did hold two consensus conferences on procedural sedation.

For those conferences, we lumped them together, pediatric and adult, but divided up efficacy and safety in two separate conferences, a little different than what we decided to do for this meeting, which is divide up this meeting for adult in one meeting and in pediatric for a second meeting that we're planning on.

That's probably why Lena is here, because she's going to help me put together the pediatric sedation meeting, and John Berkenbosch, who was with us for the SCEPTER I and II meetings, is very interested in doing that, and he was unfortunately not able to make this meeting.

We published our first paper in Anesthesia and Analgesia on Patient-Centered Outcomes in Clinical Trials of Procedural Sedation.

Page 19

Part 1 was efficacy. That paper immediately got an editorial by John Butterworth, and not an altogether complimentary editorial. He ended up by saying what is this group that has the presumption to actually make a recommendation for something. But he nailed it because at the end of his paragraph saying what does it actually speak for, he said, "Alternatively, should we regard these recommendations as well-intended advice from a group of interested investigators and consultants?" And my letter back was, "Yep, that's exactly what this group is."

I think it's an advantage. We're not representing any particular organization. We're not representing any particular agenda other than a group of, as he put it, well-intended interested investigators and consultants. There's a lot of expertise in this room in a variety of areas, but there shouldn't be any particular political agendas, and I think less so for this meeting than there was for the procedural sedation meeting.

The next paper, part 2, was safety -- the

Page 20

first one was efficacy. We didn't get an editorial for part 2, but it was selected as the article of the month. So we moved from a critical editorial to a complimentary article of the month selection for our two papers.

In the first two meetings -- and I want to suggest we think of something similar for this meeting -- we took the IOM reports that talked about the healthcare quality domains: safe, timely, patient-centered, effective, efficient, and equitable.

We decided for procedural sedation that equitable and timely weren't necessarily important areas, so it shouldn't be any issues about either of those for procedural sedation, but the other four were important, to be safe, patient-centered, and effective. And efficient was perhaps a little less so, and we didn't address efficiency quite as much. That may be more important in ICU sedation.

Patient centered is both patient and clinician centered, and there is overlap. This was a slide we used for the SCEPTER I and II meetings.

Page 21

There are things that the patient is very interested in. These pretty much apply to the ICU sedation also, things that the clinicians are interested in, and a lot of overlap. So when we say patient centered, it has to be patient centered, and it also needs to be centered about what the clinician needs, but the clinician side is the efficacy and efficiency side.

ICU sedation is complex. I'm not an intensivist. I'm an anesthesiologist, respiratory, physiologist, clinical trialist mainly in phase 1 type clinical trials. But I've learned a lot in last I guess almost 9 months in organizing this meeting, and I've done a tremendous amount of reading and a few emails from new and old friends to help me figure out what's going on.

This review paper by Reade in the New England Journal back in 2014 had a diagram that I couldn't resist putting up on how complex ICU sedation is. One point I want to make is pain and agitation, unpleasant awareness, is the important pieces that analgesia and sedation is trying to

Page 22

accomplish in the ICU.

We've had a lot of discussions before the meeting about delirium. For the purpose of this meeting, delirium is truly an important outcome, but it's not something that we're really going to have to be able to discuss about treatments for delirium, per se, either preventive or treating once it's right. But clearly it's a piece of the important employment outcome of ICU sedation and analgesia.

We didn't do a systematic review before this meeting like we did on SCEPTER I and II, and that was because my friends said, well, we've really already done that, and this was last fall, saying the paper's going to come out; it's going to come out soon. And in fact it did. It came out last last fall.

So the PADIS guidelines published in 2018 really provides a lot of the details and systematic review that we perhaps would have done prior to this meeting if we hadn't been so lucky for the PADIS guidelines to come out, and we're fortunate

Page 23

that many of the authors of these guidelines are here with us today.

That's what we'll start out with, with our first panel being, really, a discussion of what PADIS found that perhaps could be improved methodologically and why there are things that PADIS recommendations couldn't have been made because there wasn't methodologically adequate studies to provide the evidence.

Housekeeping, Valorie, who you all met at the front desk, is standing and can wave back there. If you need anything, she'll fix it for you. By the way, the places at your desk, a red light goes on if you start talking. This meeting is being recorded and transcribed, so when you make a comment, please talk into the microphone. Make sure the light comes on so we can get the recording. Speak clearly and, please, every time give your name.

By the end of the meeting if you don't know who you are already, we will know who you are by the end of meeting, but the transcriptionist

Page 24

doesn't. And when I go back and read the transcript, it's nice to know who it was that made that comment, so please say your name before you make your comment.

I guess it goes without saying that this is being recorded and transcribed, so you may want to be careful what you say. In fact, Bob, do we put it up on the Web?

DR. DWORKIN: Yes.

DR. WARD: Yes. So it will actually be put up on the website for the public. It's actually buried a little bit, so it's not easy to find, but it is on the ACTTION website. So you may want to be a little careful if you don't want your comments put out there for everybody to find on the internet, anyway.

Please sign in daily at the registration. These are Val's things. Obviously, silence your cell phones. It's being audiotaped; directly in the microphone. Restrooms are outside to the left. WiFi, select the Western meeting rooms on your browser, and ACTTION with 2 T's is the access code.

Page 25

Lunch and dinner is upstairs where we had dinner last night in the Mayfair Court. Our breaks will be done right here.

Any questions, comments, concerns what we're going to try to accomplish in the next two days?

(No response.)

DR. WARD: Okay. Nobody's had enough coffee yet.

Our first panel, Doug is going to moderate, and John and Yoanna are going to review where we are at this point from what PADIS came up with to get us started as the background.

Presentation - Douglas Coursin

DR. COURSIN: Good morning. I'm Doug Coursin, for the record. I'm taking my blazer off for my friends from Vanderbilt, but I do have a tie. I wasn't sure as a moderator what the role really was. I also wasn't sure if I was allowed to have slides. And I figured by the end of this we might be PowerPointed to death, so I was going to take a shot at doing it without slides.

A discussion moderator is a person whose

Page 26

role is to act as a neutral participant in the discussion. I have no biases, nobody's paying me to be here, know nothing as Alfred E. Neuman once said. But I try to hold the participants to a time limit and try to keep them from straying off the topic of the questions being raised.

Fortunately today, we have two of the world's experts in sedation and a host of other ICU related issues. They have significant experience in study design, reporting studies, guideline development, and publications in this area, and I will introduce them in a minute.

This is a broad area. And just to provide a little historical perspective, we live in the ongoing tsunami of guidelines. There was a guideline how to get here today, how to get on the metro, and how to get into the hotel. And we often encounter competing guidelines.

These guidelines were developed initially in 2002, and there are two survivors of the three generations. I'd like to recognize my good friends from Portland, from Maine Medical Center, Rich

Page 27

Riker and Gilles Fraser, who managed to give birth and participate strongly in all three generations.

In the late '90s, early 2000, the first generation was pulled together with a collection of experts, and the focus was purely on sedation and analgesia. Like so many other guidelines, I think ACLS is the best of all of them. I expect in another generation or two, they'll cover all of critical care. But in the case of the sedation analgesia ones, the whole area of delirium, altered mental status in our critically ill adults became an additional focus.

One of the buzzwords, which I think really is the core to what we're going to talk about here today and tomorrow, is patient comfort and safety, because I find when I participated, they jettisoned me after the second generation; probably a good move. I'd either lost so much hair and my beard turned so gray that I just couldn't stand up to the pressure

As they expanded things, they began to look more at the spectrum of what we do with our

Page 28

critical care patients and what the critical care patient brings to us with their comorbidity, their mental status to start with, and the like. In the third generation, they expanded from the SAD guidelines -- sedation, analgesia, and delirium -- to the PADIS guidelines, and to this they added in immobility.

I think the elegant work that Mona and others have done and the Australians have done, they're encouraging us to get people the heck out of bed and get people to maintain at least their musculoskeletal function as best they can and maintain their respiratory function as well, but also I think moving the clavicles up for us to more aggressively address our cognitive function in the intensive care unit.

With that, they added the "S" to the PADIS, and that is sleep, which is a whole other topic to discuss, an incredibly complex topic, and I don't think is going to be a particular focus of the group here.

So I just wanted to provide, the first

Page 29

guidelines came out in 2002. The next gestation was incredibly prolonged and painful. It came out in 2013. John and Yoanna did a spectacular job in herding an incredible cross section of cats to produce an expanded deeper guideline.

Each of the generations, in the first one, we didn't have anything like Cochrane analysis or grade, or PICO, which I think Yoanna and John will talk about. That came out in the second generation. That allowed us to focus and that facilitated trying to come up with evidence-based guidelines.

The problem with all of that has been where's the evidence? Show me data. Not the money, but show me the data, and show me the data in my patients, whether it's in Portland, at Tufts, or across the border with our friends in Montreal and elsewhere, what is the data? And what's your population like in it at all? Medical ICU, or adult surgical ICU, or God forbid, it's a subspecialty ICU.

Critical care is becoming more diffused.

Page 30

It's, unfortunately, in my humble opinion, likely, if we're not careful, to be more siloed. We're likely to have the CT surgical group, and the neurosciences group, and the widget group over here. We really need to work within that context because my drug that I advocate for may be totally reasonable in my population but not yours.

I think another thing that we have to take into account as we look here is that most of the drugs, save one that Mervyn and others developed, was never developed for the ICU. It was imported from someplace else. And I think from a development viewpoint, if I was a pharmaceutical executive, what would be my motivation to develop an ICU drug?

I think coming away from that, one of the things I'd be interested from the experts here to discuss is what do we really want in this hodgepodge of sedative, analgesic, anti-delirium, comfort-inducing medication? What are the indications? We haven't even gotten into the whole area that we have limited PK/PD data on many of

Page 31

these drugs. We extrapolate them from much smaller, much more tightly controlled studies in healthy volunteers or in much briefer exposures.

We give opiates for prolonged periods of time. Surprise-surprise; we see side effects. We see tolerance. We give sedative medications, and hypnotics that may cause physical dependence or potentially withdrawal. These become incredibly complex.

So I think the things that Yoanna and John can also point out to us is how they came about to come up with a host of recommendations but they really nicely identify what are our gaps in knowledge, and they are not insignificant. I think as we come out of a meeting and a lively discussion like this, that's really something we want to focus on as we try to move ahead and the future studies with the expertise of methodologists, biostatisticians, and of course clinical experts across the spectrum from physicians, nursing, pharmacists, physical therapy related individuals.

Critical care for those who don't practice,

Page 32

it is a team sport, and there's good data that as a team sport we do well if we have team leaders and team expertise. But most importantly, we do our best work when we have good communication.

So just in closing, I'd raised the following questions that I hope we can address at this meeting, if I can find where I listed them. What do we want from the medications and non-pharmacological interventions and protocols that we generate?

What properties would the ideal agent, or agents -- if they're going to be pharmacologically mediated or non-pharmacologic approaches, what would they look like? What would they give us? What's likely the best way to get at developing something new or taking what we already have and identifying the right patient, the right intervention, and the right outcome?

With that, I'll close, try to maintain my neutrality, and turn to the experts. Our experts today, Yoanna Skrobik from McGill in Montreal. She has so many titles and degrees I can't go into all

Page 33

of them. At some level, She's a molecular genomicist. She's an intensive care physician. She's becoming increasingly an addiction specialist, which I think is very germane to our practice in the ICU considering the average ICU patient is receiving what, Gilles? Would you say 10 to 14 medications a day?

DR. FRASER: That's the bottom.

DR. COURSIN: Many of them as continuous infusions; many of them with very under-recognized central nervous system effects, and I think we need to keep that into account.

Our other expert -- and they were the co-authors, and John was really the driving force in this and had agreed, last night I heard, to do the next generation. Thank you very much, John.

(Laughter.)

DR. COURSIN: On behalf of the board of directors of SCCM, we thank you. John is a professor at Northeastern and a professor at Tufts Medical School and brings a wealth of knowledge.

Yoanna, I think you have slides for us.

Page 34

DR. MAZE: Can I just ask a question of you, Doug?

DR. COURSIN: Yes.

DR. MAZE: As an interloper who hasn't been involved in either the previous iterations, I'd like to understand how you went from SAD, to PAD, to PADIS, because I think that is clearly not the end of the acronym.

DR. COURSIN: Rig

DR. MAZE: And I say that because if you now have, as you've identified, a post-injury condition, surely there's something you've missed doing or not doing in that critical care period.

DR. COURSIN: Mervyn, I think you've hit a nail right on the head, and I'll have the other speak to this more eloquently than I can. But SAD, I think the beginning, which was sedation and analgesia, and they tweak out from that the use of paralytic agents, neuromuscular blockers.

There's a whole guideline on this, and I think Gilles and John and others from the PharmD world would agree that the use of the paralytic

Page 35

agents in the ICU has gone like this [gestures] over the ensuing two decades, unless you're occasionally using something like cisatracurium, therapeutically, not just to paralyze patients, with really severe ARDS.

So they started I think with a very specific focus, and I think we had a naive -- Gilles and I were heavily involved in that first iteration with Judy Jacoby, a former president of SCCM and a very gifted PharmD. Our focus was, well, we came from an era where everybody got high-dose morphine, high-dose valium, and high-dose vecuronium.

Probably 10 to 20 percent of our patients back in the '90s were being paralyzed, so everybody worried what could be worse than paralyzed and not adequately analgesed [ph] and sedated, and we began to see a lot of very strange things occur. We have of course all the issue with tolerance to the morphine. You give a big slug of valium. You give valium or lorazepam as an infusion. It's dissolved propylene glycol. You have issues with renal dysfunction, metabolic acidosis, and then you have

Page 36

drugs that have extremely long either lives or very active metabolites.

Then you look at vecuronium and you look at the world that we lived in at that time with renal dysfunction with a drug that has 3 active metabolites that are renally excreted. So we started to see these very weird post-paralysis myopathies, any one of the number of things.

We also got lot smarter in the way we ventilated people, and we woke up to the fact that maybe it wasn't such a good thing to have people just flat in bed, stone cold, not moving. We very simply moved to one of the key bundled pattern things, which is unless you can't do your patient care at least at 30, so you limit the aspects of aspiration pneumonia. You limit the development of ventilator-associated or hospital-acquired infections.

So I think the first charge, what we very simply thought -- and correct me if I'm wrong, Gilles or Rick, we thought, well, let's get a handle on sedation analgesia. Let's come up with

Page 37

some recommendations. Propofol was just coming into its own. Etomidate had fallen off the map because of its issues that you and others looked at with adrenal steroidogenesis.

We now have a short-acting benzo that, quote/unquote, "did not have active metabolites," which is not true, and those active metabolites were never going to be a problem when, one, hydroxy midazolam actually can accumulate. But we were used to giving midazolam in the operating room as a pre-med, a couple of milligrams or in the endoscopy suite and get on your way home after your colonoscopy. We were given 10, to 20, to 30 milligrams an hour of midazolam, not for an hour or two, but for days.

Our length this day over the last 20 years has gone like this [gestures]. Our length of stay in a major medical center is under 4 days. Now, that doesn't mean that you're not critically ill when you go out the door, buy you may go out the door with a trache in place that we percutaneously put in, and you're either going to go upstairs to

Page 38

our intermediate care unit where you can be ventilated on a trache, or you're going a mile away to our LTAC, long-term acute care hospital.

So I think what people saw then, I think Pratik, and Tim, and Wes Ely and others, Yoanna, folks from Britain and the continent, started to point out that people had really strange recoveries when it came to delirium and cognitive function. I think it started to come out that delirium wasn't just me up here acting out or going into withdrawal, but that it was a hypoactive delirium, that this was very common, that we were under-recognizing it.

Just editorially, I have good opinions that tell me we grossly under-recognize pain in the ICU. We also grossly under-recognize what the patients and the families perceive of things and how their needs to communicate may have changed.

So I think what happened in the second generation, two big things. One was people became aware that delirium and post-op cognitive dysfunction was a major issue potentially in the

Page 39

ICU. What was the role of sedative and analgesics, either inappropriate utilization of them or prolonged utilization, or any one of a number of factors, in playing a role in delirium.

I think one of the things you are expert in and I know very interested in is what about all the other things that have gone on, comorbidities, inflammatory processes, surgical procedural intervention, that in and of themselves may create a delirium situation or a post-ICU cognitive dysfunction.

DR. DEVLIN: I didn't mean to interrupt. I was just going to add a couple of thoughts, too. Sorry.

DR. COURSIN: A couple of what?

DR. DEVLIN: I was just going to add a couple of thoughts additionally.

DR. COURSIN: No, please interrupt if you --

DR. DEVLIN: No, no. I didn't mean to interrupt. Sorry.

DR. COURSIN: Okay.

Page 40

DR. DEVLIN: I think the other thing with guidelines, that we firmly believe it's an instrument for change, right? We want to be making change at the bedside for all these things, and they're so interchangeable that we felt, when we went to the Board of Regents at SCC and to propose this plan for having five sections, including a large immobility section and a sleep section is, again, the interchangeability in clinicians at the bedside don't necessarily put them into these particular buckets: why is the patient awake at night; why do they have the ICU cardiac weakness, et cetera?

The other thing that really came out from PAD 2013 was we had questions not necessarily focused on immobility, but Gilles and I worked on the part where we were looking at ways to reduce delirium, and of course JP Kress' landmark study had come out in terms of early mobility, and we put that in the context of the guidelines as a way to reduce delirium. But obviously the far bigger question is we need to really tackle this

Page 41

immobility thing.
So obviously with Bill Needham's leadership and many others, that's why that's all included, and we felt if we don't do this, these are bedside patient-derived issues. Again, the PADIS guidelines are all focused on patient symptoms. That's what PADIS stands for.
We thought if we didn't bring that context in here, even though realizing -- for example, with sleep, that there's just so little data and such a complex area, that if we didn't start to define that, and point out gaps, and drive people forward, at least it's going to help clinicians think about these things and what they should or should not do.
Sorry. I didn't mean to interrupt, but those are just some important --
DR. COURSIN: No problem.
DR. DEVLIN: -- that sort of came along.
DR. COURSIN: Just for Denham's benefit, that was Mervyn Maze asking us about other areas and how this expanded, and John Devlin weighing in.
DR. WARD: I can recognize Mervyn's voice.

Page 42

DR. COURSIN: Okay. Excellent.
(Laughter.)
DR. COURSIN: But I think one additional piece to the development of both SAD guidelines and then the PADIS guidelines was that they began to bring in a way to try to come to a collective recommendation and a quality of that recommendation to answer very focused questions that had not been undertaken in the first one. That's a very interesting process. The panelists will address both the PICO approach and the use of grade because it's not as if there aren't issues with that or controversies.
There's always the age-old issue in any of these of getting a collection of experts together and having the wallflower up against a strong personality or the aspect of we really don't have much data, but our constituency wants a recommendation.
I think one of the final things I'd comment on is I'm quite interested as an observer to see how these guidelines are actually applied, and I

Page 43

think that has some unintended consequences regarding either people's development of protocols or people's interpretation of the guidelines, particularly when John and Yoanna spent so much time with the collective group trying to say what is the basis of the guideline and what is the quality of data from this guideline, and how should we interpret that?
As you look through their pages of discussion on these, there aren't a lot of really high-grade 1A recommendations, and I'm interested in my own community to see how relatively conditional low or very low database guidelines, how they're applied in my institution. In very short order, in the world of protocolization, they get chiseled in stone, so as an outgrowth additionally of the PADIS guidelines, SCCM has put together A, B, C, D, E, F of a bundled guideline approach or extrapolation from guidelines, and I'm really interested to see what comes next, E, F, G, H. I'm trying to think of things for X.
DR. SKROBIK: I'd like to say, Doug, we've

Page 44

infiltrated them.
DR. COURSIN: Okay. Excellent. So we'll go from there.
Yoanna, I believe has some slides and overview and discussion, and I appreciate people's questions, and hopefully we can provide some useful information. Yoanna?
Presentation - Yoanna Skrobik
DR. SKROBIK: Thank you. My name is Yoanna Skrobik. I have the privilege of having been invited to vice chair the PADIS guidelines with John. It's daunting to stand in a room of people this smart, and it's daunting to summarize, in what I think is a short period of time, what I would like to be a summary of what we did in the guidelines and an invitation to come up with an actionable methodology, or two, to invite the next generation to do better or to do differently.
So I would like to invite all of you who are doing something else to set that aside for maybe 15 minutes and listen to the content of what we did, in summary, but more and more in the discussions

Page 45

that we've had, what we didn't do, and try and perhaps come up with one or two suggestions of your own so that when we come together into these groups, rather than have a speaker and content, to open up the discussion.

We are not just privileged speakers. There are a lot of smart people here. As I was coming up to this magnolia flower-filled neighborhood, I was thinking about the privilege of what money buys and how lucky we are to have the partnership that was set up by Dr. Denham and others to think; the luxury of being able to step back.

So I hope to be able to honor the people in this room and the process by at least helping with one or two deliverables, and considering how many ideas I have in my head and how chatty I am, it is going to be a challenge. So what I would like to do is summarize very briefly what we did and highlight what we're proud of.

When we brought it rehabilitation, the reason that John said GP's important work, some of us had small children also, so we thought if you

Page 46

move people, they're going to be more relaxed, and then they're going to sleep better at night. It's artificial to dissociate sedation and sleep. We said we wouldn't address sleep today, but you all know that in the clinical environment, one of the most important reasons to administer sedatives is sleep.

So all of those topics, as Dr. Maze said, our confluent, and as Dr. Egerod softly pointed out to me last night, it's not perfect, and there's some uncertainty.

I think we were delighted to have patients both as collaborators and c-oauthors because we learned so much, and we brought in experts from Europe and Australia to a traditionally American bastion. Dale sweetly pointed out that we had not included other continents, but I think that it was something to be proud of that we had at least broadened it a little.

We were particularly also proud of saying not only what is but what isn't, and saying why it isn't, because we thought that was really

Page 47

important. So it's the first guideline to do that within the SCCM. In the supplemental material, we hid in the 84 pages of the online supplement, we also published how people voted on each of the recommendations. We snuck comments into the grid because we thought it was really important to not be proscriptive, and if we were going to consider the reality of contextual and clinical variability, it's not true that there's always one right thing with a capital R and capital T.

So those were our attempts to do better. We also came up with some new questions. What I wanted to stay in the retrospective after all that work, we realized that we had gone over some of the topics anew, and some we had not, and I'll give an example of that over the next slides; 37 recommendations. It was 2 ungraded practice statements and 32 ungraded statements, and I'll speak to that very briefly.

We use the grade method for ranking data, and therefore favored RCTs. We did not consider qualitative data. We didn't find a way to

Page 48

integrate the patient experience in a way that was methodologically sound, and we acted as if when intervention in a heterogeneous population actually applies to other people.

We finalized PICO questions and came up with what we produced in 5 electronic databases, 2 awesome librarians, a lot of people who supported the effort on every level, and we had the discussions in the five sections.

Each of the discussions were attended by John and I, or both, and then came together for the final wrap-up in Hawaii in 2017 and spent a day and a half hammering out what it was we meant, because one of the things we learned the most clearly in the process is that in the communication piece lay the clarity in the final recommendations. So part of it had to do with explaining how we had written was actually intended, and I think we achieved that

We had a lot of agreement, but it wasn't perfect. I really like the fact that it wasn't perfect because there are reasons to do things differently, and I think we should celebrate that

Page 49

diversity. I want to acknowledge that without John's rigor and enthusiasm, because I've always wanted to have what he puts in his coffee in the morning, we would not have had the performance metrics that we had.

We delivered the guidelines on time and with a hundred percent participation in each of the recommendations regardless of whether people agreed with them or not. And I think that is to the creditor of our fearless leader. I think we also honored the ICU survivors, and I'll never be sure whether we did it enough, but we tried.

We used PICO questions for the recommendations that we made, and I wanted to give you two examples of how we did that that are relevant to the sedation issue, and then give you one more recommendation to think about.

The pain assessment and management question, should a protocol be used, was one of our PICO questions. We said you have to differentiate between analgesia first and analgesia-based sedation, meaning you do your analgesia first and

Page 50

then say, "Ooh, do you still need sedation?" We're playing you soft music or rap music, if that's what you prefer. We're massaging your feet. Do you want a drug on top of that, versus using an analgesic as a sedative.

When John and I were discussing this more recently, we thought, well, that would be an opiate, wouldn't it? Would we have phrased it the same way now that opiates are front and center as being potentially problematic and potentially problematic in terms of their effectiveness as analgesics.

So we then delivered improbably this good practice statement that pain should be guided by a routine pain assessment. I think it is actually extraordinary that all we could do is come up with a good practice statement because it would seem humanistically that it doesn't make sense to do anything else.

So you see where I'm highlighting all of the caveats because if you use a framework that requires RCTs, what kind of caregiver wouldn't want

Page 51

to palliate pain? In fact, the data suggests that not only don't we evaluate it very well, but we don't manage it all that well.

So here we are talking about sedation and analgesia for sedation, and maybe the analgesia part and opiate part are not so straight forward as we thought. So we suggested -- and this is the content of the guideline -- that there be an assessment driven and protocol based approach, but we lumped analgesia and analgesia-based sedation.

The process that we used meant that all of the patients ranked things according to priority, so the pain part before giving sedation was hugely important to patients, and I just want to highlight that. If patients thought that that was so important and we say we're doing patient-centered care, how are we going to incorporate that in what recommendations actually say or don't say? Where is the place for the patient's voice?

I'm switching gears now to the actual notion of managing agitation and sedation. We've come to understand -- thanks to the work of several people

Page 52

in this room, probably the most compellingly -- that sedatives are actually not very good for you short term; not long term. We have some notion that long term, they may increase cognitive dysfunction, but how and so on.

We state in these guidelines that a specific indication for giving sedatives is imperative. Nobody asks the question. Your patient rolls in, they get delivered a drug, but we stated that pain should be addressed first and then sedatives should be given; that there should be a reliable scale; and that adverse events should be thought about.

In the gaps that we identified -- and this list is very long -- pathophysiologic state, so inflammatory states were blood-brain barrier permeability may not be the same for drugs that are potentially toxic like sedatives; reduce drug clearance; PK/PD that have been studied extensively in children but not in adults; drug-drug interactions, which some of us have modestly been interested in and every pharmacist knows about but aren't necessarily integrated into how we practice;

Page 53

how individuals respond.
I think Pam Flood rattled me. She described the subjective sensation of getting dexmedetomidine versus propofol during one of the panel guideline meanings and described her husband's reaction to the two drug exposures. We had no room in our guidelines for integrating how you feel, so all of you who have taken an opiate, a sedative, know that different ones do different things to you. Where is that in the way that we practice, and does it really matter? Does how you feel about it matter at all?
Of course, genomic epigenomic factors are huge because we are starting to understand that they play a huge role in drug metabolism. All caveats that we listed with specific, we were not able to address or answer these questions.
We looked at short-term outcomes in the 2013 guidelines. We tried to look at long-term outcomes in the 2018 guidelines to speak to Dr. Maze's question, and we hit the wall of the lack of information, and the lack of precision, and the

Page 54

lack of rigor of consistency across studies and how that was done.
We looked at all the topics in all the sections based on rank order. The experts said this is what we think is important, and then we handed it to the patients. So the order, for instance, for the pain section was dramatically altered by the patients; most of the others were not.
Here are the most important ones for the sedation group. Sedation and clinical outcomes was considered to be the highest ranking, and then the sequelae of lighter versus deeper sedation. I'll speak to the light versus deep sedation because it also highlights some of the questions.
We are looking at 15 years of literature. We were able to find 8 RCTs and 3 observational studies. So we're making recommendations for the universe based on a relatively small amount of data and end up saying we should be using lighter sedation versus deeper station, but we don't know what light means, and Pratik could talk about this

Page 55

for an hour and a half, I'm sure, if he was so inclined, but he's not.
How do you define where the harm line is? Is it an average, over 48 hours, or is it one moment where you're completely -- the anesthetists in this room will tell you that you get post-operative cognitive dysfunction.
How is that different from the exposure of the sedatives or the opiates that we given in the intensive care unit where we give more drug longer than most places. How that changed over time and impacts people in the long term is also not clear, and how do we describe what happens to patients?
I was listening to Dr. Brown casually say describe the fallout from the intensive care experience that he had and that his family experienced. How do you measure it and how do you say that it matters?
Judy Davidson from the family-centered guidelines taught me that 25 percent of families from ICU survivors are not back at work 6 months later because they are too burdened by the caring

Page 56

and the psychological fallout of having had someone you love go be near death. How do you measure that economically? We're looking at hospitals costs. What about the impacts on society, and should that matter? So we didn't go there, and I think that there's the patient specific factors.
When we asked the specific questions, we said for medical and surgical ICU patients, so non-cardiac surgery patients specifically, should we use propofol or benzodiazepines, or dexmedetomidine versus benzodiazepines, or dexmedetomidine versus propofol? We sat and talked. They said, okay, so it's meaningful. What would make you choose one or the other?
That was one of my favorite discussions. What do you think? What does your nurse manager think? If you're occupying your bed for 4 more hours or 4 less hours, it doesn't change the nursing shift. And the definition that we came up with were an agreement between the patients and the clinicians saying if you lighten up in that much faster -- 4 hours faster is what we decided,

Page 57

completely, and I think it was reasonable but we made it up -- what do we think is a significant shortening of extubation time?

The patient answers to this were the most interesting man. They said, "Well, I don't really care." So it highlighted that all of our metrics, duration of mechanical ventilation, mortality, the patient said, "Well, if I am better and I have to spend one more night on the ventilator, then I don't really care." I was thinking, "My God, and I couldn't talk and express myself," so there you go. So much for my understanding.

So the recommendation was that we use either propofol or dexmedetomidine because benzodiazepines had problems associated with them that are well described. But we were not able to define what long-term and patient-centered outcomes were, and the meaning to survivors was something that we couldn't quite put our finger on. We learned from Pam and others that patient perceptions were something that we were not able to methodologically capture, and the pharmacology piece was hugely

Page 58

missing.

Costs were the subject of huge discussions also. Here we were with an Australian, or two, or three, and Europeans, and Canadians, and the Americans. I don't need to tell you that melatonin costs a very different amount in each of these places; that each of the drugs cost something different in each of these places and how does that compute into what you end up deciding.

I think Dale has also highlighted the Third World's application of what we say. They're cost limited in a way that we don't consider, and the whole question of analgosedation and patient subgroups. These were the gaps who identified for the sedation choices.

I want to acknowledge all of the people who made this possible. It was hard work, and when I was on phone calls with a group beside my dying father, there were times when I wondered what I was doing there. My father was a man who liked things that would be delivered so that they would serve to build something else, and I think that all of the

Page 59

people in the PAD guidelines who spent these five years with us all had moments where they could have and would have been doing something else and chose to contribute.

So I would like to think that we can honor this work and take it a step further, and I would particularly like to thank the patients who were not only part of it but engaged to the very delivery of the manuscript and contributed to it even more through that. Thank you.

(Applause.)

DR. SESSLER: Dan Sessler. Did you address how to measure sedation?

DR. SKROBIK: Pratik, I don't know if you would like to speak to that. We had addressed the scales in the previous guidelines, so we had done the psychometric qualities of the sedation measurements. How to measure sedation is a wider question then that. Pratik led the sedation group. I don't know if --

DR. PANDHARIPANDE: Pratik Panharipande just for the recording purposes. We did tackle it with

Page 60

regards to the scales, and because there was no difference between evidence between the 2013 guidelines and the 2018 guidelines, which recommended using either the SAS or the RASS as the two scales with the greatest psychometric properties, we did not address that separately in the 2018 guidelines because there was new evidence to suggest anything should change with regards to that.

The area that we tried to delve in deeper was within the context of the scales, how do you define light versus deep sedation? I'll touch on some of that tomorrow as well, but that was an area of debate and a fair amount of discussion because the literature doesn't clearly articulate what is the best definition for that. In each of the studies that targeted light versus deep levels of sedation either used different scales or used different cutoffs for that.

So in general, we taught, based on what we read, somewhere between a minus 2 to plus 1 on a RASS scale and equivalent [indiscernible] on other

Page 61

scales was what people defined as light sedation. But again, that was an area that was relatively nebulous because none of the studies actually targeted that. So that would be an area that we will discuss tomorrow as far as what may be ways to try and determine what is a definition of light versus deep sedation.

DR. DEVLIN: The other thing to add -- this is John Devlin -- Rich Riker led an important descriptor question, too, on objective sedation assessment as well. It wasn't an actionable question, but I think it was a great summary of where we're at in some of the pluses and minuses of incorporating that in the ICU. I just wanted to add that.

DR. COURSIN: Denham?

DR. WARD: Denham. Thank you. This was a great summary, and it brings lots of questions if I was sitting here with a new drug in my pocket that I wanted to get approved. I'll just start with one to follow up with Dan's.

I've decided to use RASS as my measure in my

Page 62

clinical trial for a new agent. How much do I have to worry about the training and quality assurance of my people who are measuring RASS? Do I have to in my clinical trial -- my experience came from procedural sedation, and Roche when they came out with midazolam and flumazenil and advocated the use of the MOAS system, they actually had a training video that they produced.

You had quite a long training video. Then they had examples at different levels of scores, and you had to score them and compare it to the experts' scoring. Then as you used the scale, there was a quality assurance program to go through and make sure that your clinical trial person was actually using the scale properly for that.

So if I'm designing a clinical trial and using RASS for my sedation measurement, how much do I have to worry about the training and the quality assurance of the people who are making that measurement?

DR. SKROBIK: I think at the sake of sounding like I'm always making things more

Page 63

complicated, if you don't know that your pain assessment was done properly, how are you even going to go down the sedation road? The data that we have that are current from Lisa Burry's work and others in the Netherlands, and in Canada, and in the community, and in academic centers, suggests that nurses assess pain maybe 50-60 percent of the time in ICU patients. And when they do, their documentation of it is different than what the patient reported.

So to answer your question, I think I would add a layer to it and say you would have to mandate in every sedation protocol that pain be measured first and that it be tracked because in the same way that I think I drive a car better than --

MALE VOICE: Doug.

(Laughter.)

DR. SKROBIK: -- we all have the sense we do things well.

MALE VOICE: You've got your license?

DR. SKROBIK: I came close to losing it on the way to the airport. This was not Yoanna

Page 64

Skrobik.

So if we measure what we're doing and then compare each other in a trial specifically -- if you're doing a multicenter trial, I think it would be interesting to say how often and how reliably are you measuring whatever it is that your bedside metric is, and I would hope that that would improve the overall pattern of care.

DR. COURSIN: If we could just hold on a second. Steve and David and Claudia I know have questions. I'd just like to make two comments. One, Dan, I think you're absolutely on the money having a standardized and reproducible technique, and I think there are three others in the audience I'd like to hear from about that in the studies they've done, SEDCOM and others

I think the other issue is, is there something available objective in the way of a neuromonitor that would give us -- and I think that is a whole hand grenade verse of conversation.

But Pratik and Rick and Gilles, in your studies, in multicenter studies, how did you

Page 65

control for the quality of SAS or RASS, if you happen to be a Richmond guy or a Maine Medical Center guy, in your analysis? How did you control for the quality of their subjective scoring?

DR. RIKER: Riker. For SEDCOM, as part of our startup meeting, we actually had the folks from Vanderbilt who developed or validated RASS and developed CAM-ICU, spend time with each of the research teams to train them in that process. We didn't do secondary confirmation of reliability or anything like that at each site. We didn't go that far, but it was included as far as our startup meeting for training.

DR. COURSIN: Sir, yes? You had a comment to that?

DR. SHEHABI: I just wanted to add, Rich, I think it's very important that the sites get trained specifically on site to control the quality of conducting a pain and sedation and delirium assessment. Like what Rich did in SEDCOM, in sites where we ran it in 74 ICUs around the world, the team visited every single center to train them on

Page 66

how to conduct these tests.

We left them with videos that they can use at the bedside with PowerPoint presentations, and then we had a study monitor who visited every single site at least twice during the conduct of the study for the quality control of the data and how they're doing it. I think that's very important in terms of making sure that the frequency visits are done as supposed to be done and they're done in a standard fashion across all sites in a large multicenter trial.

DR. SKROBIK: Can I ask, other than the social engagement that you make when you connect live, do you think it's feasible to do that more cheaply through electronic platforms or through more pragmatic -- we talk about the cost of doing RCTs and how huge it is for results.

DR. SHEHABI: I think it took quite a while. It took us two years in doing that. We introduced the site in a target fashion, so they were not all started on April 1, and we had multiple people who were doing that, visiting the sites. Like for

Page 67

example, at the Malaysian site, there were 11 sites there. It would have been impossible to do that on a video call.

That engagement at the beside training for the research team and the senior clinicians was very, very critical for them to understand what they're expected to do. Even in the UK, the sites there preferred onsite training, so we conducted at least 5 centralized meetings in the UK for that purpose.

DR. COURSIN: I want John just to get a chance to jump in.

DR. DEVLIN: I think the other thing that Yoanna and I have had a lot of discussions, two particularly with delirium assessment, is nurses I find want to know does my patient have delirium or they don't, and they're challenged, and it can add a little bit of stress to them as "I'm not really sure."

I think it's important through the education to give them that knowledge that it's okay that they're not sure exactly what the RASS score is if

Page 68

the patient's CAM positive, but then to seek out someone else in the unit because this could be a night nurse who might be maybe a better trained colleague that can really help.

Certainly in the research I've done it at Tufts, we've really promoted that, and it seems to have prevented a lot of not able to assess or not really sure. I haven't done research on it, but I think it's helped the validity of some of the assessments or at least let the investigators know the next day that they weren't sure. That was just one thing to add.

DR. COURSIN: I think in part your local research coordinator has to be working. I'm in a unit probably typical of this group. We have 93 nurses in our unit, and we have float nurses and we have nighttime replacements. So it's not a small thing, but I think you need a series of champions as well.

I just wanted to give Claudia a chance in the back. She's patiently been waiting to comment. Thank you.

Page 69

DR. SPIES: Thank you. I have one comment to the discussion right now. We're doing a study now with the German Ministry of Health involving several centers. And our impression was in the beginning that all these measurements on sedation, analgesia, and delirium didn't work if you really tried to give it by guidelines or give it by e-learning.

So we implemented a blended learning concept where we have e-learning as a beginning so people know what they are talking about. So that's very important because of the different professions involved at the patient's bedside. And even the relatives and the patient, him or herself, are always concerned. So if you train the people, they perform better.

Second, we have a simulator-based concept, so people are not corrected at the bedside. They don't feel annoyed. Sometimes they feel annoyed if you do that. So I think that's not good to do it that way. So we have a simulator-based concept, and at the end you do supervising at the bedside.

Page 70

This at least decreases the inter-rater. Also the inter-rater, we have a lot of variability in that setting, and I think it's very important that we -- some think we do the same things, but we don't.

The other point I would like to address to the methods of the studies. If you want to address pain, I think you need a pain measurement, and that's not done in many of the studies. The second point, if you want to try sedation, you need -- all the things that are not clearly stated in the beginning. And that's giving you a lot of confusion because sometimes if you measure pain, you measure side effects of sedation and not pain itself. So this is complicated, and I think there are a lot of things we need to consider together to release all that out.

DR. SKROBIK: If I could just add to that, I think the other caveat that we thought about later is that the notion of benzodiazepine withdrawal, for instance, is not something. In children and in the pediatric population, opiate withdrawal and

Page 71

benzodiazepine withdrawal are routinely measured.

In the trajectory of an ICU with all of the drugs that we deliver, we actually don't test for it very often in adults. To my knowledge, there are two or three studies describing it, suggesting that the incidence is as variable as 15 to 50 percent, so not an insignificant amount, and I think in the don't knows, I think that's part of it.

In Canada, there are three provinces now that have electronic registries where you can't close your days charting if you haven't done the pain measurements and the sedation measurements. I'm not sure the content of what's written in there reflects what the patient actually has or doesn't have, but I was grateful enough to be part of the quality assurance of setting it up for the delirium component and for the three components of sedation. So there are ways to make sure that the assessments are done, and in each of those platforms there's an uncertainty box.

DR. COURSIN: You have to be careful, I

Page 72

think, of the old VA jokes that the patient had normal vital signs an hour after they were declared dead. Any of the experts who are boots on the ground in developing these scores in the audience for implementing or making sure that they're implemented have comments? Michele?

DR. BALAS: Michele Balas. I think one of the things that we talked about significantly on working with the PADIS is the importance of conceptual clarity in what the symptom is that we're measuring or the syndrome.

For example, one of the most frequent reasons clinicians report giving benzodiazepines for is for anxiety, and many of the studies that we've reviewed or that we include that involve sedatives have rarely measured anxiety. Anxiety gets lumped in with agitation. Agitation sometimes gets lumped in with hyperactive behaviors, and sometimes the patient that's deemed agitative is actually having a normal response to being restrained and having 15 tubes put in their body.

So I think it's really important. From a

Page 73

clinical trials perspective, I absolutely agree that fidelity monitoring and the inter-rater reliability is something that we definitely need to build upon in the ICU community, but also the need for that conceptual clarity when we're looking at the symptoms; not just the outcomes but the symptoms that we're looking at.

DR. COURSIN: Thank you.

Steve, you were jumping in, and, David, we'll get to you. Steve's been patiently waiting, and so has David.

Steve Shafer?

DR. SHAFER: I'd like to step back for a second. Steve Shafer from Stanford. I'm not an intensivist, but certainly your paper from last year in Critical Care Medicine is just a wonderful piece of work outlining both recommendations but also the gaps in the knowledge.

One of the things that jumps out to me is there are so many gaps in the knowledge and so many things. I went through and made a list of all things where it says low-quality evidence. And

Page 74

since we're here to talk about clinical trials, I think that one of the things we're here to talk about is how do you fill in this low-quality evidence that really dominates both sedation and analgesia, particularly, in my view, the sedation piece? Because it costs money to fill in this evidence, and the question is what's the economic driver for it?

I can envision several economic drivers. If there is a new drug that a company can come forward with, that will drive more studies. Hopefully from the company's perspective, it will be studies that show favor of the drug, but it will give us more information to fill in these knowledge gaps.

The problem is the drugs that we have are generic. They all come out of anesthesia, and there have been a number of attempts in the last 20 years to come up with the next generation of propofols and dexmedetomidines, and other sedatives, and nobody's done it because the market isn't big enough and the regulatory path seems too burdensome. And frankly, the drugs we have are

Page 75

awfully good, so they set a high bar.

That's one of the drivers. I don't see that happening. Other drivers would be physicians, perceive a gap in care. And I'm not one of these physicians, but I don't get a sense that our ICU doctors are saying we have these huge gaps in care. You've identified the gaps and the knowledge, but are the physicians saying -- the boots on the ground in the ICU -- we need these gaps of knowledge filled or there's an economic gap.

We could do better and we could pay for this work if we could save money by doing these things, and that would fund the studies. What is the economic driver to fund the research to fill the pretty overwhelming knowledge gaps that you identified?

DR. SKROBIK: I think that what you speak to is exactly that. We have a 4 percent error rate across our medical systems no matter where or how you look. We don't acknowledge it. We don't talk about it. We don't apologize for it. We don't fix it. So in addition to saying we are not perfect,

Page 76

you want us to say and maybe we don't deliver, and it doesn't make sense.

On the other hand, you talk about making money from an intervention. How much money would you save if you delivered the care according to whatever simple metrics? Not the sexy new molecule that is going to make my wrinkles go away, but the delivery of what the patient expects.

I don't know where that answer lies because when I listen to your question, I would think that incentives to make people more accountable for the delivery and the costs of the care might be an interesting perspective.

DR. DEVLIN: I just wanted to add really quick, I think with the caregiver, it's a really, really good point you brought up. I think it also depends on the paradigm perspective of the clinician, what they feel is the goals of care and whether they truly are well versed on some of the dangers of deep sedation and the mobility, not being able to mobilize patients and as such.

Obviously, in the U.S. at least, we're

Page 77

dealing with a system where we're focused on the cost of care in the hospital, but obviously this spills over post-ICU, and readmissions, and everything else.

DR. COURSIN: Steve, one observation I'd make -- you've raised excellent points. Who's going to spend half a billion dollars to bring a drug to the market place that doesn't have the multiplier of the next generation of statins, or Z-Pak. The second piece to that, though, I think that for the most part, in critical care, we are pretty satisfied with what we have and what we have that comes out.

I think where we have real blind spots are -- what we've clearly outlined nicely how the patient and their families are functioning, and in a sense, the ICU with the shortening of stays and whatnot, and the throughput that we have, a lot of the problems that we're discussing here, they're out of sight or out of the mind to the critical care patient at the bedside.

I don't realize that 25 percent of families

Page 78

are not back to a functional state working. Their quality of life is impacted a year later, and we haven't even gotten to the root cause analysis of what about the huge percentage of patients in the United States, anyhow, who shouldn't be within 5 miles of an ICU, but yet get admitted? So I think those are key issues.

JP, you had popped up with something. I wanted to make sure I didn't oversee it.

DR. KRESS: One thing -- and maybe we'll talk about it later -- this gap in what we currently have and what we're seeking in terms of quality there or continence and the recommendations. I think it's important that if you look at the way that the grading system for these consensus statements is used, it's a really, really high bar to get a strong recommendation.

If you look at the published guidelines for many, many different areas, what percentage of grades are low quality or weak recommendation compared to strong? I would submit it's probably more than 10 to 1, and maybe that's because the

Page 79

system is intentionally stringent.

If we're trying to improve on the quality of our recommendations with better data, that's a really high bar to jump over. Having sat through these meetings and these consensus guideline writing, you almost have to hit a grand-slam in a particular area to get a strong recommendation. So if we're hoping to get the next generation of these guidelines with all strong recommendations, I think it's almost, if not certainly, impossible based on the system we use to give studies and recommendations grades.

DR. DEVLIN: That's such an important point, JP. I think the other thing too is we're framing our guidelines, PADIS, for all critically ill adults. This is just one example. So then we downgrade things when there's not a patient population that's been well studied, which is all different subtypes of patients of critically old adults. So that's an automatic downgrade when there could be a great randomized study and a good answer potentially in a subgroup of, say, a certain

Page 80

type of surgically critically ill patients. We ran into this all the time in our guidelines.

DR. SKROBIK: Could I just speak to the grade comment? We have huge discussions over the grade methodology over these guidelines, and I think it's a very interesting and important point. I'm not sure how it influences trial design because if we're going to be asking the questions within the trials, perhaps that imperfect metric should be set aside altogether because in itself, it is the best tool we have so far. But for the very reasons and many others that you've point out, it has major limitations.

DR. COURSIN: Dan Sessler?

DR. SESSLER: It's true that with some current systems, it's hard to make strong recommendations, but that's not a fault of the system. It's because we don't have the underlying data, and there is a bit of a history of groups coming out with fairly strong recommendations that didn't hold up, and you only have to look at the recent World Health Organization recommendation on

Page 81

supplemental oxygen, which defies available data, so there's a lot to be said for being rigorous.

Along those lines, you presented a very formal way of developing consensus of doing a full systematic review, grading everything, voting, recording how people voted, making sure that people don't vote if they have a conflict, which might even be defined as having done for relevant research in the area.

DR. COURSIN: And we did do that.

DR. SESSLER: Okay. That is becoming a standard. It's the way we develop the Canadian Society of Cardiology guidelines. It's not what we're doing here, which is just something to think about. I mention it because I was involved in a PCORI consensus process and papers, and we got huge pushback from reviewers that basically said this is no longer the way it's done, and frankly, I think the reviewers were right.

So going forward, we might think about doing this a little more formally so that we are at the current standard of care.

Page 82

DR. SKROBIK: I think it can be said of patient representation, if I could just add that, because Dale has done some very elegant work using groups of patients that were representative of populations in using Delphi rounds and going through a very rigorous process.

We got slammed in the guidelines for having randomly apparently selected patients who just happened to wander in and want to donate that much volunteer time. Can we agree? It couldn't have been -- if you would have asked a representative sample to do this thing, they would have told you where you can get off the bus.

I think the most elegant response to that was Cheryl Misak's, who said, "What is your presentation anyway?" And who speaks for whom and under what auspice? And I think what Dr. Ward was talking about earlier, benevolence is not a small -- it shouldn't be set aside.

Pratik and I had a very lively discussion over the day yesterday in terms of what intellectual conflict of interest means. The key

Page 83

is the transparency and the communication. If you can say we did this transparently, this is what we did, this is exactly what we did, you can knock it, but you know what it was.

The communication among those who share ideas the way we are has to be about the content and not about your opinion and whether somebody should be getting sedatives. We had one interesting intervention, and that was why would you give somebody a sedative anyway? You say that to most people, and --

DR. COURSIN: David, I wanted to give you an opportunity. I apologize. David Brown.

DR. BROWN: Again, I'm going to wear my patient hat a little bit more and family hat. I noticed -- and Yoanna, a great job on summarizing -- it was only survivors, and I wonder about the families of deceased ICU people to bring to bear because my experience in this area of working with people with advanced illness, one person has an advanced illness, the whole family has the advanced illness. Maybe I missed it.

Page 84

DR. SKROBIK: We had both, but we didn't declare it up front.

DR. COURSIN: Dr. Shehabi?

DR. SHEHABI: Yahya Shehabi. I think in terms of the recommendations, I think it's important to link that to the outcome looked at and make a recommendation. I think when it comes to the sedation group -- and Pratik, you could speak to that -- I think we made the bar very high in terms of the outcome.

Instead of saying we're going to accept a short and mechanical ventilation to make a sound recommendation for X versus Y, which now it has to reduce 90-day mortality, or it has to do reduce such and such and such at 6 months. I think that probably what led to a lot of recommendations being made conditional, a low recommendation, because there's just simply no data on that.

DR. DEVLIN: Yes, and that's a really important concern. That's one of the things obviously grade requires, is we did vote on our highest priority outcomes for our PICO question,

Page 85

and there was some pushback as people get familiar with the data, and later on as we're trying to make recommendations, "Well, can we just change some of these?" And we were basically, no, we can't, but that's such an important point, because the data, it did drive a lot of lower-level quality recommendations.

DR. COURSIN: Pam?

DR. FLOOD: I'll add on to what David said. I want to recognize --

MALE VOICE: Could you speak into the mic, Pam?

DR. FLOOD: Sure.

DR. COURSIN: That was Pam Flood.

DR. FLOOD: I just want to add on to David's comment -- Pamela Flood, Stanford -- that we are absolutely -- particularly those in this room who are ICU survivors -- not a representative sample. Not only have we survived, we survived intact, and we were relatively healthy academic physicians before all of this happened. So what's important to us and our families might not be important to

Page 86

everyone, and other things might have greater importance.

DR. EGEROD: This is Ingrid Egerod. I'm a qualitative researcher, and I have some concerns about the way the patient representatives are used in research because we discussed this at length about whether it's representative when we do our qualitative research. I have a feeling that that whole layer of discussion has disappeared when we have patient representatives, and patients suddenly become a representative to a much larger degree than they really should.

In qualitative research, we always put in all these important discussions of can this be generalized and so on and so forth. I think that's another thing that we need in research, is to really define or discuss how to use patients and families so it makes sense.

DR. COURSIN: If I could, Dale Needham, with your expertise, would you comment for us on that?

DR. NEEDHAM: This is Dale Needham from Johns Hopkins. I do think, and I did perhaps say

Page 87

to John and Yoanna, sure we have some patient representation, but it's pretty tiny. I think that our knowledge in how to do this is evolving. When I present in an hour or so, I'll talk about one approach that we had to try to have in a formal consensus methodologies and Delphi, and try to have about a quarter of our representatives be patients or families.

I think we've still got lots of ways to learn and how to do that, and I was sharing some of that with John and Yoanna as well, but I'm not sure that we know the answer yet. But I think it's important that we continue to bring this up to ourselves, continue to think about how we should do that and recognize that often patients are going to talk about an experience of one person, and we need to put that in context, too.

I've seen sometimes where I think we just give too much weight to what might be an outlier or one representation. I say that just because the research that I've done for 15 years looks at long-term outcomes of ICU survivors. So we've done

Page 88

assessments on thousands, thousands of assessments, so I think I have a little bit of a feel. Of course, my bias as well, but sometimes one voice may not always be the representative.

So that's a couple of thoughts, and I think when I present, I'm happy to share some of our learnings and thoughts that have come out of that.

DR. COURSIN: I want to get Leanne Aitken's thoughts as well. I'm sorry. I'm trying to get both of you.

DR. AITKEN: That's okay. Leanne Aitken. Some of you may not be aware, but within the UK research funding environment, you basically won't get any government funding without a reasonable patient and public involvement process, and that includes some sort of consultation with PPI, as well as PPI members as co-applicants on the grant with you.

In the current study that I'm a co-app on --

DR. SKROBIK: Sorry. What's PPI?

DR. AITKEN: Sorry. Patient and public

Page 89

involvement.
DR. SKROBIK: Not protein, pumpkin --
(Laughter.)
DR. AITKEN: No. And there's another PPI related to insurance. It's not that either. In the current study that I'm a co-op on that's comparing dexmedetomidine versus clonidine versus usual sedation, through our PPI process, one of the outcomes that was considered most important to a group of about 20 was how well the patient could communicate with the family member.
We would never have thought of that. We would never have put emphasis on it. That's a single example, but I think it's an important example of how we do need to think differently and make sure that we get that voice. Now, we have a group of between 18 and 20 that we consult with regularly, and we have two patients and public on our co-op team, but that's the process throughout the whole of UK government-funded research.
DR. SKROBIK: If I could just speak to -- this is Yoanna Skrobik -- the reproducibility

Page 90

within this small effort, what struck me was the cohesion between the 5 patient representatives that we had, who the exception of their preference of depth of sedation spoke with one voice despite their different experiences. I know it's not the thousands that Dale refers to and not the long-term outcomes, but within these specific topics, I was struck by the homogeneity.
DR. COURSIN: Ingrid, you had a comment?
DR. EGEROD: Yes. I just wanted to add to that, that I think one of the problems is that we're trying to generalize and maybe we should just accept that we can't generalize and that that's okay.
What we're doing is we're giving a lot of good examples of what might be meaningful to patients, but we're in a different paradigm, and I think we really need to keep remembering that it's all right that they're not completely representative. People are different.
DR. COURSIN: Mervyn, you had a comment?
DR. MAZE: Yes, and this is not meant to get

Page 91

into wordsmithing, but I want to ask about the way that a recommendation is framed either in the negative or the positive. And I'll give an example of what you've said, but I won't read word for word. You say something must not be used in all patients. How does that differ from it is useful in some patients and start defining what that some is?
DR. SKROBIK: I think that's an excellent point, and on the last slide that I showed, the subgroup comments spoke to that, the gaps being if people are all different in terms of pathology and in terms of how they respond to whatever intervention, being pharmacological or not, how do you tailor?
You're not looking at a cohort with an average when you're looking at the patient in front of you, and if that subgroup hasn't been studied or that personality profile hasn't been studied -- the Israelis published a beautiful study looking at whether being a controlling person made you more likely to develop delirium. I have to say it was

Page 92

one of my favorite papers.
So how do you make that transition? And I think that's where the -- subgroups specifically, because we made recommendations for pain management in the previous guidelines based on two studies in two very specific subpopulations, and said opiates were all the same because 2 opiates were compared in each of those studies; so an imperfect example of how you make a recommendation.
DR. MAZE: But are you therefore saying that unless the subgroup is not identified, that it's better to frame it in the negative?
DR. SKROBIK: No. I think you should think about it and just express it clearly. This is a very personal opinion, not the SCCM; this is a very personal opinion. I think you get so caught up in the naming of the conditional strong, weak, blah, blah. If you have the patients to not read the summary but read through the content of what created that recommendation, then you get an idea of what you're talking. Based on these two groups of this profile of patients, these were the

Page 93

results.
You can then take that away and apply it in your patient or not. But we live in a world where medical information triples every 10 years. So you take the guideline and you take the summary because you couldn't possibly be an expert in sedative exposure, and mechanical ventilation, and -- you couldn't. So you take it, and the problem is in the summarizing of it and in the words that you use in the summary.
DR. DEVLIN: Again, another thing with grade is you're forced to make your conditional recommendation for or conditional recommendation against, or strong for or strong against. That's where you're parsing this divide of risk versus benefit and all the other factors that came into the recommendation space. So we have some that look like they're negative and then some that are positive, and that's simply how --
DR. SKROBIK: But if it could have consensus, it's artificial.
DR. DEVLIN: We had a comment in the back.

Page 94

I'm sorry.
DR. TANG: Hi. I'm Wing Yu Tang, and I have much more experience on more of the real-world data on research side. So I'm very curious per a lot of the comments being made about generalizability, and certainly sample sizes I think comes into play as well, in terms of generalizing either subpopulations or small populations to a much wider and generalizable audience.
We also talked about things like productivity and absenteeism, which we capture a lot as well, that are more I would say real-world outcomes that aren't necessarily typical clinical trial endpoints. These are actually realities that I would say are limitations that sometimes clinical trials can have.
So I'm really interested in thoughts about -- we talked a little bit about qualitative research, but there's obviously a really growing number of real-world data, phase 4 studies, other kinds of prospective works which are targeting much more larger sample sizes, and the maturity of it

Page 95

clearly is evolving quite rapidly. And I'm wondering thoughts about where that falls into the current discussion on generalizability.
DR. COURSIN: Comments on that? Claudia?
DR. SPIES: Maybe surely. I'm also heading the whole medical society development in Germany for the guideline development, and there's a lot of discussion also with the Guidelines International Network on the quality of the guidelines. I learned a lot from our guideline from this networking that AWMF is having 180 medical societies included. The Guidelines International Network is a huge society giving standards and including all the stakeholder representatives to qualify a guideline.
At least from my perspective and doing a lot of guideline research a lot of times, I'm mainly stuck in the methodology. I think it's very important that we have these people who help us really to qualify our guidelines and really to get that implemented because only with them is it possible to get that implemented.

Page 96

My question is, is that moderated, all the guideline development here in the U.S., by Guidelines International or by a U.S. specific guideline network that's really having all of these people involved. One has to look at the throughput model at the end, or the patients, the relatives, the organizations, the system. It's very context specific before an intervention goes. And because it's not so easy to understand, I think it's important I think to use that help people have in different other guideline developments.
DR. DEVLIN: Yes, I can speak to maybe just a little bit of that. There's a great working group that obviously postulates and promotes the ways clinical practice guidelines should be done. Cochrane is involved as well. But there could be inherent biases from those, those organizations.
Currently, from what I've seen talking to other critical care organizations, is it's a little bit fractured in terms of the societies that support these guidelines have their agendas for doing them. I don't think there's a lot of

Page 97

cross-talk. Even within SCCM, I'll be honest, there's varying level of methodological support and the focus of how these guidelines are done. So there's just an incredible amount of variability in quality and how they're done. These are really, really big issues from that practice guideline thing.

The one other comment I wanted to make, which I think goes back to the comment there is, when we're looking at choice of sedation, this came up a lot within our group is, we're focusing a question on which sedative the patient is to get in the ICU, but that patient stay obviously could be quite dynamic throughout the ICU stay, and maybe there's a choice of sedative that's better on day 3 than the first day they get intubated if they even need the sedative, and that dynamic process is not brought into the guidelines at all.

We do bring it up as a gap, but I think it's a really important one for this group because most of the studies, you randomize patients to one sedative or the other, and you keep that sedative

Page 98

going unless there's an adverse event, or a safety concern, or they're extubated.

DR. COURSIN: Thanks, John.

As the moderator, we're coming kind of toward the end of this session, and a question comes up, a very logical one, about longitudinal database follow-up.

Frank, any comments on that as far as creating these databases and looking at them over time, particularly with evolving practices or competing guidelines?

DR. DEXTER: Frank Dexter, Iowa. I understand longitudinal databases for endpoints such as work or something like that, but it's really hard, if it's difficult to measure something even in a randomized clinical trial, to begin to think about longitudinal measurements. I kind of find that to be very difficult. Even if you were to say take databases that already exist currently, if you can't in a randomized trial measure something reliably, having more data isn't going to make it reliable.

Page 99

DR. COURSIN: Dan?

DR. SESSLER: Data quality also tends to be poor in registries and controlling for confounding is challenging. That's not to say that registries are useless. We do lots and lots of registry studies, but they sure don't have the reliability of a controlled trial.

I think the point that even controlled trials in this environment are difficult is valid and important, and a solution to that is not a registry.

DR. COURSIN: Thank you. I'd like to keep going on with that, but I have one final kind of burning question I'd selfishly like to ask. I'm going to direct this to Steve.

Steve, you outlined what are the problems; where are we unhappy with things; who's going to pay for this? It would seem to me the key action items coming out of this meeting would center around clearly identifying what we need, whether we can fill the gaps in or not, but what do we need and who the hell's going to pay for it, and who's

Page 100

the advocacy group?

The force I would try to get out there, the last data I looked at 1.4 percent of the gross national product is spent on critical care in the United States. That's a lot of dough. Is there a way to leverage what we're talking about here in a manner that we could make effective collection of data, analysis to that data, and implementation?

Steve?

DR. SHAFER: I'm looking up to the session that I'll be moderating at 4:30, and that's the same question I've had, which is when we're talking about clinical trial designs, you can't really talk about that in the assumption that there's unlimited funding. And a clinical trial design has got to identify a problem worth solving, and the worth, I hate to say this, has got to be defined in dollars or whatever the currency is, but it's got to be defined.

DR. COURSIN: We'll vote on Brexit later today.

(Laughter.)

Page 101

DR. SHAFER: Yeah.
DR. SKROBIK: Dollars and glory, no? Is it not dollars and glory?
DR. SHAFER: Somebody's going to write a check, and they aren't going to write a check for your glory.
DR. SKROBIK: No, write a check so that your finding -- so there's the academic and journal driven study; wow, we have this new thing. Nobody does granular metrics.
DR. SHAFER: That will motivate all of us to put in our time and effort, the glory part, because we don't really do it for money; we do it for the contribution we make. But in terms of funding the cost of doing a study, if we can identify -- here's the cost of not knowing. The cost of not knowing is X, and it's going to cost some number smaller than X to fill in that gap and give you this return.
So I think we have to identify the costs in that -- and as you say, 4 percent is a big number, but what is the cost of not knowing -- what are the

Page 102

costs of these gaps in knowledge and gaps in practice?
DR. COURSIN: Pam?
DR. FLOOD: I was just going to add to that, that in a way, it's fortuitous, even though it's sort of horrible. But at least in the ICU, there's a big pot of money. We're spending an enormous amount of money in ICU, end of life, and hopefully not quite end-of-life care. The concept that this could be done more efficiently and better means that there actually is money to be saved there.
DR. COURSIN: And again, I'm looking for how do you garner the issues and get an advocacy group. And the real people we're looking at who seek glory that could make this happen is the political force. And what does this group see as a way to pool the data together and the conversations here and get that kind of information out there.
Well, there seem to be a few comments. We'll start here.
Rick?
DR. RIKER: Riker. I think one of the

Page 103

things, Steve, in response to your comment, as clinicians, I would be interested in a study -- we see this resource costs versus outcome benefit or whatever. There are four quadrants. So I'd be interested in a new approach if it got as good results for less money or less resource utilization, or if for similar costs, I got better outcomes.
Those are the kinds of things we're looking at to make decisions about what do we do with our patients. If there's not a new drug that comes down the pike that the pharmaceutical industry is going to pay for, for the research within our own societies, or our governments, or AHRQ, whatever we use for pathway, we've got to look for those kind of outcomes.
DR. COURSIN: David Brown?
DR. BROWN: I think one of the big challenges for all of us is $550 million a year is spent in lobbying in this city on healthcare.
DR. SKROBIK: How much?
DR. BROWN: $550 million is spent on

Page 104

lobbying in healthcare. Almost all of it is spent to keep the system the same because there are so many people making so much money inside the system. So that very idea that Rick brings up, is that generating the political will to make a change, you're stepping on so many toes.
So I think, Doug, you're exactly right. You've got to use political maneuvering, and you've got to have a face to it. There's a group here that our firm's a member of, C-TAC, Coalition for the Transformation of Advanced Illness Care. Dave Longnecker is their chief of strategy. It's focused on better end-of-life care, and there may be something for some of the critical care groups to have a little larger role in that group. It's a nonprofit spun out of AAMC.
DR. COURSIN: Dr. Shehabi?
DR. SHEHABI: The Australian Ministry of Health and the Medical Research Council, which provided a lot of funds for the ANZICS Clinical Trial Network investigating an ICU. I've recently looked into the return on its investment into the

Page 105

clinical trials by ANZICS clinical trial group, and the biggest return was learning what not to do in ICU.

The point that Yoanna made -- I think you, Doug, made that point first, that a lot of what we do in ICU came to us from outside ICU and wasn't actually designed for ICU. So it's really important to examine what is it that we're doing and what is it that we need not to do because that's where the real saving is.

DR. WARD: Thank you.

DR. TANG: Just for the record, I think it's important that was noted before about health economics being an important player and balancing that conversation so that you can translate appropriately clinical outcomes to what it means it health economics, I think that's definitely an arena.

The reason I brought up earlier about the ideas of real-world databases and registries is not to say that it is in any way going to replace or even the supplementary, but it's offering more data

Page 106

points to consider when we're talking about actually seeing how these patients are flowing through, and if we're not capturing them, how we can better capture them, and the idea of creating that more as a baseline of how we look at things rather than pointing at all of the concerns.

The idea is that we really want to make sure we have as wide as data possible. And if we have already these infrastructures that speak to insurance with claims and we have EHR databases, can we look at those as different avenues that are less resource intensive to add to more of that data information and system forum?

DR. COURSIN: Thank you. We're at the little after 10 o'clock mark. I didn't want us to fall behind. I wanted to thank Yoanna and John for their expertise. I had told Denham I would do my best to get this shy group going, and I appreciate your coming through for us.

DR. WARD: We do try to have fairly generous break times. A lot of the discussion takes place not in the formal meeting setting but over a cup of

Page 107

coffee. So I hope we can continue these discussions over the next half hour and then come back at 10:30, where we will continue with the panel discussions. Thank you, and thank you, panel.

(Whereupon, at 10:02 a.m., a recess was taken.)

DR. WARD: After that nice background discussion of where we are and where are the gaps both in evidence and methodology, I really wanted to start back to the patient's perspective. I think that's one of the pieces, when you talk about patient centered, that can drive a lot of parts of understanding what we need to do to fill the evidence gaps and what the right methodology is. There was a lot of great comment of incorporating patients in clinical trials, particularly using qualitative research methods to understand the patient's perspective.

We've got two speakers and a panel before lunch, and I would like to start out with Dave Brown.

Page 108

Presentation - David Brown

DR. BROWN: Thank you, Denham.

We're going to be talking about patients and families in the ICU, and our previous discussion was on a population you're in now, which is going to be one patient and one family, and I'm going to tell you my story.

I'm an outlier. If you see me weaving and bobbing a little bit up here, it's because my critical illness neuropathy knocked off some of my proprioceptive function, and so that I can keep my feet on the ground, I move a little bit side to side. It's especially bad if they turn the lights out. It's an amazing process that we almost never talk about. It's not so much what to do, but what to avoid and side effects.

Cures. Diseases come with a lot of side effects. You can see that I'm an old chair. I'm a former chair. I led a lot. There's actually faculty in the room here that I was blessed to work alongside and others that I've been blessed to work alongside nationally. But I now run a little firm.

Page 109

What we're going to do today is talk about patients, and more importantly families, from what I learned, and then I'm going to share with you some heartfelt lessons that I think I take away from this.

Now, my declaration of interest, I'm the CEO of a firm called Curadux. I wanted to call that from Doc and the Family, but a Japanese guy owned the domain name, and I couldn't buy it from him, so I Latinized the English care guide, and that's what Curadux is. I'm a board member of NeuroTherapia, which is a unique molecule, a cannabinoid 2 compound that we've worked on for 15 years that just had IND approval, and we hope to get in phase 1 trials by June for Alzheimer's. It's a very unique drug.

I'm an academic medical insider. At the time of my illness, I was one of the directors of the American Board of Anesthesiology. I sat on the ACGME's executive committee. I ran the RRC for many years. My clinical background is pain medicine and anesthesiology. So that's who I am.

Page 110

Here is a graphical cartoon of where I was. I started out first in the Air Force; then went to Virginia Mason; then to Mayo, where I ran their quality program; and then to Iowa, where Frank Dexter and I worked together; then to MD Anderson Cancer Center; and then to Cleveland Clinic, which is where I had my most notable certificate, and that was my critical illness and multiorgan failure. Then I went on after that to do two years of graduate work at Loyola in Chicago in bioethics and health policy in preparation for what I'm doing now.

Now, I'm going to sit, and the star of this about 4-and-a-half-minute video is actually my wife.

(Video played and transcribed.)

"DR. BROWN: The story goes back 35 years. I was active duty military in the Air Force. We were on the southwest side of town, and we got gunshot and knife wounds, frequently transfusing many, many units of blood. At one of those points, it's theorized that I had a needle stick, and back

Page 111

then we didn't know the term hepatitis C.

"The FDA approved a new treatment regimen that gave me a larger opportunity to be cured if I went through the chemotherapy.

"KAREN: The first day it was fine, and then looking at the number of pills he took every day was shocking, and they affected him emotionally, physically. He couldn't get up the stairs anymore without crawling. It was very, very difficult.

"DR. BROWN: Then as we got about 5 and a half months into the 7-month course, I developed sepsis and required admission to our surgical intensive care unit.

"KAREN: I remember the first night he was in the ICU. It was like being on Mars. They don't speak my language. The noises are very strange. I remember asking what am I supposed to do?

"DR. BROWN: Time [inaudible}. I spent 3 and a half weeks unconscious, had a heart rate somewhere north of 140, and I [inaudible] had pancreatitis, and my liver took a vacation. Of course, there were concerns that I bled into my

Page 112

head. Then I had respiratory failure. I had more than 6 organ systems out, so that predicts somewhat around 100 percent mortality.

"My family had a couple different discussions about end-of-life care and whether to do not resuscitate or should [inaudible]. I found it very interesting in a very detached professional way to watch myself dying.

"KAREN: I remember just thinking, just tell me what are the odds here? The wonderful nurse asked if I wanted to lay down with him, and I thought it was the last time I'd ever lay next to him. My decisions were based on more than medical knowledge. It was based on the hopes I have, and the prayers I have, and the faith that I have, realizing I didn't know the outcome.

"DR. BROWN: Then, late on a Friday night in one heartbeat, I became myself. I became cognitively intact. I became myself just as fast as you can snap your fingers. The first phrase I typed out was, 'I've never been more alive.' People walked with my family. That was very

Page 113

important.
"KAREN: And your family is very important. Every person is different with different challenges. So I would say equip yourself before you meet a circumstance that will require tremendous fortitude and faith.
"DR. BROWN: Caring for the family, letting them know that we're going to protect their loved one, we're going to do our very best, was keenly important. To know that the family of that individual is a human being with God-given dignity that that human being has [inaudible], sometimes a patient will tell me what they're worried about, and I can put a hand on their shoulder and say, 'A year ago, I was where you were.
"There's really nothing so beneficial and almost a sacred commitment that we have with our patients, to respect them and try to relate their pain. I always considered myself very empathetic, and I thought I was, and I probably was. But this illness has raised my degree of patient-focused empathy to another level. If I followed the book

Page 114

on algorithms, I'd probably be dead."
DR. BROWN: This was used by the Cleveland Clinic in part of their empathy series. Just some framing for you out of the patient experience. I'm in my own ICU, so everybody in there worked for me at the time. I have a daughter that's a doc, a son that graduated from Georgetown Law. My wife's really smart. I had all the advanced directives done. Our money was in a trust. I had planned ahead. Everything was planned ahead, a real connected medical insider, and our family struggled.
So I'll tell you a little more about me. This is my hepatologist note the day I was admitted. You see in the upper left, blood pressure's 78 over 45 in his office in a wheelchair, and that's the night before I knew I was failing, so one of my buddies came over from an outlying hospital near where I lived on the west side of Cleveland.
Started an IV, hung the first bag of lactate, and then I hung the next 2 bags of lactate

Page 115

to give me a little boost till morning. My secretary sent a car for me, and I laid down in the back until I got to the hospital. Then I was sent to the emergency department even though I was on my way to the ICU, because that's what we did at the Cleveland Clinic. You couldn't be admitted directly from any office to the ICU. Is that patient centered, 6 hours in the ED?
This is a week later. We don't know quite what's wrong with me yet. This was the day they figured it out. I had vasopressors. My heart rate was north of 140. My respiratory rate was about 40. And I also looked as if was developing a viable proliferative disorder with a node biopsy positive for a lymphoma. Oh, and by the way, my EBV titer was greater than 5 million per cubic millimeter, so I had developed from my immunosuppressed state an EBV. We don't know if it was just simple EBV sepsis or EBV hemophagocytic syndrome. That was a weekend.
So overall, you've heard a little bit from the video. I had hepatitis C for 35 years. I

Page 116

picked it up in the military and had done pretty well. And I thought, my family, probably I'll die of heart disease before hep C will get me, so I kind of avoided it because I'm genotype 1, which is what most of us in the U.S. are, and the cure rate was about 25 or 30 percent with the old chemo.
Well, the interferon, ribavirin, and a protease inhibitor moved that up to about 80 percent. Let me tell you, if you ever get offered interferon, don't take it.
(Laughter.)
DR. BROWN: Find another way. But I had the DIC for 4 days. My platelets were about 18,000. That's why they gave me contrast through the portable CT to see if I'd bled in my head, and they knocked my kidneys off at that time with hypertension. I was delusional through much of this. I can tell you, for part of the time, I was in the ICU, delusional unconscious, for 3 and a half weeks. I was in a European ambassador's place with China plates on the wall. I've never been able to find those China plates in that ICU room.

Page 117

Also, I will speak to you in a moment about a piece I wrote for Anesthesiology in the Mind to Mind section, and I'll tell more of that story. But the most suffering I had in the ICU was every time an alarm went off in my room, it signified to me and my delusion that one more patient was entering an unethical research trial I was leading. It's not an exciting run. That actually caused more suffering than any of the physical things I had.

Intubated and ventilated, I had ARDS. I had a 50-pound weight loss, and that will become important in just a moment. My marrow wasn't working. My albumin was 1. I had what looked like a term belly that had to have ascites drained. Having that much ascites with ADRS and a 50-pound weight loss is an exciting run.

So here's what I think I learned out of this. When you lose 50 pounds out of your core, every time these wonderful nurses turned me to clean me, because I was incontinent of anything, my shoulders and my hips subluxed. And I came up to

Page 118

my critical care docs and nurses afterwards and I said, "Do you guys know about that?" They said, "No. Nobody ever lives that is in that setting."

So that was painful. In this morning's discussion, I am betting that it looked like that was agitation when it was actually really severe pain on shoulders and hips. I had an 8 o. tube cut at 22 centimeters, and it's like breathing through too small a straw. If you think of an anesthesiologist intubated with ARDS, that's a bad setting because you understand it.

Then after the nurses and respiratory therapists come and suck your tube out, they suck your FRC right out into that tube. Now, you have to cough back up your FRC, but you don't have the core muscle strength to cough, so you feel like you're suffocating, and it hurts, and you're short of breath.

Nasojejunal tubes, they hurt. They're sewn in so they don't pull out. Fluid overload between dialysis. When the young nephrology fellow comes by and says, "Oh, your numbers look pretty good

Page 119

today; I don't think we'll dialyze you," "I just wonder why I'm short of breath?" They were going to treat my numbers rather than treat the patient.

I can tell you waking up -- I'm still in the ICU 7 years ago today. I was extubated 3 days ago. I'm in the ICU. I wake up in March Madness. There's not much better time to wake up than watching basketball from a warm bed in dialysis where they take 2 or 2 and a half liters off of you. You breathe better. They feed your graham crackers and orange juice and put warm blankets around you, and they're very kind individuals in dialysis.

Three weeks into my ICU admission, 3 and a half weeks about, I have a dream. I don't think it's a delusion because I'm an old pilot, and I flew right up until my chemotherapy. I flew the airplane you see in the lower right all over the country. I rarely flew commercially. I flew myself. But I'd had a dream over the previous decade, not flying, but they're in trouble.

I walk up, take the right seat, pull back on

Page 120

the stick, avoid the high line wires, squeak the wheels onto the runway. I have this dream, and as you can see there, it was actually 2:40 when my eyes opened up, and I'd been out of it for 3-plus weeks. I became myself. I became this guy, very weak, but cognitively I thought I was intact.

Here's what was going on, on that day. I had no sedation at the time I woke up. I was not on sedation at the time I woke up. But during my ICU admit, they used propofol or dex. They used some benzos early on. And that comment I made in the video, I remembered to this day, I woke up. My nurse was a former army medic from Iraq war and went to nursing school, barb wire, weightlifter; couldn't scare him at all.

Jose came into my room. His eyes got about this big, and I got the alphabet board out, and my little old tremulous fingers typed out, "I've never been more alive. Call my wife." And then it really got exciting because I became -- I talked about lobbyists here in DC. I became the principal lobbyist at the Cleveland Clinic for extubation of

Page 121

one single individual.
(Laughter.)
DR. BROWN: I had had such an up and down course, they didn't quite trust me. "I think we'll just -- we don't want to do it. Let's let him prove himself."
I wrote this article. It's in the October 13 issue of Anesthesiology, Fantastic Delusions, Futility, and a Family's Love. My wife is an English major in her first life, and she tells me, I wrote in the genre of stream of consciousness, and it goes through all the kinds of things I was thinking. It tells a little bit of the story.
My son, who is an attorney, his sister, his older sister, who's a physician, told him the first night I was admitted, bad things happen in hospitals at night. He never left my room at night. He caught 3 drug swaps that probably would have hurt me over those 3 and a half weeks. But he asks me, "What do you want us to do?" And I said, "Let's do the next thing."
When they could rouse me out of the

Page 122

delusions, I had mostly a detached clinical look at my condition. My wife went to the family consult room on two occasions where they talked about end of life. My critical care team thought I was suffering too much, and certainly it's gone on, and on, and on. Should we just provide comfort care? And my wife actually said, "Let's give him a couple more days. He's always been a fighter." Well, I'm very thankful that she did.
This is the story of -- that's not a delusion. That's my recurring dream of saving the aircraft. I have to tell you, during my illness, my auditory function was maintained pretty well. I have lots of intermittent memories of hearing even though my face was masked with unconsciousness or delusions. We know in anesthesia we always say don't talk about patients even when they're anesthetized because auditory function often is maintained. Well, I can tell you, even in critical ill patients, auditory function is maintained even though I couldn't speak into the world.
Being extubated, I still can feel the

Page 123

emotion of being extubated. I was an old hurdler in college, and I actually used to be pretty fast before the neuropathy. But I still remember occasionally they'd make me run 400 meters because they needed somebody for the relay, and I had fast-twitch muscles, I didn't have slow-twitch muscles, and I needed a shorter distance. And I'd finish a 400 meter and just be gasping, and then it felt so good to stop. Being extubated is the same thing. It feels so good because they pull that tube out.
Hippocrates had it right. I could sit down right now. I won't, but I could. It's more important to know what sort of person has the disease than to know what sort of disease the person has. If I have a criticism of modern healthcare -- and I have many as a patient. I get care at the VA hospital. I get care at Mayo. I get care at Cleveland Clinic. I get care at Marshfield Clinic. So I sample lots of different centers.
If you look at me in my medical record, I

Page 124

look like a walking dead man because they keep every comorbidity in there so they can up-code when the reimbursement structure allows them to. So I've had to create my own personal health story to share with physicians when I have to meet with them so that they find out I still get 17,000 steps a day, and my dog and I go hiking quite a bit, and I actually am pretty active.
Carl Hug, many of you know Carl. Carl sent me an email. This happened in March of '12 for perspective. This was in October of '12, we were getting ready to go to the American board. And you can see what Carl said. He said, "You screwed up my ICU lectures." He taught ethics down at Emory. He said, "I always thought after 7 days, we ought to just provide comfort measure, and you taught me something. I look forward to seeing you."
Roger Williams was a PhD biochemist at the University of Texas, and I think he hits it pretty right when you're thinking about me as the patient. Medicines for real people, statistical humans, are of little interest. What happens is we let our

Page 125

policy discussions get too close to individuals' bedsides.

I'm going to tell you now about my family. I have 4400 pages of ICU records from the Cleveland Clinic. A little secret, I'm in my own ICU. I'm everybody's boss. When I was most critically ill, 10 days my vital signs were identical in the record. Do you suppose somebody was cutting and pasting? Do you suppose?

(Laughter.)

DR. BROWN: I can't validate it. My daughter, she and I have always been fairly close. I used to coach her in long jumping. She's a pediatrician and did special needs pediatrics at MCW. Here's what she posted. She's the Facebook generation. "Nurses are angels sent from God. Certain doctors can be angels when they listen to you and actually come examine your family member rather than making decisions from outside the room." If I reflect on what's most missing in intensive care units, it's time; what's missing in all of medicine, but intensive care, it gets

Page 126

amplified because of the critical nature of it.

Text message. My children saved all their text messages. I have 20 hours of recordings at my bedside that my son used his iPhone. He worked on Capitol Hill for six years, so he's all attuned to that kind of stuff. My son, Cody, daughter, Sarah, they're going back and forth. "He's going for a pancreas ultrasound. Is this all pancreatitis? No."

Ten days. "It doesn't look like Pops has much more time left; supraclavicular node." She concludes, "I think he knows his lymphoma won't be treatable. I'm trying to do what dad taught me to do for a patient rather than do to a patient." And this old father doesn't even remember that I taught her that, but I take credit where credit's due.

Now, my daughter is a pediatrician. What do pediatricians do? They worry about kids. I had two grandsons at that time. I now have four all out of this family. She wanted to bring the boys to see me before I died. So that you can see the punch line is she never been happier for following

Page 127

her gut and kind of forcing the children into my room to say goodbye.

The point I'd make here in this group is if a family member has an advanced illness, everybody in the family has an advanced illness. The ICUs often are exclusionary of -- because we're busy. I mean, we're that way. But it really was important to her.

My daughter also finally said, "Can we get one doc to be the quarterback, at least in communicating? Because everybody's telling us slightly" -- it's that telephone game. You whisper to one person, and what comes back around. Family really wants to have a single individual to share the message.

Later in the course, I'm still uncertain if I'm going to live at this point. I'm really sick, and my daughter says, "I can list by name the ones who've made our stay comfortable." So if you think families don't know what it means to them, they do know. It's a very personal journey.

Agitation, sleep, and PADIS guidelines, my

Page 128

son slept in the recliner in the corner of my IC room for almost 3 and a half weeks. He says from my nonmedical perch, he needs some peace and quiet. Every time they leave him alone, his numbers get better. I can tell you, after I recovered, I was back at work in 60 days. I'd actually showed up the first 3 weeks after my illness.

I drove to the Dairy Queen 2 weeks after I was discharged. That was the test, 2 miles from our house. I was on weight gain diet because the dietitian told me eat a thousand extra calories a day. And I can tell you, if you ever get in the position I was in, I was on an apple fritter times 2 daily diet, a thousand extra calories. The only hard part of that diet is stopping.

(Laughter.)

DR. BROWN: Even now when I walk by the bakery and I see a fritter lying unattended, I struggle.

They would text ventilator settings back and forth. My daughter had to go home to her practice and her family. This was on the day that I

Page 129

awakened. He says, "Dad's communicating; told the ICU doc, 'I won't fail. Extubate me.'" And I have to tell you, I had the little minimal strength to extubate myself. I refused to extubate myself because it was my unit and I wanted to set an example.

(Laughter.)

DR. BROWN: I think that's crazy, but that was inside thinking. Since this is an ICU sedation conference, my wife would send emails blast every evening. She's an introvert actually. So she would go home and sleep. My son would stay. Daughter would go. She'd send an email blast, and on this one on the 24th, the night I woke up it says, "I think he bothered his doctors so much, they put him on sedation for their sake, not for me." And I would bet you she had that parsed just about right.

So let me tell you now lessons learned, and I'll get off the stage. Algorithms aren't always expert. Physician judgment, two physicians in particular probably saved my life. Many

Page 130

contributed to saving it, but two in particular were not going to give up. And it's not that they beat me, but they saw further ahead than some of the busy docs did during this.

I've told you before, utilitarianism, we all let it creep next to the bedside. We all let it creep next to the bedside. Everybody needs an advocate. Often you need more than one advocate. It may be your nurse that night, it may be your family member, it may be a physician the next night, but you all need an advocate.

These thoughts are what I came up with my first week home after coming. "Care for me as a unique human being. Look me in the eye." And I say that mainly, the 7 months that I went through chemo with interferon and my hemoglobin was about 6 and a half or 7. My white count was often down in the 200 range. So I was getting intermittent Neupogen and erythropoietin.

The most meaningful experience I had was one of the laboratory techs that drew my blood weekly. She'd gone through an osteogenic sarcoma when she

Page 131

was 18. She was now about 28, and Gina and I, she'd always take me. She wouldn't let anybody else take me, and that was probably one of the most meaningful experiences in that illness, was having that personal connection.

"Don't harm me if you can help it. Don't let pain overwhelm me. Provide me with cutting edge care yet appropriate for me, my values and goals, and my finances." We almost never talk about that, but these were my thoughts the week after I got home at night. And I logged it in just so I wouldn't forget and start to make it sound better than it was.

This isn't my slide. This is out of the Harvard Business Review. In December '14, Drs. Mate and Compton-Phillips, Dr. Mate lost a mother at MGH, falling through the cracks after a hip fracture, I believe was what the setting was. The problem in our healthcare system, it's fragmented. Somebody said earlier, I think Denham did, that we sometimes work in silos. And for those who haven't heard me share the story, one of

Page 132

my mates used to say, "Those are cylindrical centers of excellence, and we all our cylindrical of excellence is the dominant one."

Advanced illness forces families and individuals to become their own care managers. Our blog this week out of Curadux is actually on healthcare coaching. The paper, Dr. Beamer is a researcher at Mayo and just published a paper on Healthcare Coaching at the Mayo Clinic.

I practiced there for seven years. They used to pride themselves on navigating anybody through the system. There were general internists that were the coordinating docs. If Mayo is starting to coach individuals to be their own care managers, the system is just about tipped.

Powerful incentives, mainly revenue, top-line revenue dominates the unique values and goals and individuals. Futility and suffering rapidly increase. I can tell you that suffering is real in an ICU, but it's not all physical suffering. There's a lot of emotional suffering at that. But this is not my words. This is Mate and

Page 133

Compton-Phillips.
So with that, I thank you all for listening. And at this point, my illness was a gift, and I see it in that fashion.
(Applause.)
DR. WARD: David, we'll hold questions till we do the panels.
Now, I think we've had the patient perspective, and now I think Dale will talk about how do we measure that patient perspective. Anecdotally is important, stories are important, but the broader perspective of multiple stories are also important.
Dale?
Presentation - Dale Needham
DR. NEEDHAM: A very hard act to follow, so I'll try my best. This is a case study not directly applicable to sedation, but hopefully there's some generalizable concepts. The work that I'm going to talk about was funded by an NHLBI R24 grant. That's a grant mechanism to create research infrastructure rather than to do original research.

Page 134

With that grant, we're looking at creating outcome measures that should be used in evaluating the post-discharge outcomes of ICU survivors. We also were interested in how to retain these survivors in longitudinal research, so state-of-the-art cohort retention, and then statistical methods because these data are hard and complex, and Elizabeth Colantuoni will talk about some of that separately.
So what I'm going to focus on is one aspect of Aim 1, and I'm going to try to go through these points. I'm going to start with a scoping review to tell us the size of the nature of the problem that we're trying to address.
Within critical care, as you can see from the figure, there's a growing number of studies evaluating survivorship experience. You see the figure with the graphs going up. So we're very much interested in this, but out of the 425 papers that have been published on this topic, we're all measuring different outcomes.
Quality of life seems to be a pretty popular

Page 135

outcome; 65 percent of the 425 studies evaluated that, but only 6 percent of studies actually evaluated physical functioning through an in-person assessment of patients. So large variability. There's sort of no standardization in how we think about the survivorship experience.
This makes it very challenging for us to have comparable and consistent comparisons and representation of survivorship experience, and I think it really reflects that we don't know what's important, what our patient important outcomes. We're all sort of just measuring different things, but are there a core set of minimum things that we should always be measuring if we want to understand survivorship? So that's a key question.
The next question is how the heck are we going to measure these outcomes? Across 425 papers, there were 250 different measurement instruments. Within post traumatic stress disorder, for example, there are 70 papers that evaluated it. They use a whole host of different measures. This is like if we're going to measure

Page 136

temperature and somebody's going to use Celsius and somebody's going to use Fahrenheit and 10 other instruments with no crosswalks between them. We really can't bring the field forward if we're going to continue this.
Also, there's a big chance that important outcomes will simply get missed when we don't have a consistent minimum approach. It's difficult to compare results and to meta-analyze. These are our big issues, so the scoping review helps us understand the nature of this problem.
Now I'm going to talk about one approach to addressing these core sets and a number of subpoints, so a little bit of jargon. A core outcome -- and this isn't my idea. This is something that's happening across all fields of healthcare that lots of people are interested in what are called core outcome sets. A core outcome is a concept, health-related condition, or aspect of health that always must be measured within a field, so it's what you should measure.
A core outcome measures how we're going to

Page 137

measure it, so how should we measure it. So the what and the how, those are two different questions but related. Then the core outcome sets that minimum set of outcomes that all of us agree to measure as a minimum within a specific field of study, and a core outcome measurement sets the minimum collection of measurement instruments.

Importantly, this doesn't restrict investigators from measuring a hundred other things if you want. This is supposed to be a small, feasible minimum set that we all agree to do. Within critical care, this work that we did in our grant is not telling researchers that everybody has to measure patient outcomes after the ICU. Lots of people got sort of upset about this. This is just saying if you choose to do this, if this is relevant to you, would you consider measuring this minimum set of core outcomes with these measurement instruments? That's sort of where this is trying to address things.

To do this, we're going to need to understand a few things. We're going to need to

Page 138

understand, first of all, what are patient important outcomes, how we might measure them, and how we might make decisions. I want people in critical care to stop saying we don't understand patient important outcomes. I'm going to present you a series of arguments to tell you that I think we do.

I'm going to go through each of these points and start with a qualitative research study that our group did sampling patients from across the U.S. We wanted to understand this experience of acute respiratory failure survivorship. We had 48 survivors recruited from 35 hospitals. They're followed up around 9 months after follow-up, starting with open-ended questions, and then probing after the open-ended questions using the PROMIS framework, that I'll talk about a little bit later, to make sure that nothing seemed to have been overlooked.

These are some of the experiences that survivors reported to us in their qualitative work. A 34-year-old man a year later said, "I have the

Page 139

tendency to forget a little bit more. My brain's a little bit more scattered," so thinking about cognition; a 67-year-old male at 6 months talking about mood, "I'm a useless person and basically a parasite. I have this emptiness inside. You wonder why I should even wake up," then a 63-year-old woman at 6 months -- or 9 months, still having difficulty with swallowing and talking about how she needs to relearn how to swallow her food so she didn't choke. This is just a little bit of examples.

The key findings to synthesize in a slide is that patients' experiences seem to fall within these categories, having physical impairments, problems with mobility, pulmonary symptoms, stamina, having mental health symptoms that we thought fell into depression, anxiety, and concerns around getting sick again; and then social health, which really hasn't been looked at so much in the empirical literature, but changes in employment, and changes in being able to do your valued activities.

Page 140

That's the first bit. I'm going to go through each of the bullet points. The paper that I keep showing at the bottom of this slide synthesizes everything that I've just said in a single paper for people that are interested and don't want to look up all the individual papers. It's free, full text.

Our qualitative work was just a one person's study. We then did a systematic review of all the published qualitative studies that had been done. There are 21 studies at this time, and the key findings from these studies looked at to synthesize for physical function, mental health, and social health; and then also, some patients, as we've just heard, having a positive experience after their critical illness, describing gratitude, changes in outlook, this being a gift having been critically ill. So again, something that's not captured so much in empirical research.

To triangulate our one qualitative work with everybody else's qualitative work, there seemed to be consistent signals around patients having

Page 141

impairment across a lot of different domains, that some survivors had a positive impact, and that social health was important such as return to work but not often captured in quantitative studies.

Then we want to take a different angle to understand patient important experiences. We wanted to understand these measurement instruments that we commonly use, are they capturing what patients think are important experiences. Here we had a unique opportunity. We'd done that qualitative study, and those same ARDS survivors happened in a separate study to have had standard patient-reported outcome measurement instruments performed.

What we did is we independently looked at these qualitative findings and tried to characterize what some of the themes were from those, and then compared these patient-reported outcomes in those with and without symptoms that were self-described in the qualitative research.

Patients may have described something that sounded like mobility impairment when two

Page 142

independent people looked at the qualitative research. And we said for example, those that that qualitatively described a mobility impairment, how did their objective scores or their patient-reported scores look differently? We did the same with mental health and cognition.

For example, those patients that endorsed having problems with mobility had much worse scores when it came to two different measures of physical functioning, the SF-36 physical component Score, the EQ-5D mobility score. Those seemed to capture the patient's experience.

Again, patients that qualitatively described what independent people thought were anxiety or depressive symptoms also had worse scores on objective measures, HADS anxiety score, HADS depression score, and ISR score for PTSD. So those measurement instruments, again, seemed to have captured the patient experience. But then when it came to cognition, interestingly, patients endorsing memory impairment compared to those who didn't had virtually identical median scores across

Page 143

objective performance-based measures of cognition.

Common measures that are used and actually happened to be part of the core outcome measurement set seemed to reflect patients' experiences of mobility, anxiety, depression, and PTSD, but when patients are reporting cognition, they're reporting something different than what we pick up with objective cognitive testing.

So I think we need to think very carefully about that. Patients may have objective problems on cognitive testing, but not actually have any insight into that, and vice versa.

Now we're going to move and think about what clinicians perceive. Here we have two independent Delphi consensus projects. These were sort of test runs for our big Delphi at the end. Here we took an international audience in the United States. We had a hundred clinicians that responded to a poll, and 44 of them were able to come to an in-person meeting for a second round of voting. And after we finished that, this exact same Delphi project was completed in Australia, mainly with PTs, but the

Page 144

same project, a completely different audience, at a different time, but done in the same way, so we could look at comparisons between two sites.

What we have offered up to these clinicians were what we could think of as 19 different domains. We've spent a lot of time thinking about what are going to be outcome domains that might be relevant for the patient experience during ICU survivorship, and we used lots of different ways to populate those 19 domains.

We used this NIH-PROMIS framework, which is a comprehensive measurement framework for patient-reported outcomes. It's a whole system that the NIH has had millions and millions of dollars into. You can see some of the domains that they talk about there and then here. So we used that to populate these 19 domains.

We also used the SCCM's post-intensive care syndrome framework to, again, get more things for those domains that clinicians voted on, and then we used the WHO's ICF. So we had lots of different ways that we triangulated to get 19 different

Page 145

domains to ask about.

This compares the American-based Delphi work with the Australian based. Importantly, there were signals across two different continents, two different populations, that these clinicians' perceptions were that research studies should always be measuring survival, physical function, cognition, and health-related quality of life. We ask patients in a whole bunch of different ways. We now ask clinicians in a whole bunch of different ways to figure out these core outcomes.

Then finally, I'm going to talk to you about a survey that we did. We had 279 respondents. We had about 80 survivors from across the United States of ARDS and acute respiratory failure. We had 80 family members and 55 pairs of patients and families from across the U.S., and then we had 121 clinical researchers in this field from around the world, predominantly from Europe, some from North America, and Australia.

We all asked them the same question. We gave them those exact same 19 domains, and we asked

Page 146

them, should these be measured as a minimum measurement in every single ICU survivorship study? Interestingly, the patients and family members thought 18 of the 19 outcomes are really important and should always be measured, so these have a lot of face validity with them.

Of course the researchers recognized that that probably wasn't feasible in terms of response burden, and universally researchers, except for survival, thought all of the outcomes were a little less important than patients and families. But if we triangulate and say what did they both agree on, they both agreed on physical function, cognition, mental health, and return to work or prior valued activities.

So this is a whole program of research with lots of different lenses to figure out what seemed to be important outcomes.

To triangulate across every single study that's been done, then, over several years, it seems like important outcomes are survival, physical function, cognition, mental health, return

Page 147

to work, and quality of life, across a whole series of different kinds of studies and different perspectives.

So we've maybe got some thoughts around that, but we're also going to need to think how are we going to measure these if we're doing quantitative, empirical research? We actually did another systematic review, and we found that there are only 20 studies ever published in critical care that looked at the measurement of any instrument for ICU survivors.

There's a dearth of data, and most of the studies using COSMIN reporting weren't high-quality studies or high-quality reporting. That spurned us because we had this grant from the NIH that allowed us to do at least a number of other psychometric studies, really aimed to help populate and provide data for the upcoming Delphi across a number of mental health fields and physical fields. We at least gave some more data to inform the field, and probably maybe almost doubled the number of studies that had been published ever before.

Page 148

All of that's leading up to an international Delphi consensus process. All of that is just the prelude for the main thing. For people that don't know what this Delphi process is, it's a way of achieving consensus among experts when there is no empirical data or inadequate empirical data. We're just trying to get expert opinions and expert consensus.

To do this, we have to have a panel of informed experts, which we strongly believe needed to be patients and family members, as I'll talk about. Everybody in the panel needs to remain anonymous. The panel members didn't know who else was on the panel because we don't want one person to influence another person. What makes that different than maybe this meeting is the loudest voice or the most influential person has the same say in a Delphi as everybody else. We feed back iteratively results of the Delphi after rounds, and people can reconsider their results if they want, but it's all anonymized, and then we have an a priori criteria for what consensus is.

Page 149

What this figure shows is, first of all, who's going to be on this panel. We used a lot of different things to figure out who might be on the panel. The way we finalized it is about half of the panel -- there is I think 77 members -- half of them were clinical researchers who were our target audience, but a quarter of them were patients and family members. A quarter of them were clinicians, and then a few were U.S. funding bodies.

We defined consensus such that one of those minority groups, patients and families or clinicians, could totally veto us reaching consensus. Even a portion of those family members thought that we're out to lunch, that we could not reach consensus. So they were about a quarter each and strongly empowered.

Because we used the InFACT umbrella organization, we had an official representative from every InFACT member group around the world. That means that the Asian critical care trials group, the African group, the Latin American group, the Greek group, whatever; every single

Page 150

organization provided a representative to give us some international coverage in terms of clinical researchers.

We also recognized that not everybody's part of a clinical trials group, so we also had some key leaders in outcomes research, and we randomly sampled corresponding authors from that database from the scoping review. We had federal U.S. funding bodies. And then we had patients and caregivers from Canada, Australia, the United States, and the UK.

We also had official representatives of critical care nursing, critical care medicine, and critical care PT from the same four continents. In the U.S., we also had international critical care groups, an official representative from the SCCM, CHEST, ATS, et cetera.

Those were who were part of this Delphi panel. We then presented to them outcomes. I gave all this stuff, but then we went to the panel with all that information I just presented and said, okay, here are 19 different outcomes. Are there

Page 151

any missing?

So they suggested 8 others, but as it turned out, none of those 8 others made it into the consensus. And we asked them to vote using all the data that I've just presented to you on what they thought were the core minimum set of outcomes that should always be measured, and they voted without thinking about any measurement properties of an instrument.

There might've been an outcome that had no instrument because the focus was what are the important outcomes? There may not be an instrument; there may not be a valid month. Let's just talk about what's important as an instrument. Then, there are 2 rounds of Delphi for that around core outcomes, and we're so fortunate to have 97 and 99 percent response rates across the 2 rounds for the core outcomes, even with patients and caregivers.

We went on and did three more Delphi rounds to look at the measurement instruments, and what we did there was we presented to them 38 measurement

Page 152

instruments. The panel then suggested, well, there's 37 additional ones that they thought we should think about, so they got put into the mix for voting.

Every single one of those measurement instruments were summarized in a standardized way that we hoped were understandable to patients and caregivers as well. They often had videos showing how you do a 6-minute walk test, how long it takes, what's involved, what's the cost, and what are the psychometrics of the instrument.

Here for these 3 rounds of voting, we asked them to specifically look at feasibility, cost, measurement instrument, measurement properties, et cetera, and we had 91 to 97 percent response rates across the 3 rounds there.

How did this all turn out after this 5 years of work? The Delphi panel agreed on these 8 outcomes as outcomes that should be measured in every single critical care survivorship study: survivor; health-related survival; health-related quality of life; mental health; pain; pulmonary

Page 153

function; muscle and/or nerve function; physical function; and cognition.

How do we measure these? That was the last 3 rounds of the Delphi. For survival, we didn't ask groups; we just suggested that you measure a date and a location of death rather than dead or alive at 90 days. No; you measure the exact date of death so you could do survival analysis if you want.

Health-related quality of life, they agreed on the EQ-5D measurement, which is small and easy. For those that want more detail, they agreed on the SF-36. They reached consensus; two different measures for mental health, hospital anxiety and depression scale, and impact of events scale revised that specifically measures PTSD.

For pain, their consensus was don't have a new pain instrument; use the EQ-5D pain measure. On the bottom row, they reached consensus that there is no feasible way to measure pulmonary function. They didn't think surveys where appropriate, and they didn't think that we could

Page 154

mandate spirometry for instance so that every study had to include spirometry. We all know that that really isn't so feasible. So the consensus was there is no appropriate way to measure it.

For physical function and muscle or nerve function, I read every single comment from every single participant, and there's this big tension between people wanting to do performance-based measures. So you have to have the patient in person, and you measure their strength physically, and you have them do a walk test. People thought that was the best way to do it but felt that it wasn't going to be feasible, and we shouldn't make that a mandatory minimum measurement.

So there was no consensus, but if people are able to do in-person testing, the greatest consensus around standard manual muscle strength testing, grip strength testing, and the 6-minute walk test, but the group didn't feel that those should be made mandatory because they wouldn't be feasible in large-scale studies.

Then for cognition, there is data showing

Page 155

that in our population of acute respiratory failure survivors, the mini mental status is very poor measurement characteristics, so there's no consensus reached. The greatest interest was in the instrument called the Montreal Cognitive Assessment scale, which has had very little valuation and didn't reach consensus because there's not a lot of data. There's a little bit of preliminary data that's come out that shows us that it may have some challenges.

To put this in an easy to understand way, the Delphi panel agreed on measuring using these three instruments: EQ-5D, HADS, and ISR. This would be 42 questions, take 12 minutes in ICU survivors, and cost about a $1.50 per assessment. If people want to add on cognition, which we didn't reach consensus, it could add the MoCA BLIND, take a little bit longer, go on to deep dive and quality of life that could add in the SF-36, which would increase the cost, and the time, and the questions. If they want the Cadillac version, then they could do all of those at a cost of around $3 and about 26

Page 156

minutes of time. All can be done by phone.

It can be done in 15 different languages. All of these instruments happen to be available in 15 different languages. There are lots of other research agenda that I won't get into for the sake of time, but there's lots of other things. We recognize this as a very early start. There's a whole years and years of research to make things better.

We also are actively seeking input from research participants, once they've gone through the core outcomes set, what did they think of the experience, and the same from the research staff. You just administered this small battery; how do you think it went?

This is the article that I talked about, the full-text article. We've got lots of information at our website if you're actually interested in this kind of research. And importantly, that website, if you're interested in measurement, this website has lots and lots of different measures, these standardized instrument cards.

Page 157

I get emailed most days of the week with, "Dale, how might you measure this? I'm interested in doing a study of ICU survivors. I want to measure cognition. How might I do this?" This website gives all sorts of guidance on that. And if you want to do your own core outcome measurement set in Delphi, this gives resources.

If you're doing these longitudinal studies, how the heck do we keep patients in these studies? Again, there's lots of free tools. We've got a database with more than 600 ideas for cohort retention strategies based on unpublished studies. We've got lots of tools that are free. This is all funded by the NIH, free, everyone can use. We've got checklists how to search for patients that have become lost to follow-up, and then we've got lots of statistical things that have been published that Elizabeth will talk about a little bit later.

So lots of things out there. Hopefully this is -- really, it's a case study. I know this isn't directly related to sedation, but gives one way of incorporating patients and families, one way of

Page 158

thinking about consensus, and one way of trying to think about measurements. So hopefully that helps inform things a little bit. Thanks.

(Applause.)

Q&A and Panel Discussion

DR. WARD: We'll take some questions and add Ingrid to the panel.

DR. SKROBIK: Dale, thank you a very nice summary of a complex topic. Do you know if there is a relationship between all of the elements that you captured and societal impact and cost?

DR. NEEDHAM: Yoanna, is the question, if somebody has problems in mental health, is there an association between that and societal impact?

DR. SKROBIK: If I could take issue with mental health as a term, I think psychological wellbeing might be -- sorry for the -- because you're traumatized beyond belief, and then you recover, right? I don't think it's a mental health issue. It doesn't make you schizophrenic. But I think -- sorry for the -- I think we as physicians don't talk about psychological wellbeing, our own

Page 159

or that of our patients, enough, and some of your work has.

So we know that it happens often. We hear stories of how burdensome it is. Is there anyone who's linked the two to reflect our earlier discussion about societal costs and the argument that this may be a financially interesting dimension to pursue. Do you know that?

DR. NEEDHAM: I can't think of a lot of -- there's other smart people around, and Mona in the back, too, to chime in. I can't think of any rigorous empirical papers that have shown those. I think many of us who see these patients routinely, in clinical or research studies, know that they're intrinsically linked. We know the person that doesn't survive --

DR. SKROBIK: So Margaret's shown it, for ARDS survivors, but you've got much larger non-ADRS --

DR. NEEDHAM: Yes. But the issue was around quantifying in dollars; I can't think of it. We've

Page 160

published a couple of papers looking at return to work, both from a national sample and a Baltimore-based sample, and we actually put the economic valued lost income. We worked with a labor economist to try to put dollar signs on that.

DR. SKROBIK: But do you have a grading of the HADS that say in relationship to whether you do go back to work? I think that's the --

DR. NEEDHAM: We have a graduate student in Seattle, University of Washington, who's looking at return to work in these kind of factors, using our data. But other people -- am I missing the study that people can think of?

DR. FLOOD: It's an important aspect. I just want to jump in, Denny [ph]. I'm Pamela Flood. I know many of you personally. Denny asked me to give you kind of a quick summary of -- I'm here as a patient, of my experience as a patient.

I became ill around the same time as David did, and I was working. Mervyn was my boss at UCSF. I was the director of OB anesthesia and had a background in preclinical and clinical research.

Page 161

I'd been a plenary lecturer in Hong Kong the night before, and like an idiot, flew back, and I was on call the next night.

I felt awful. I thought I had a sinus infection. But nonetheless, I showed up to call as good anesthesiologists do, but I was vomiting uncontrollably in the sink, and my very wise colleague said, "Sorry. You can't take call if you can't stop vomiting."

(Laughter.)

DR. FLOOD: So brought me down to the ER where I was vomiting and had a small fever. They thought I had some sort of virus; maybe SARS because I had been in China, so they kept me aside from everyone else, and then shipped me over to Mount Zion, which is the less acute care hospital. But I had a severe headache, and a stiff neck, and a fever. And my husband showed up and insisted that they do an LP, which they did, and I had no cells but very, very high protein.

So the long and the short of it is I had an autoimmune encephalitis. I was intubated for about

Page 162

a week. I have some memory of being lightly and deeply sedated with propofol. I have an amazing memory of my extubation on dex, and I have to say, Mervyn, it's a great molecule; I love it.

I remember very, very clearly, very specific details. I remember Mervyn coming by. I remember Ted Eger [ph] coming by. I remember Larry Saidman coming by. I remember wondering if Mervyn was mad because I had missed some calls.

(Laughter.)

DR. FLOOD: I remember feeling my chest go up and down and thinking, "Wow! That's so interesting. I'm on positive pressure ventilation." I remember I had some amnesia. I have a master's degree in neuroscience, and I remember, even intubated, thinking, "Wow! I have plastic hippocampal amnesia." It's really very short-term memory loss. But I didn't care. I did not care at all. None of this was worrisome to me at all. I found it intellectually fascinating, but I did not care.

My experience, I was never -- well, I guess

Page 163

I told my husband -- they told my husband I might die, but as soon as I woke up, I knew I wasn't going to die. So perhaps my husband should be here to give you a family perspective because, frankly, I had my own struggles coming back to work and recovering. But I think he has PTSD, and I think that's an important consideration. So just a quick summary to put my being here in context.

Mervyn?

DR. MAZE: Yes?

DR. FLOOD: By the way, were you mad at me for missing call?

(Laughter.)

DR. MAZE: No.

I have a question for Ingrid, and this follows on something I mentioned early to a couple of people, which is that you guys have done such a good job with these guidelines assessing the evidence, that you'll be taken over by machines pretty quickly, i.e., machine learning will do this for you or do this for subsequent generations quite well. But I worry about the patient qualitative

Page 164

aspect of it, where the evidence and measuring the quality of the evidence is going to be so challenging, yet so important, and something machines will not be able to do.

So I'd like to get your perspective of whether the qualitative behavioral aspects of the work that you're doing, will it achieve a level at which we now look at RCTs and meta-analyses, and so forth?

DR. EGEROD: Thank you for that question. I was very happy to be invited here because I have felt for the last 20 years, as a qualitative researcher that I'm a nurse and that very few physicians would ever read something written by a nurse. So I think it's tremendous that you're even asking the question because I've experienced that some doctors say, well that's interesting, but it's overwhelming that there is very little interest in the kind of research, asking the patients how they do -- if they want to have that kind of research, they want someone else like an anthropologist or some real scientists to do it.

Page 165

So anyway, I'm very happy that there is concern about that. I definitely feel that it should reach status, maybe not equivalent to RCTs, depending on what is your measurement. I'm very aware that basically survival is number one, and you don't learn about survival necessarily from qualitative research, but I think it is so important.

Also, I feel that your whole study, Dale, shows that, yes, you can do these big triangulated studies where you do get something generalizable from a qualitative research, but it might not always be the goal. I think the human reaction is so individual that you always have to have remember somewhere in there that there is something that cannot be generalized and that has to be seen and understood in context.

So I think the two things should always go hand in hand. They're important in different ways, but they're definitely important to understanding why we want to survive.

DR. BROWN: If I could just weigh in, I

Page 166

think Ingrid has said it very well. It's both/and. We need the narrative pieces in here, but we need the algorithm to know where to fit a narrative in. But if there's anything I wished the health care system would do right now is put my values and Goals as a unique human being in the chart. You can't find it. You cannot find my values and goals in an electronic health record of Epic or any of the other commonly used ones. I'm just like everybody here, and you're just like I am, and we're very clearly not.

DR. FLOOD: I don't know how you would put your values and goals until an electronic health record.

DR. BROWN: You'd actually ask a question. You'd ask a question, and you wouldn't have a billing document serve as your clinical record, which is what our electronic health records are, billing documents. I will now muzzle myself because I have some passion about that.

DR. FLOOD: Mona Hopkins at the back.

DR. HOPKINS: I want to go back to Yoanna's

Page 167

comment about the link between our outcomes and social outcomes, employment and funding. We don't have a lot in ICU. We have some hints there are interrelationships across these domains. But we shouldn't ignore the huge literature in traumatic brain injury that shows cognitive impairment is directly related to return to work and your financial status.

We shouldn't ignore the work and caregivers, and how that impacts family values, and we have some of that in the ICU and impacts their financial income. And we shouldn't ignore the work that comes out of our veterans and other people with PTSD, showing that that directly affects their ability to return to work and affects their financial outcomes.

DR. FLOOD: Steve?

DR. SHAFER: In terms of looking at the long-term outcomes, a lot of it is clearly based on reaching people by phone and by following up. I'd just like to ask you if that's becoming increasingly difficult. This is sort of a nuts and

Page 168

bolts question, but I watched Pamela, who's doing some phone outcome studies, curse at the number of people who won't answer their phone because it's going to be another spam call. And it seems that that's actually really consequential in trying to pursue this.

DR. NEEDHAM: So you're right. This is Dale Needham; a couple of methodologic things. With our studies -- and Mona Hopkins was a principal investigator with me on a lot of the work I presented -- we would have 2 call centers, for example, in Utah and in Baltimore, for phone, and then we would have a subset of the entire research network where we do in-person assessments as well.

I think they're both complementary and they're both necessary, but in-person assessments aren't feasible to do across a thousand patients with our current funding budgets. They're feasible. It's just people don't have enough money.

DR. NEEDHAM: Then how do you get people to answer the phone?

Page 169

DR. SHAFER: So that's actually why, when I created this grant almost, whatever, 8 years ago or something, I insisted that aim around cohort retention methods. Peer reviewers weren't keen on it, but having done so much of this, I knew it was absolutely critical.

So we have lots and lots of different approaches to doing it. It takes an awful lot of work. For example, one of the key things is collecting contact information at the beginning of the study and making sure the contact information actually works.

We've got a study going on right now. We were given 3 numbers for a patient in Nashville. Two of the 3 numbers didn't work, then we were down to a single lifeline to connect to that patient. Two of the numbers didn't work to start with at all because the patient had memory problems, and when they provided them to the research staff, they gave them wrong numbers.

So there are a number of best practices. And in the studies that Mona and I've done that

Page 170

have enrolled more than a thousand patients from 48 hospitals across the country, we've had cohort retention rates of 97 percent at 6 and 12-month follow-up. So it is possible, but it takes time and persistence.

JP Kress was talking about it last night with follow-up work that they're doing in Chicago, and in-person follow-up work in disadvantaged communities that there are safety issues, and challenges, and things as well. But it is possible when we treat it like a science. We also design research strategies and budgets for doing this and hire the right kind of people.

The staff that do in-patient ICU studies are often not the same kind of staff that should be doing follow-up. They just don't have the mind-set, and it's just something that most of them find uncomfortable, which is why we found that centralized call centers with specially trained staff often are a much more successful approach.

DR. FLOOD: I've already downloaded your deadlines and cohort retention.

Page 171

Tony?

DR. ABSALOM: Tony Absalom from the Netherlands. A chap in the Netherlands has developed an app that allows patients to do longitudinal, quality-of-life assessments. It sounds attractive to me, but I'd be interested in your opinion because for myself, I worry which patients would be the ones that would respond and how many would. How many would respond to an email? How many would even look at email?

DR. NEEDHAM: I think it's very challenging. I think there's a huge selection bias. I think that our patients -- as JP and I were talking about last night, some of our patients that are the hardest to contact often fall into two categories. Some of them are completely great, back to work; "Why are you bothering me? I'm fine." And they don't even appreciate that the vast majority of survivors are not fine. They don't recognize they're an outlier, and they're busy, and they're back to their normal life.

The others are patients that have an awful

Page 172

lot of challenges. We've had patients say things to us like, "I just couldn't pick up the phone. I felt so down in the dumps that I just couldn't." But we are persistent, and then say, "Thanks for not giving up on me." After we've done our 50th phone call, rather than saying this to us, cursing at us, they say thanks for not giving up on me, or I knew it was important because you didn't stop calling. When we get trained in our social interactions, we believe that somebody is going to curse at us, right? But in fact that almost never happens, and it's the exact opposite.

So there's a huge selection bias around that. For each patient, we do need to take an approach that will work. Some are email responses and some use apps. In our experience, most don't use any of those things, and we found phone is the best way to -- and then you've got an idea of how much time and effort people are putting into the answers and are they understanding it. You're having a human connection. So we found that much more successful than an app, or email, or trying to

Page 173

mail a questionnaire out to people.
DR. FLOOD: Yoanna?
DR. SKROBIK: I just wanted to ask all of you what you thought of the value, the therapeutic value, of the narrative in follow-up studies and how to capture that. I was surprised when I did the Towards RECOVER study with Margaret Herridge in Canada, that patients were grateful for the capacity to tell the story. And I learned -- because, like most people, I knew everything at 30 -- that if you tell the story in your own words, that's part of the journey back.
I'm curious about people talk about the burden, and in the Canadian critical care trials group that I belong to with Lisa and others, the nurses always worry about burdening patients with follow-up studies, whereas my observation is that some of them don't care, don't mind, but there's a spectrum. Who are we harming, who are we burdening, and are there any that we're helping in those evaluations? I welcome your thoughts.
DR. EGEROD: I think one problem we have

Page 174

with a lot of the kind of narrative responses and other interventions we do to try to help the patients, like ICU diaries and other kinds of follow-up, is that we often measure it on SF-36 to get the quality of life, and they always show nothing. It's very distressing, that we know there's something out there. We know there are some values out there. We know it's good to tell your story, but we can't find the instrument that shows the value.
DR. FLOOD: There are a couple of hands.
DR. BALAS: I think to follow up on that question -- Michele Balas -- I'm wondering is it safe, Dale, to assume that the core outcome measures that you're suggesting for the long-term follow-up, would those same measures be applicable to use, like the pre-ICU, before ICU. I wonder if we have that core set of measures because that's obviously one of the challenges that we have. How do you know this is different from their baseline?
DR. NEEDHAM: Yes, and that's a very tough issue. We didn't tackle it at all, so I don't know

Page 175

what the answer is. I can give you an opinion. Sometimes we try retrospective recall using the same -- a couple of the psychological instruments, you can't use that, but the SF-36, you can say think back to before the onset of the illness that brought you in the hospital and score it.
We have some results that show that proxy and patient results are quite different, so we generally rely on patient rather than proxy. And we've got some results that showed dramatic differences that seem to have face validity, but of course it's tainted by recall bias, and your current state may influence how you see the past.
But I think that's sort of a starting point, but I think it is a really big problem and issue. I think there are some innovative studies happening where there are ongoing large-scale prospective studies -- Lauren Ferrante is one of many people doing these -- where they're just prospective studies and things are measured. And some of the patients happen to end up in the ICU, and therefore you have a truly valid prospective. But that takes

Page 176

large-scale studies. You need to enroll an awful lot of patients to get a few that go into the ICU.
DR. BALAS: Then I guess the question also comes up with the validity and reliability of the recommended core outcome measures in terms of a patient that has known or preexisting cognitive impairment. So now you have patients that have cognitive impairment, and are the anxiety and the depression tools valid and reliable to use with someone with cognitive impairment? We always get this by reviewers, and I don't know how to answer it.
DR. NEEDHAM: Yes, exactly. I know that there are no data, at least based on our systematic review, in ICU survivors around that. I guess whether they're preexisting, and then certainly many patients have post-ICU cognitive impairment, it becomes a judgment around are the answers consistent, and it becomes sort of judgment, which stresses the importance of training the people who are doing that, having lots of contact with Mona and I working together -- Tim will chime in, in a

Page 177

second -- when Mona and I talk together, if there are things that we're not sure about, we'll have a conversation as PIs and say what do you think of this?

Tim?

DR. GIRARD: Tim Girard. I actually didn't have a comment but a question. I feel like it probably will sound like a loaded question, but it's not, and this is for Dale and Mona and anyone else. You've alluded several times to the lack of psychometric data on many of these measurements in this population. Can you tell us -- and when I hear that, I feel like you're implying that it needs to be done and that the measurement qualities may be different.

Can you talk about why that would be? When I think about that, I'm not sure that they would or they wouldn't. Mona just got through saying we've got all this data from other populations about various aspects of what we're discussing.

DR. NEEDHAM: I think a classic one -- and Mona can chime in; she's more expert than me -- is

Page 178

the Mini-Mental Status Exam is the world's most validated, most used screening question and seems to work well in lots and lots of populations. We then administered it in ARDS survivors against a reference standard cognitive testing and found that the performance characteristics were very poor. That's one example.

Other examples would be -- and why I like that mixed-method study that we did, that looked at the measurement instruments, patient-reported outcomes versus qualitative, is some people, as Ingrid said, feel like the instrument itself isn't capturing what's really important.

Mona, do you have more?

DR. HOPKINS: Yes. When we look at the MMSE, these were developed for elderly patients to identify dementia. And if you look at the ICU outcome studies, with one or two exceptions, the mean age in those studies is 52, which does not anywhere near meet the criteria of elderly. Certainly, most of the people who are 52 that are healthy and have no other disorders don't have

Page 179

cognitive impairments.

So one of the problems is we're taking a geriatric measure designed to detect a degenerative disease and applying it in an ICU population. So the Mini-Mental Status may be uniquely different, where something like measuring depression doesn't seem to have those differences as much across populations.

So on some measures, I think it works better than others, so I think it's important to get data -- and Joe Bianvenilla [ph] has done a lot of the [indiscernible] Conley mental health through psychological wellbeing, although anybody that designed those measurements wouldn't call it psychological wellbeing -- and look at these instruments in specific populations because you can't just pull something like the Mini-Mental Status design for a specific purpose. And in other populations where they've been used, they've done these validation studies as well.

DR. NEEDHAM: For people who are critical researchers, can we embed a study within a study?

Page 180

So you're already going to be using these instruments. Can you find a way to embed a reference standard or something? And that's what we did. We had a psychological, semi-structured diagnostic interview for PTSD embedded into another study, so we had sort of a psychiatric diagnosis of PTSD in a subset of patients while we're using this screening instruments.

So can we build those things in, partnering with our colleagues who are experts, and say, hey, we've got an ongoing study; could you contribute to it and do this so that we learn? And we're doing that now with HADS, a depression screening instrument and a semi-structured diagnostic interview for depression, for instance.

DR. FLOOD: David, I know you had a comment, and then a question from Richard.

DR. BROWN: I'll make it very brief. Yoanna had asked a question about narrative, and another personal experience. I think I wrote about my illness a lot. I think that was very helpful, but I can tell you, having watched my wife who I think

Page 181

still has PTSD because she thought she was going to have to take me to dialysis for the rest of my life, she finally, 6 years after the event, wrote about it. And I can tell you, it was quite liberating to her to get her feelings out about it.

So I think there's some healing that goes on in those narrative descriptions.

DR. FLOOD: Richard?

DR. RIKER: Yes. A question for David and Pam. You both implied a little bit the difference in the sedation quality between dexmedetomidine and propofol in a manner that I don't think we would ever capture with a RASS score or time to extubation. I wonder if you can embellish your descriptions a little bit, or in physicians, how would we capture this? What is it and how would we capture it?

DR. FLOOD: Well, propofol, of course, it depends how deeply sedated you are. While I was deeply sedated, I have no memory at all. While I was lightly sedated, it wasn't that it was unpleasant, but I was very aware, for instance, of

Page 182

time now, and I was very aware that I was not sleeping. I appeared to be sleeping, and no one could tell that I wasn't sleeping, but I was not, and I wished I could go to sleep. In fact, I felt fatigued.

Then on dexmedetomidine, I felt that my mind was much clearer, and in fact I was even aware in which ways my mind was not normal and not clear. So I much preferred that feeling. In some settings, a patient might prefer to be unconscious. Something truly awful might be happening to them, and they might have 5 million tubes coming out of them, and that might be a period that would be better to forget.

FEMALE VOICE: It depends on the person.

DR. FLOOD: Yes, it very much depends on the person.

DR. EGEROD: We have a non-sedation regime at one of our Danish hospitals, and we invited one of the patients there to tell her story. She happened to be a nurse from the same department, so she also knew both sides.

Page 183

Her message was, being non-sedated is the worst thing she's ever tried, and that kind of went against everything we were working for. But we have to listen to all these perspectives. But it is difficult when you get a lot of patients that say it was good to be awake, it was good to interact with the staff, and then you have some that say it was just horrible. Then what do you do?

DR. FLOOD: Well, did this individual have pain? Because I think there's a huge differentiator. I did not have a condition that caused pain.

DR. BROWN: Part of mine was I was too hemodynamically unstable to have much propofol. And when I was really hemodynamically unstable, they didn't use dex either; it was just native. But the part that Pam talked about, it took me 2 months to sleep more than about 20 minutes after recovery, and that was probably one of the single things that was troublesome as far as the fatigue and starting to feel cognitively intact.

Page 184

DR. WARD: Last question.

DR. FLOOD: Denny, do you want to have the last word? I think we're getting to the end.

DR. WARD: I don't want to delay lunch, but to bring us back to the sedation in the meeting, Dale talked about a lot of validated, from many directions, outcome measures, but he started off with saying, well, these are optional. If you want to do this, these are the things that you can use.

Is there a strong enough correlation between what goes on in the ICU as far as sedation that will provide a measurable signal in these outcome measures that you've talked about; that we should move beyond saying, well, if you want to measure some outcome measures, here's some good ones you can measure. And say if you are investigating sedation in the ICU, you should be, must be, measuring these outcomes because there is a signal there that is measurable, either qualitatively or quantitatively.

DR. NEEDHAM: I'm going to say yes, but I should share the voice with Pratik and/or Tim. I

Page 185

absolutely think this is so important. Our patients have a legacy of problems from their critical illness. Some of it they bring in, some of it's their comorbidity, but some of it is related to what we're doing.

DR. NEEDHAM: There's so much that goes on in the ICU, and sedation is a small -- analgesia is a small piece of it. Is there a detectable signal in these measurements given everything else that goes on in the ICU, or are we just going to pick up noise and we're not going to be able to differentiate propofol versus dexmedetomidine versus something new coming along in these?

DR. FLOOD: Pratik?

DR. PANDHARIPANDE: So going back to the guidelines, when we are creating the guidelines and creating the priority list as far as outcomes and as a result of which we have a lot of the conditional recommendation and low evidence, all the outcomes that were deemed important align very similar to the outcomes which were in the core outcome group; not the set of instruments but the

Page 186

basic teams of cognitive impairment, mental illness, et cetera.

So that's one point. I feel that there is a lot of similarity between what was identified as a priority area in not a formalized Delphi method, but a prioritized scoring that the experts and patients did in the PADIS guideline.

So that was one point. I feel that there is a fair amount of similarity, and there should be little reason why we have to go far away from some of the things that Dale presented.

The other thing is, looking at Dale's outcomes, which are related with acute respiratory failure, if you look at one of the strongest indication, at this point at least, that people tend to use, they all seem to be linked. The majority of patients in the ICU who are mechanically ventilated are sedated, and they are mainly in the ICU for respiratory failure.

So I think they're all, again, hand in hand with that regard, that these are all related, so there should be very little reason I think for

Page 187

having very different outcomes versus what was presented by Dale.

DR. WARD: We'll go to lunch. We need to be back here at 1 o'clock.

DR. FLOOD: Comment?

MALE VOICE: I'm sorry to interrupt. I just wanted to extend a real thanks to Dr. Brown and Flood for sharing these very personal stories. I think to the extent that we can incorporate these kinds of really deeply personal patient perspectives into our research activities, we can try to approach this aspiration we had as medical students to be the kinds of doctors that are taking care of real people. Those stories were really thought-provoking, and I just want to thank you for sharing them.

(Applause.)

DR. WARD: Back at 1 o'clock.

(Whereupon, at 12:07 p.m., a lunch recess was taken.)

Page 188

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

(1:06 p.m.)

DR. WARD: I've asked Rich to -- there have been a lot of great studies out there already done, and there are some lessons that we can learn different than the lessons from PADIS this morning, but lessons really in the methodology on the great studies that have been the foundation for recommendation; and then Marti [ph] from the FDA, who can help us a little bit with the FDA perspective on the stuff we're doing, because some of this does end up on the desk of the FDA.

As Bob talked about, one of the initial ideas at the FDA head to start ACTTION was how can we improve the quality of the clinical trials that are coming to us that we have to look at to make the approval or disapproval for a new drug or indication for that. So this is all about improving the quality of clinical trials. Then we'll have a panel to kind of put this together and talk about the current controversies and unmet needs.

Page 189

Rich, if you'll start us off.

Presentation - Richard Riker

DR. RIKER: Sure. Thanks, Denham.

Well, I've got the unenviable task of taking us from lunch, so hopefully I can try to keep you awake. It's a little bit daunting to give this talk. We all have very different perspectives, and there are some things that we're going to agree to disagree on, but I think we all carry a lot of evidence that it guides our decision making and also our study design approach.

What I'm going to try to do is to summarize not so much what the results were but maybe some things on the second or third level that may have confounded or potentially confounded some of our outcomes or our ability to interpret some of the studies.

So I'm going to go through some of the older studies and some of the more recent ones, tell them what you're going to tell them, tell them, and then tell them what you told them kind of thing. The control group is critical. Targeted level of

Page 190

sedation is important. Sedative versus other drug therapy, timing is everything, and then in the anticipation of the FDA holding the microphone next, I want to ask some provocative questions.

This is an old summary systemic review by Ostermann. The thing I want to have you look at, it's hard to read, but about the fifth or sixth column is mean percent time at sedation target level. If you follow that down through the rows, you can see that every study has some of that information. But if you look further to the right, time to extubation, length of ventilation, ICU length of stay, the majority of the studies don't have that.

So if we think back 15 or 20 years, the standard primary outcome for sedation studies was how often are you at that target level of sedation. If we look at some of the more recent studies, I think Pratik in their MENDS study was really one of the first to look at something more meaningful perhaps. So they looked at 12-day delirium-free, coma-free outcome. In SEDCOM, we kind of fell back

Page 191

to the percent time and target. MIDEX and PRODEX, the combined dexmedetomidine versus midazolam or propofol studies, looked at percent time and target but had a noninferiority design. Then Yahya in his SPICE study took the real leap and looked at 90-day all-cause mortality.

So we've seen a wide range of primary outcomes that have been targeted for these sedation studies, and I think it prompts a fair discussion about what should the primary outcome be as we move forward. So I want to back up now to one of the real pivotal studies. It's a little bit daunting to stand here and tell JP what he learned in his study, but I'll do my best.

This was really a groundbreaking study that randomized patients to either daily interruption or standard sedation, and also randomized to midazolam or propofol starting 48 hours after enrollment. The target sedation level, which is Ramsay 3 or 4, and in the group that was in the intervention arm, midazolam and propofol and morphine were interrupted daily. The patients were awake

Page 192

following 3 to 4 instructions or became agitated. If they did become restless or agitated and sedation needed to be restarted, it was started at half the previous rate.

I go through that in agonizing detail because many follow-up studies use this same approach, so I'll refer back to JP's study as the methodology for some of the other studies that we're going to talk about.

As you all know, daily interruption prompted a dramatic reduction in duration of ventilation, median ICU length of stay, and the need for diagnostic testing. One of the important things for this study now looking back is that if you look at what percent of the days on study patients were awake, in the intervention group, that was 85 percent of the day. But in the control group, it was 9 percent. That means that 91 percent of the days on study, the control group was never awake. That's an important control group aspect that we need to keep in mind when we look at these outcomes.

Page 193

Interestingly, the drug doses were dramatically lower, intervention group with the daily wake up compared to the control group for midazolam, but there was no difference in drug doses between propofol, maybe reflecting to some degree the duration of effect that we see with those two drugs.

Now contrast that study, where the conclusion was clearly daily sedation interrupted improves outcomes, to a more recent study that used the exact same type of intervention. The study drug was interrupted. Drugs were not controlled. The main difference in the study was that the targeted level of sedation was much lighter. Instead of a Ramsay of 3 or 4, it was a SAS of 3 or 4 or a RASS of minus 3 to 0, but the interruption protocol was exactly the same.

As you can see, the outcome here is the number of patients or the proportion who are extubated, and you can see those curves overlap. The sedation scores, the mean scores were exactly the same in both arms of the study. There was no

Page 194

time to extubation difference, no difference in any of the other outcomes, but there was a difference in the amount of doses of drug that were given and the number of boluses that were given, and the workload for nurses was greater.

The conclusion from this exact type of study was the opposite; sedation interruption doesn't make a difference. In fact, it makes it more drug doses, more work for the nurses, and more bolus doses. So how do we reconcile those two things; same intervention to different conclusions? It's the control group. We need to really be thoughtful about designing a study and incorporating that control group. Is it the standard of care? How do we really want that comparison to look?

We know that there are a lot of studies out there in critical care where the control group difference made a big difference for the study. For partial liquid ventilation, remember the control group did much better than expected. For early goal-directed therapy, the control group did much worse than any other study of sepsis. So

Page 195

incorporating those kinds of concepts as we design these studies I think is important.

I think another component we need to be thoughtful about is the level of sedation. This is Pratik's wonderful study, MENDS study, where they randomized patients to either lorazepam or dexmedetomidine; allowed bedside clinicians to determine the level of sedation as was the standard of care at that time; and then post-study, grouped them into deep with a RASS of minus 3, 4, or 5 versus light, and found that the dexmedetomidine patients had more days free of coma and delirium or just of coma, but there were no differences in ventilator-free days, ICU length of stay, and almost but not quite 28-day mortality.

This is just the median and IQR bar graphs for the coma-free days, delirium-free days, or both. One of the findings of this study, which was different than the prior phase 3 and other dex studies that had been published up to that time, was that the dexmedetomidine group actually got a lot more fentanyl than did the control group. Many

Page 196

of the prior studies had suggested that dexmedetomidine was actually a fentanyl sparing type of intervention.

Pratik and his colleagues did probably one of the best graphs I've ever seen -- oh, how did that get in there? I'm missing a slide here; oh, it's here -- one of the best graphs I've seen, where they looked at the light sedation group, RASS of minus 2 to 1, and then on the bottom half, it's the deeper sedation group.

This compares the fentanyl doses in the two arms. It's a little bit hard to see, but the round dots to the left of each number study, day 1, 2, 3, 4, 5, are the dex patients and their fentanyl dose. The one to the right in triangles are the lorazepam patients and their fentanyl doses. And if you look at the bars that represent the median across the top there, they're very similar for the light sedation group, whereas for the deep sedation group, you can see a dramatic difference in that the dexmedetomidine patients were getting much more fentanyl.

Page 197

I think the interpretation you guys had was that it was primarily -- because it's hard to get patients on dexmedetomidine deeply sedated, and the fentanyl was being used not so much as an analgesic but to try to get them into that target level of sedation.

Is that fair?

DR. PANDHARIPANDE: I think a little bit of both.

DR. RIKER: Yeah. I think the take-home here is that although the study was randomizing for two different medications, the range of sedation targets may have affected the dosing of some other medications.

So let me go back here because I think I got things a little bit out of sequence. Within the SEDCOM study here, one of the things to take note of is the stuff in blue. There we didn't let the bedside clinicians identify the level of sedation. We said it's going to be a light level of sedation in both arms of the study, so a RASS of minus 2 to plus 1.

Page 198

When we look at that, those patients were at that level of sedation to the same degree in both arms of the study. So because this was our primary outcome, it was a negative study. We didn't have a higher degree of compliance or time and target sedation in one arm or the other. It was ideally the same in both.

It turns out that that was probably one of the best things that could've happened because then any future differences in outcomes -- time on the ventilator, incidence of delirium, any of those kinds of things -- could not be blamed on a deeper level of sedation, more coma in one arm than another. In fact, because they were sedated to the same degree in both arms, any of the outcome differences would better be explained by the drug itself or some other factor that we didn't take into account.

So if we look at this, and with that same level of sedation in both arms, the dexmedetomidine group get extubated about 2 days faster and they had some other outcome benefits as well. So we

Page 199

went through that already.

I want to talk about another trial where maybe it wasn't so much sedation; it may have been other things. And that's the study we all know well, the Strom No Sedation protocol. Remember there, patients were randomized to either standard propofol or midazolam versus no sedation.

The no sedation group had some resources that are quite uncommon in the U.S. They had 1 to 1 nursing. If the patient was not calm or comfortable with that, they could have a bedside sitter in addition. They could receive as much morphine as needed. They could receive as much haloperidol as needed.

If they were still restless or agitated, they could get continuous propofol for 6 hours and get that up to 3 times. And if that happened, if they needed that 3 times, they would go on continuous infusion propofol. About 20 percent of the patients in the intervention group actually ended up back on continuous sedation.

I think that's an important take-home for

Page 200

this model. They may have traded sedation for human resources to keep those patients calm and other medications, besides the sedative, to keep those patients calm. The outcomes were quite striking, more ventilator-free days, shorter ICU hospital length of stay, and almost a mortality benefit.

Interestingly, I think this was not ideal from my standpoint. They excluded 27 patients who either died or were extubated in the first 2 days. Those are kind of important outcomes. I wish they had left those patients in. The whole intention-to-treat analysis is critical, but why did some of the patients get extubated and why did some of the patients die? I think those are two outcomes we don't want to exclude patients for.

Yahya I think in his series of SPICE studies has shown us that timing is critical, the timing of sedation when we look at what kind of sedation we're giving and when in the ICU stay are we talking about. Almost all of the studies that I talk about up to this point have been done with

Page 201

enrollment starting somewhere in the 24, 48, maybe even 72-hour time frame after being intubated.

SPICE looked very early at these patients. Data started within 4 hours and really looked at during that first 48 hours in the ICU, a time that most of the other studies had ignored or not enrolled, was deep sedation a significant problem? They treated deep sedation as a continuous variable, the number of deep sedation events you had, and showed that time to extubation, time to delirium, time to hospital death, and 180 day mortality were affected by that incidence of sedation.

This is a very similar analogous study done in a very different population of patients, which basically showed the same thing. If we look at the bar graph in the lower left here, the black bars are the first 48 hours. You can see there's really a trend to the right where many more patients are deeply sedated in that first 48 hours. The gray bars are the rest of their ICU stay, and you can see there a greater shift towards a RASS of zero

Page 202

where those patients are awake.

I think Yahya really showed us that targeting a specific level of sedation after 48 or 72 hours may be missing a critical interval in those patients' care, so I think we'll have time to talk about that.

There was another study that was designed very differently but also looked at that early time frame. This is Gerald Chanques study where they took a group of surgical patients, primarily abdominal surgery, coming to the ICU and randomized them within 2 to 4 hours of arrival in the ICU to either standard care with sedation, which turned out to be light sedation, versus immediate interruption of their sedation.

When they interrupted sedation, they used a protocol very similar to the one that JP had designed and that Geeda Macha [ph] had used in the sleep study, where they only restarted the sedation if the patients were restless or uncomfortable. If they needed that, they could get on continuous sedation for 6 hours. If that happened more than

Page 203

twice in a 24-hour period, they left them on continuous sedation till the next day, and then they started over again.

That interruption was associated with a dramatic reduction in time to extubation, 8 hours in the interruption group versus 50 hours in the standard care group, a dramatic reduction in the incidence of coma and a reduction in the incidence of delirium as well. So I think this early time frame in the ICU is something that we need to be cognizant of as we move forward and design these studies.

I think we could draw some possible conclusions from these findings. Number one would be that the control group is critical to our understanding about the impact of intervention and we really need to look carefully at that standard care, or alternative drug, or whatever we want to design.

I think a second important take-home may be that the targeted level of sedation may alter those outcomes in that in this day and age, light

Page 204

sedation is probably the standard for many ICU patients.

The concept of deep sedation in the ICU and who needs it is an area we've got very little evidence to guide us. I think we all have our biases about who we want to keep deeply sedated and why, but the evidence supporting that is not very great and probably is another area we need to do more investigating in, not so much with RCTs perhaps, but with other design approaches.

The third point would be that the protocol must prevent or monitor bailout medications to avoid confounding our conclusions and perhaps even our outcomes. Then lastly, timing is everything. That first 48 hours is pretty critical.

So I want to finish up, and I don't know if we're going to do questions now or do that after.

DR. WARD: We will bring up the panel.

DR. RIKER: Yeah. A couple of provocative questions, and I target these to each of us in the audience and also to the FDA who will be coming up next. One would be can we take placebo-controlled

Page 205

ICU sedation studies off the table? This is a standard, embraced, religious almost approach to study design that doesn't work in the ICU. It's so cumbersome to try to do a placebo-controlled sedation study. It has its own problems. It's nowhere close to the standard of care we provide.

So if we're going to include a placebo-controlled group, I think there are many issues with it we need to consider. And I would pitch -- again, I'm being a bit provocative here, not necessarily telling you what I think. I would propose that we take that off the table.

Are we beyond time in target sedation zone as the primary outcome? I think we probably are. I think that's no longer a reasonable primary outcome. It's not all that important. It's an important secondary outcome. We need to know how compliant people were with the various sedation strategies, but by itself as a primary outcome, I don't think we're there.

This one is maybe a little bit more controversial. Is mortality too high a bar for a

Page 206

sedation study in the ICU? I would pitch that it is, that if we have a negative study with mortality as the outcome, there may be many, many more meaningful outcomes that we could consider advantageous to us as clinicians, to patients and their families, that we lose if that's our primary outcome. I don't know what the other alternative right one is, but I would pitch to you that maybe mortality is too high a bar.

This is another controversial one. Does ICU sedation really impact late outcomes? Our patients are so complicated with sepsis and renal failure and a bunch of comorbidities that they come in with, and various complications that are occurring during their ICU stay that might or might not be related to sedation.

How much of poor long-term outcome, poor functional status can we blame on sedation? Some? All? None? I don't know the answer to that, but I think it's worth asking the question.

I think to challenge the FDA a bit, is resource utilization meaningful? I've heard them

Page 207

say at this meeting in the past that they don't view it as a meaningful outcome. I think it is a meaningful outcome to us as clinicians. I think it probably is to patients and families.

So if we can get patients off the ventilator or have greater ventilator-free days, similar for ICU length of stay, discharge to home or rehab versus death or skilled nursing facilities, those are maybe more functional types of outcomes; looking at short-term functional outcomes. Then as we talked about this morning, the great range of patient-focused outcomes and priorities that we need to consider probably need to be included there.

I'll stop there. Thank you.

(Applause.)

DR. WARD: The perspectives from when all this stuff ends up on your desk.

DR. SKROBIK: Can I just ask a clarification question, Rich? When you pleaded for no studies where the control group gets placebo, you didn't mean that every patient in every trial should get a

Page 208

sedative, did you?

DR. RIKER: Riker. I did not mean that, but to use a placebo for an arm of critically ill patients for sedation, I think --

DR. SKROBIK: So not make it available is what you meant.

DR. RIKER: I would not design a study where placebo was part of the design.

DR. SKROBIK: Where one arm was unable to get a pharmacological intervention.

DR. RIKER: Correct.

DR. SKROBIK: Because the way you presented it, I had some -- it wasn't clear to me whether you meant that each should necessarily get a pharmacological intervention.

DR. RIKER: No. If a patient doesn't need a pharmacologic agent, I don't think they should get one.

DR. SKROBIK: And that can be part of what you consider a no-placebo group.

DR. RIKER: Yes.

DR. SKROBIK: Thanks for clarifying.

Presentation - Martha Van Clief

DR. VAN CLIEF: Well,, it's an honor to be here.

I found the conversation and the presentations very challenging to me on an intellectual basis. It's an incredible group of people, so thank you for letting me come here today. I'm an anesthesiologist by training. I don't have a huge ICU background, and it's been a while since I was in training. I'm here to give you a regulatory perspective, and I hope that I can add something to this conversation.

This is my disclosure statement that's required. This presentation reflects the views of myself and should not be construed as representing the views and policies of the FDA.

Just as a brief outline of what we're going to discuss today, I want to start off with some regulatory concepts, and then we'll move into talking a little bit more about defining the effect. After that, we'll talk a little bit about measuring the effect, and then we'll finish up with

some requirements for marketing approval.

What does the FDA regulate? Obviously, you know that we regulate drugs. In addition, we regulate medical gases, which is kind of an interesting segment that we control, and also devices.

This is actually a timeline. I thought it was an interesting timeline because it gives you a perspective from the late 1960s to roughly around 2000 as to what drugs were approved that are used for sedation. A lot of these are used for sedation off label. The only drugs that are on label for ICU sedation include the propofol, the midazolam, and the dexmedetomidine, which are highlighted in red.

Obviously, since 2000, we're almost two decades later and we're still -- we haven't come up with any new options. We would love to see some new drugs come out to address the ICU sedation challenge as well as just sedation in general.

Among the compressed medical gases, there are gases called the designated medical gases, and

these gases include the list there that you can see. The interesting thing is that nitrous oxide is the only agent with properties that might be useful for sedation, however, it's typically used for short-term procedural sedation, and that's an off-label use of that drug.

As an anesthesiologist, I was amazed when I first arrived at the FDA and started learning about these medical gases, that inhalational anesthetics are not medical gases; they're actually drugs. I thought that was a unique perspective.

With respect to devices, the FDA also clears devices for uses, and these devices would potentially provide an objective measure of brain function that might be helpful in the setting of ICU sedation. An example is the BIS monitor, which was cleared in 1996, primarily for use in sedation. It's been around for quite a while, and it's been studied in several different settings. I did find one publication in 2018 that looked like in patients with severe traumatic brain injury, that the BIS had some benefits over the RASS.

When we look at the FDA, the indications that have been used that incorporate the concept of sedation include the sedation, anxiolysis, and amnesia during therapeutic and diagnostic procedures, and that's the procedural sedation that's already been addressed through this organization, and also the sedation of intubated mechanically-ventilated patients for treatment in the ICU setting.

Let's skip this slide because you guys already know that. So I'm going to go on now and talk about defining the effect. This is a slide from probably -- most of you have seen this. It's the sedation continuum. It's like what we were taught as anesthesiologists about sedating patients for therapeutic or diagnostic procedures. Of interest, it's a pretty simplistic looking diagram. It's like, okay, it make sense, but again, how do you define minimal, moderate, or deep?

The FDA has never really even evaluated a medication for minimum, moderate, or deep. And as you mentioned before, a lot of these agents came

Page 213

out of anesthesia that were then used in the ICU. That's another, I think -- I think there was some bias early on because they were using this type of approach.

Thank you, Denham. I know that I'm using the slide you used earlier, but I found this slide incredibly fascinating and a bit overwhelming. What I did want to point out from this, which I thought was very interesting, is that this author described this triad of pain, agitation, and delirium as the ICU triad. He also made analogies to the anesthesia triad. So I liked that aspect, and I thought it was worthwhile to kind of think about that in terms of how to manage ICU sedation.

The goal of the anesthesia triad was to develop a balanced anesthetic. We were taught to basically always think of amnesia, analgesia, and muscle relaxation when we were planning an anesthetic for a patient. There are lots of different ways to achieve those things, but you want to have each element to actually provide a balanced anesthetic.

Page 214

I had a colleague -- this was quite a while after I was out of training -- who decided to do a short-term case with just remifentanil in a young guy, and the anesthetic -- the vital signs were perfect, however, the patient remembered everything. So that was a good lesson in making sure you have everything covered.

The ICU triad that was mentioned in this paper includes the pain, agitation, and delirium with the goal of a coordinated approach. I know we're not really talking much about delirium, that that's not a high priority, but I felt like it was worth putting into this -- just for the concept of the triad.

Pain is typically opioids, however, if a patient has neuropathic pain, you may be adding in different medications to help address that. Regional anesthesia is actually becoming quite a prominent option for pain management. Every since ultrasound-guided regional anesthesia came about, we've been putting a local anesthetic in the [indiscernible] plane we can find. So it may have

Page 215

a greater role in the ICU setting.

Agitation is really a normal brain in an abnormal situation, and that's where we've been using dex and propofol mostly. The delirium becomes a little bit more complicated because there's an underlying pathology of the brain. The problem with that of course is that there are risk factors associated with some of the drugs that we normally would give for sedation. Fortunately, dex has probably the lowest prevalence of delirium associated with it, but it's not without its own problems.

We're going to talk a little bit now about measuring the effect, and I think some of these were already mentioned, but I'm just going to go through them quickly. Challenges to ICU sedation sedation trials would be what will be the comparator. As was mentioned in the previous lecture, we're talking about will the comparator actually be the current practice since the combination drugs are usually what are utilized.

How will the patients be randomized? Will

Page 216

they be already on a sedation regimen or will it be something newly initiated?

How will you create standardization and protocols? This I think is kind of tricky because you want to standardize as much as possible, but you have to give people a certain level of flexibility because everybody has their own bias or own comfort level I should say with certain medications. How do you deal with discontinuations? Also, some of what's been obviously a very important part of the discussion today is how to measure long-term patient outcomes.

Trial design, superiority has always been kind of the gold standard for the FDA because it's easier to interpret. But there are some other options. Besides just the placebo-controlled trial, you could use a placebo in an add-on trial and you could also use an active control.

It seems like noninferiority trials are becoming more prominent; at least I've seen more of these lately. I know you guys probably already know what a noninferiority trial is, but the point

Page 217

is that the new treatment may have similar efficacy as a standard drug, however, it may offer some additional advantages such as fewer side effects and easier to administer.

Desirable attributes of an endpoint that we look for are the endpoint should be clinically meaningful. Does it give us a direct measure of how the patient feels, functions, or survives? Does it provides clinically relevant and convincing evidence directly related to the trials primary objective?

Is it reliable, which means consistent and reproducible? Is it sensitive, which allows you to detect changes in the treatment effect? Is it readily measurable and does it reflect accepted norms and standards in the field? The endpoint should be carefully defined in the protocol with its rationale just to make sure that you're really measuring what you're planning on measuring.

What are the considerations when defining an outcome measure? These are also known as clinical outcome assessments, and we want to know is the COA

Page 218

appropriate for a clinical trial intended for drug development? Is there an appropriate target population? Can it identify signs and symptoms that would constitute a clinically meaningful benefit in the target population if improved? And will it allow you to establish the magnitude of change in the score that will provide convincing evidence of a clear benefit?

One thing I've learned a lot at the FDA is statistical significance and clinically meaningful or not the same thing, and we are much more interested in seeing case reports that show us really a clinically meaningful benefit.

I went back and looked at our previous marketing approvals. It was pretty slim. Midazolam was approved in 1985. Diprivan, again, was initially approved in '89 primarily for anesthesia, but then it was approved in 1993 for ICU sedation. Precedex was our most recent approval, which was 1999.

Of interest, the assessment tools for the most recent approvals, Diprivan and Precedex, they

Page 219

used the Ramsay scale. However, as you are well aware, there are limitations to these drugs. Propofol has accumulation as well as the risk of PRIS. Dex can cause tachyphylaxis and adrenal suppression, and midazolam also has a problem with accumulation, and it may be a risk factor for delirium, so we're really in need of some new drugs.

I wanted to talk a little bit about the Precedex trial to give you an idea of how this was approved. Sedative properties of Precedex were evaluated into adequate and well-controlled trials. It was dexmedetomidine compared to a placebo control, and they evaluated the manner of rescue medication required to achieve a Ramsay sedation scale of greater than or equal to 3. One of the trials they used midazolam for rescue; the other, they used propofol. The duration of the trial was 24 hours.

We think that probably 24 hours is too short of a time; 48 would probably be more appropriate. Obviously, I'll talk a little bit more about the

Page 220

Ramsay scale on the next slide.

What was interesting to me, when I looked at this Ramsay scale and what their criteria was, which was great than 3, I thought, wow, that's like a broad swath of sedation. That could be anywhere on that spectrum, and it didn't seem to be very patient driven.

The other thing to determine was the fact that level 6 is comatose like we discussed, so really, what value is that, unless you need a patient absolutely still and unresponsive. Fortunately, the scales have improved. As some of the studies that were previously discussed, there's a lot more granularity in this scale for two reasons. One, the Ramsay scale just has one number for agitation, whereas this gives you greater options for determining agitation. In addition, I like the fact that there are levels that respond to verbal stimulation, and then those that need a heavier level of sedation.

So I think that this obviously is a more effective tool, but I'm not going -- I think that

Page 221

I'm not going to say that this is the effective tool. It's just a newer tool than the Ramsay, so obviously it was designed to give more information.

I'm not going to talk about this very much because you all know about all these assessment tools, and I've just created not a comprehensive list but just an example of the tools that are currently available.

Now, we'll talk a little bit about the requirements for marketing approval, and I want to talk about the CDER Clinical Outcome Assessment Qualification program to finish up. Marketing approval typically involves these elements, a robust clinical program; adequate and well-controlled trials, and typically it's two trials; to provide independent substantiation of the results. However, if it's not a new molecular entity, we may be okay with a single trial if it's a repurposed drug. We would just need a rationale for that, but what's going to be your clinical outcome assessment and is qualified?

Qualified, we'll talk about in just a

Page 222

minute. If you do create a unique scale, that's fine, but you might want to consider getting it qualified through the program that I'll discuss in a minute. You'll need an adequate safety database, and this again will depend on whether it's a well known drug that we are familiar with or if it's a new molecular entity that we have to get more information on.

I'm going to talk a little bit now about the Clinical Outcome Assessment Qualification program. This is actually the website where you would go to get some information about this program. There is even an email address there that you can communicate with people at the FDA.

I know this writing is rather small, but this tells you a little bit more about what the program does. It manages the qualification process, it works directly with the requesters, and it encourages collaboration and multidisciplinary interactions.

Just to know, the COA qualification is basically a regulatory conclusion that whatever

Page 223

assessment tool you're putting forth, that it actually has psychometric features that we're looking for. This particular program, like I said, we're willing to let you design your own, but if it's not qualified, then it may take a little bit more work for us to agree with your study.

That's all I have today. Thank you.

(Applause.)

Q&A and Panel Discussion

DR. WARD: I suspect there are going to be lots of questions, so I'd like to get a group up here on the panel who have had some experience of putting clinical trials together.

DR. SKROBIK: I have a question for Dr. Van Clief. I am heartened that an institution like the FDA would care about patients and their experiences. Do you ever invite people to -- in critical care, one of the challenges we've had over the years in doing trials is that ethics committees will often view ICU patients as being extremely vulnerable, and therefore forbid doing any kind of research rather than ask a question of these most

Page 224

vulnerable people. We've really been effective, across Canada anyway, in militating for having at least an ICU person come and pitch why it's so important.

Is there a process for that kind of clarification at the FDA? I mean it, because here you are. You adjudicate the fate of things that are game changers for people who want to implement whatever. I'm curious what your outside input is, if any.

DR. VAN CLIEF: Well, we evaluate studies that come in sometimes before -- they're called IND exemptions, so investigators take advantage of that approach. If they have a supportive IRB that feels like it's a safe study and we evaluate and we agree that it's something that we don't need to do under an IND, then that may be one pathway.

But the other pathway I think that more addresses your concern is if you submit your protocol under an IND, we have an opportunity to give feedback and see how we can work with you to maybe make that protocol safe enough to go forward.

Page 225

DR. SKROBIK: I apologize. I wasn't clear enough because I talked about to things.
DR. VAN CLIEF: Okay.
DR. SKROBIK: When you decide whether you're going to approve a molecule for use, you have an inside panel of experts like you and there are rules that you can go by.
DR. VAN CLIEF: Yes.
DR. SKROBIK: Do you ask anybody from the outside?
DR. VAN CLIEF: Yes, we do. We have -- I'll let my boss answer.
(Laughter.)
DR. ROCA: Hello. I'm Rigo Roca. I'm from the FDA and the deputy division director in the review division. So to answer your question, yes, definitely. Particularly, if we have questions about a new product, we're trying to figure out what it means, we definitely, as was being described, go through the development program with the sponsor and all that. But at the very end, we also have the opportunity for advisory committees.

Page 226

Within the advisory committee, we have a panel of experts, but there's also a patient representatives there.
There's also a section in the open public hearing where patients can come up and share their experiences, so what's important for them. The panel takes all that into consideration. The panel, the advisory committee members, then give us their thoughts and recommendations, and we assimilate the patient's information, the patient representative on the panel, as well as the committee.
So we definitely do that. But then there's also something else that we do, and recently, that you may or may not have heard is patient-focused meetings. These are actually listening sessions. We've had a couple. Most recently there was one for opioid-use disorder, which was interesting to find out what is important for a person who's suffered from opioid-use disorder. And as you can suspect, sometimes it's different than what we thought was important.

Page 227

DR. SKROBIK: I belong to an opioid-abuse community. That's why I'm smiling because the patients say things that are completely different.
DR. ROCA: Exactly. There was also another one several years back regarding -- I believe it was debilitating neuromuscular disease, and we felt -- well, it was in a different dimension, but the FDA felt that what we needed to do was have a certain degree of mobility, and I think one of them was the ability to walk a certain distance. Out of that meeting, what came out was that patients were happy if their sibling or the family member was able to just simply sit up. That was considered to be something important.
So we do have those patient-focused meetings, and we do use them to learn as to what it is that's important for the patient. As you were alluding to, sometimes it's different than we thought.
DR. SKROBIK: Thank you.
MALE VOICE: Gilles, I think you're moderating.

Page 228

MALE VOICE: Gilles or Doug.
DR. COURSIN: Well, I think someone with a bow tie looks very professorial.
(Laughter.)
DR. COURSIN: I will defer to his kind judgment.
DR. FRASER: You'll notice there are lobsters here. That's no coincidence since I come from Maine.
Today we've had I think a wonderful series of presentations about the guidelines that we've presented and the primary data that were formed as a part of the guidelines, or actually the guidelines were formed from the primary data. We've looked at the methodology, and we've also looked at the outcomes and the metrics that were involved in getting those outcomes.
What I would like to open up with in this particular session is where do we go from here? What do we need to know in order to further the science?
DR. SKROBIK: Just like that?

Page 229

DR. WARD: A question for Marti. I just went on your website and looked at your COAs, at least the PDF file that's up there. There aren't any for any sedation. There's some for pain, which is just a numerical rating scale or a visual analog scale, but there's none that would apply to the things we've been talking about through ICU sedation.

Is it worthwhile to get some of these scales that we've been talking about qualified? Should the Ramsay scale be a COA, and is that worthwhile to help future clinical trials to have that done?

DR. ROCA: I'm being told to say yes, but actually the answer is yes. I think there are a lot of advantages of having a qualified. As the slide mentioned, it is a multidisciplinary team that comes in and addresses it from all different aspects. We have ongoing discussions with whoever it is that is proposing to have a particular tool or scale qualified.

So there is that ability, and then the nice thing about it afterwards is that if a tool is

Page 230

qualified, as you are indicating, then it's actually something that we have already looked through and vetted as being a tool that could potentially be used in different kinds of clinical trials. Obviously, as the last sentence in there, it depends that it's been qualified for a particular use and a particular population, et cetera, as most tools are, but still it would be something that would be useful.

Now, the other thing that was mentioned was that if you have a tool and you haven't gone through the qualification process, you could potentially still use it. We would use the same multidisciplinary team to do that and assess it, but then it's a little bit more within the review time clock, and therefore we may not be able to have as much interaction.

Furthermore, it's already a done deal, and there's a possibility that at the end of that assessment, at the end of the review clock, we may end up concluding that it probably was not a tool that could have generated the data that they felt

Page 231

it was generating. But that doesn't mean that you can't use it; it's just that it might end up not being as positive an outcome as you would have had otherwise.

DR. MAZE: Can I ask a more structural question or rather a foundational question? When you have a scale like the RASS scale, which obviously, as you said, is more granular, are the biological foundations, neurobiologic foundations, for those elements in the scale different?

In other words, when you're producing sedation, you possibly need some different neural pathways involved versus producing agitation, yet you've got them in a continuum. Is there any benefit in having a scale that is actually continuous with respect to the neurobiologic pathways that are involved?

DR. VAN CLIEF: That would be interesting to entertain as a scale. I use that scale just as an example of where we've come from the Ramsay to that level. But I do think that whatever scale is selected, you just really want to make sure it's

Page 232

going to measure what you're interested in looking at and studying. I think the scale you're describing might be difficult to develop, but it would be very good to have.

DR. MAZE: I was kind of surprised at the acronym PAD and PADIS, that sedation isn't mentioned there, but agitation is mentioned there, as if they are the same thing. I know Yoanna is going to say --

DR. SKROBIK: I was going to say the amount of discussion around the acronym was subjective, more energy than I would ever want to admit. We didn't have Dr. Dworkin around --

(Laughter)

DR. SKROBIK: -- for cued acronyms, so this was the compromise that had to do with branding with the similarity of the PAD. But just briefly to speak to the point of the sedation scale and its validation, the FDA metrics don't reflect the previous guidelines effort, where we actually went through all the psychometric elements of all of the scales.

Page 233

What I am curious about is whether this would be an opportunity because you were saying you invite people to come and testify. Here we all are talking about sedatives in the ICU. Would this be an opportunity, without neglecting what we don't know, to say, well, here's what we know; would you like to integrate it just now; the things that we have brought forward and that we have discussed and agreed on? I think that might be one.

DR. ROCA: One of the things about a meeting like this is that my role here is actually to listen, and my role would be to help facilitate the discussion, particularly if you have a question regarding the process of how do we do things, what do we need, and that I think might help the discussion. But with respect to a decision, yeah, this is what we need and this is what we should do, I don't think I can do that.

There are particular reasons for that. Number one, this is not really an all encompassing audience, so therefore it would not be appropriate for me to indicate what would be regulatorily

Page 234

appropriate or not. That would be one thing. The other thing, too, it would be definitely drug dependent, population dependent, and indication dependent. There are so many variables.

DR. SKROBIK: You're talking about the scales.

DR. ROCA: Definitely, scales as well. It depends on what the company is proposing to have their drug do. And they come to us, and they have often asked us which scale we should use, and as you can suspect, we really don't have one that we can say, yes, this is the one you should use because it really depends on what it is that they're trying to have their product demonstrate its efficacy for.

So we usually tell a company that they can choose whatever scale they want, but they're going to have to be able to provide supporting information as to what that scale, or two, is the most appropriate one for the patient population, the indication, the drug, and all of those things. We would take that into our review process, and

Page 235

then at the end, if we still have questions, we go to the advisory committee. Therefore, the ability for me to say anything regarding whether a particular scale is more appropriate than another in this setting would be very difficult.

DR. SHAFER: Mervyn -- Steve Shafer -- I want to directly address your question. The idea is, is there a neurobiology that you can tap into here. By suggesting that there isn't with a comparison to three different drugs, sedation with propofol, and sedation with dex, and sedation with ketamine, what would be very different experiences from the patient perspective, most people find propofol somewhat pleasant to actually experience. Dex seems to be neutral. A lot of people seem to find ketamine somewhat dysphoric at really high doses.

They might look the same on the scale here, but from the patient's perspective, because the neurobiology is so different, I don't think you're going to find a scale that you would put them all on. In some ways, you'll have different scales.

Page 236

One is are you clinically achieving the effect that you want from what you can measure in the ICU, and that would be something like a RASS scale perhaps.

The other is then a more patient-centered thing; what was the sedation experience like? And as we heard from the earlier presentations from Dave and Pamela, that can be quite different with different drugs, and that would perhaps be an orthogonal scale that might be captured as well.

DR. FRASER: In order to get to that point, I think what you'd have to do is allow for wakefulness so that you can gain some feedback from your patient. And that is what I think is the next step in terms of the sedation scales. They really don't evaluate wakefulness, and they don't gather data or feedback specifically from patients.

So I would ask this group at some point in time, if there's appetite for revision of RASS or revision of SAS, to include a wakefulness algorithm such as what JP Kress actually developed in the New England Journal of Medicine.

DR. SHAFER: You don't need a different

Page 237

scale on access. This is an orthogonal access to assess something quite different.

DR. FRASER: Right. So you could use RASS and then supplement it with wakefulness.

DR. EGAN: Talmage Egan. It seems that one of the problems with these sedation scales that have arisen for use in the ICU is that they don't seem to have methodologically as rigorous a foundation just in terms of how they were validated. In the procedural sedation domain, although it's got problems, the Modified Observers' Assessment of Alertness and Sedation, the so-called MOAS scale, has really sort of become the main one used in clinical trials.

The reason is simple. There's quite a rigorous methods paper that quantified the inter-observer variability, and there are also some training materials that are available -- this was alluded to earlier -- that one can use to train the study personnel.

I've seen that there's some room for that here. There are these various scales. They seem

Page 238

to be used because it's what other people have used and there's lots of clinical experience with them. But perhaps some quantification of the inter-observer variability and some training materials would be useful, especially as it relates to quality controlling of the studies for regulatory purposes.

DR. RIKER: There is data available for both of the scales. The 2013 PAD guidelines highlighted some of that, then there was a separate publication that looked just at the psychometrics of the sedation scale piece. There have been a number of inter-rater reliability studies and some validation studies. There are educational things out there. So it may not be at the level of the MOAS scale, but there certainly are things out there.

DR. EGAN: Experts in the area that do these trials, are you guys satisfied with the overall robustness of the scales? Are they missing some of these attributes? What's the key piece that's missing?

DR. RIKER: I'll give you my opinion. This

Page 239

is Riker again. I think Gilles put his finger right on one of the issues, and that's what do they really measure? If you open your eyes, but that's all you do, what is that telling us? Are you awake? Are you able to follow commands? Each of the two scales that have been highlighted look at different things to get to their endpoint.

I think one of the issues is in addition to the complexity of reliability, is it really measuring -- or can we both say the same thing, and then validity, is it measuring what we think it is? As trial designers, how do we use that information? What are we targeting? How do we best apply that level when and in what way? So it's kind of a pharmacokinetic/pharmacodynamic kind of thing. We can measure it, but then what do we do with that information and what are we trying to do with that information.

DR. COURSIN: Well, but there's wakefulness and wakefulness. I mean, are you looking at wakefulness with cognition? And if you're looking for cognition, what level of cognition? I mean, we

Page 240

can give them a computer game, and can they flip cards quickly and tell us they have 21? Can they make executive decisions?

That again also morphs into what Steve's referring to, which seems to be, I'm awake, but I'm delirious, and that seems to have two factors I want. Ultimately, in the ICU, we don't want you jumping out of bed and hurting yourself. We don't want you in bed if you don't have to be in bed. And we don't want to be giving you something if you don't need it.

Now, those are three simple statements, but I'm not sure how to put them into a MOAS type scale. Clearly, just tapping somebody's glabella and having them blink was pretty simple; never particularly well validated. That's Ramsay, which had been the gold standard. I think the RASS and SAS scores are a good stride beyond that, but I'm not really quite sure either what we want by saying wakefulness or whether we're necessarily going to be able to quantify what we want.

DR. FRASER: JP?

Page 241

DR. KRESS: I think Doug spoke to the same question I was going to ask, is wakefulness a bivariate outcome or shades of gray? So I don't need to reiterate that.

I actually have a question for Rich. You talked about the importance of the control group in your review, and I think it's really important. As we move forward, of course the competition with the control group continues to get tougher and tougher because we learned things. That's good.

As we think about moving forward, should we think about control group as a regimented approach or some people talk about this so-called wild type where you just basically let the care providers do as they wish. I wonder if you have any thoughts about that.

DR. RIKER: It's a great question, JP, and I think the idea of pragmatic trials or adaptive trials, I hope we're going to talk about that later on in the meeting. I think it looks like we will be. But the old intervention control, RCT, power sample calculation, I think we're really bumping up

Page 242

into the limits of that for our population.

You look at the complexity of our patients, the varying populations we have, and I think it makes it really hard to -- the concept that a general ICU patient is interchangeable with another general ICU patient, I don't think that works as well anymore. So splitting, lumping, which is the right approach? It's pretty darn complex, but I think you're right on target that we have to ask that question. I don't know what the right answer is, but I think we have to ask that question.

I want to say one other thing. Wakefulness may mean different things depending on how we want to use that information. In other words, if we want to see is our patient awake enough to tell us how much pain they're having, or is the patient awake enough to do a delirium assessment, that might be a different kind of wakefulness assessment than if I keep my patient above a certain level of non-wakefulness, do I reduce their long-term outcome problems? So different issues may need different levels of wakefulness and potentially

Page 243

different assessments of wakefulness.

Tim, excuse me, and then Pam. Tim, you had a comment?

DR. GIRARD: This is Tim Girard. I agree with what Rich just said. Just to take that thought even further, I feel like one area that we have a gap is the relationship between all of the various ways that we're describing, looking at wakefulness or consciousness and the various outcomes that we and patients care about.

For example, I think, Gilles, you're referring to the process that JP used in his early [indiscernible] sedation trial following commands. There's definitely a lot of value in being able to follow a command. But for example, if your decision is whether or not to extubate a patient and if they're alert enough for that, I'm not aware of any data that suggest -- even though it's sometimes used at the bedside, I'm aware of no data that suggests that your ability to follow commands predicts your likelihood of passing and extubation.

Alternatively, there may be other very

Page 244

important patient-centered outcomes that are related to your ability to follow commands. Certainly, a patient who is delirious often does not follow commands, and there are a lot of data suggesting that delirium is related to both short- and long-term outcomes.

So the issue is quite complex, as we all said, and I doubt that there's a single, easily applied scale that's reliable that can capture all of this, the content of consciousness, the level of arousal. It's unlikely, in my opinion, that a single scale would do that.

However, the two scales that are recommended by the SCCM guidelines -- and I was not on any of the guideline panels, so I don't have any, I don't think, bias in this respect. But both of those scales were very well validated. The reliability has been studied in numerous environments and in numerous studies, and it's been shown that both the SAS and the RASS are very reliable and that they are valid in terms of measuring the constructs that they were intended to measure against multiple

Page 245

other reference standards. So I think at least for what those tools are supposed to do, which is measure level of arousal, then they are valid in that context.

DR. COURSIN: One thing from Pam as well is the question of, okay, we want wakefulness. What about restorative sleepfulness? I'd like you to comment.

DR. FLOOD: I'll answer that second. Not to make matters more complex, but I was going to speak to cognition because I don't think anyone really wants to play 21 in the ICU. Well, maybe if they're intubated, it's something to do. But there's pleasant cognition and unpleasant cognition. There's being peacefully sedated and being aware of your surroundings, and then there's being frightened, and distressed, and so on and so forth.

So you might think of that as being described with the continuum of sedation versus agitation, but that's only the behavioral manifestation. You might not know what the -- it's

Page 246

very hard to know what the patient's feeling. I think David and I both spoke to the feeling that everybody thought we were asleep, but we weren't, and we weren't able to sleep, and we were very fatigued.

So getting to your question, the more and more I know about the nature of sleep makes me realize I know less and less about it. But getting real sleep in an ICU setting, at least from what I understand in terms of people who study sleep, this is next to impossible. So I think a better question is what can you do to do the best you can with that and to limit fatigue.

DR. COURSIN: Denham?

DR. WARD: In your discussion, you get a little bit on inclusion criteria and exclusion criteria. Well, one -- let's see if I'm quoting this right -- would be that the first 24 hours is important to the outcomes. For most of the studies that you looked at it and I've looked at, too, that's usually not an inclusion criteria; usually it's after 24 hours. So it would seem like one of

Page 247

the concepts that we're coming to is that patients should be included much earlier into a sedation study that is within the first 24 hours.

Are there other inclusion/exclusion criteria that -- I reviewed a number of studies just to educate myself, and one of the things that I rarely saw was a history of drug or alcohol abuse as an exclusion criteria, that is, is withdrawal going to be complicating the other measurement of these things? But most studies never mentioned opioid-use disorder as a premorbid condition or alcohol-use disorder as a premorbid condition.

What's your thinking about inclusion/exclusion criteria?

DR. FRASER: The more we exclude to try to provide homogeneity in our cohort, the less generalizable that information is. Maybe efficacy, effectiveness, there are a lot of issues that go into what you're trying to accomplish with your study.

I'm speaking way over my head here, and I hope when Dan gives his presentation or our other

Page 248

future discussions about design, we'll get to this. But we've heard of adaptive responsive kinds of studies and platform design studies that may allow us to recognize specific risk factors and emphasize them or better understand the role they play.

Hopefully, as we move into the future, we get away from this black and white intervention control thing and more into design that allows us to try to answer some of these questions, not by excluding those patients but perhaps by including them and building that into the design, so I don't know.

DR. COURSIN: Sir, in the back?

DR. DWORKIN: Rich, my recollection is that you and I first met at an FDA meeting on sedation about half a dozen years ago -- it's a long time ago -- that discussed both procedural and ICU sedation. My recollection is that one of the conclusions of that FDA-sponsored meeting -- this was before ACTTION had anything to do with sedation -- is that in the ICU setting, the target, if you will, that patients would find most

Page 249

desirable is calm and comfortable.
So being from outside the field and sitting here all day, I'm a little surprised that that meeting six years ago ended up with calm and comfort being objective of ICU sedation. I haven't heard that so far today. I've heard a lot about sedation and a lot about agitation, but nothing about calm and comfort.
Is that a reasonable measure to think about developing, ICU calm and comfort?
DR. RIKER: Yes. I think everybody in this room is going to give you a little bit different answer, but I think from my perspective, the evidence, especially six years ago, that supports that claim is quite thin. It's a thing that makes sense. We know the evils of deep sedation; we try to avoid those. We have a little understanding about the evils of not enough sedation, and probably for the majority of patients, we err more on the side of too much sedation rather than not enough sedation.
But I think we've heard even within our two

Page 250

patient representatives today how complex that issue is and that there may be patients who are awake and don't want to be that awake, or patients that are calm but not cognitively intact enough to be comfortable and want to be more awake, and our ability to assess that and understand that is quite limited right now.
So it's a great question, but I don't know how much evidence there is supporting that concept.
DR. COURSIN: Dale, you had a comment?
DR. NEEDHAM: Just as a clinician, not sort of an expert, I want to reflect back what I'm hearing or my biases. I think that we've talked a lot about sedation scales, but I think most people agree that they're not patient-centered outcomes. I think I've heard people say that we probably need development of a patient experienced measurement, which would be totally patient centered around the type of sleep that would be complex to develop and acquire time to develop, and think about, and validate, and reliability.
But then we need to reflect on what is

Page 251

patient centered. And for FDA or other purposes, are we okay with something that's health care centered? There may be something that there's no patient-centered impact, but it reduces our mechanical ventilation duration length of stay. Are those four accurate?
DR. RIKER: One of the things I really liked about what you said is that patient-centered assessment tool. And ideally, that would be a real-time assessment tool as well, not a retrospective how was your stay in the ICU, so that we could respond to that answer.
DR. NEEDHAM: To give you an example, we've got an R01 from NINR looking at laryngeal injury, and in fact when patients are awake, we're asking about symptoms related to potential laryngeal injury. And we've had to take other instruments and figure out how can you do it in a patient with an endotracheal tube in order to try to understand the symptoms that patients are feeling and whether those symptoms are then relevant to a subsequent outcome; so I think a little bit about a process

Page 252

there and how we may need something like that perhaps to understand patient experience.
DR. COURSIN: Tim?
DR. GIRARD: Tim Girard. In theory, I agree completely, Dale, but in practice, I think there's a huge problem, which you, and I, and Elizabeth have discussed extensively, which is that to measure something like that in the setting that we're discussing, you will have a huge amount of missing data because there will be patients who cannot respond at various times, and that missing data may very likely be differential between different treatment groups.
So I would agree that using a patient-centered, real-time response would be a helpful adjunct to understanding what the effects of the different therapies that we're studying are, but I would argue that it could not be a stand-alone because you would end up with too much missing data, and that that different data would be differential.
DR. COURSIN: Michele?

Page 253

DR. BALAS: I'm going to have to agree with that comment as well. I think it would be wonderful to have such a measure, but we're doing a small study right now, and we're just trying to measure anxiety -- again, the reason people give for giving sedation -- and we're missing it on over 85 percent of the patients because of their level of arousal.

So to have a patient-centered outcome report, the patient would have to have some level of arousal, some level of consciousness, however we define that part, just to measure these other symptoms or to get their perspective. And what we're finding in clinical practice and with our work with the SCCM ICU Liberation outside of clinical trials, patients aren't at that basic level. Even though everybody's charting our goals, 0 to minus 2 right now, when you go in and you do those direct observations, they're charted minus 1, minus 2, and they're still deeply sedated. They're still in a coma, most of them.

So there's a huge disconnect between I

Page 254

think -- the validity and reliability of our tool I think is solid in terms of research, but I think that inter-rater reliability again, even in clinical trials, is kind of suboptimal, and clinical practice, I'm going to argue it's pretty nonexistent.

DR. COURSIN: Dale?

DR. NEEDHAM: I would agree with everything. I think it couldn't be a primary outcome. It may be something that allows us to get some insight into that. I was talking earlier around meeting for mixed methods study so we actually get a qualitative experience, and I think it's maybe a tiny bit like delirium, where there may be a group of patients where we can't assess it and then there's a group that we can, and then we need to figure out what is the statistical method to look at these two different -- like where one group of patients can't even have it assessed, and that may mean something in whether it's compensated, or a two-part model, or I don't know what.

DR. COURSIN: John?

Page 255

MALE VOICE: I don't think it could ever be a primary outcome because of that problem.

DR. DEVLIN: The other quick comment I wanted to make was in our PAD guidelines, we, obviously as everybody knows, found widely divergent restraint use, highly prevalent in the United States, very low in Europe. So I think that kind of plays a role; and with that, the nonpharmacologic things that could affect agitation, I think being certainly rehabilitation or mobility, and that whole interface that has really nothing to do with what we're giving for a sedative or could drive sedative use.

DR. COURSIN: Steve, you have a comment?

DR. SHAFER: It's a question actually. I'm having a little bit of a challenge here. It's a question for the entire panel. Let's say that I'm a magician and I can actually produce a drug that does anything you want. I'm trying to figure out in terms of what we're talking about here, what claim would you want that drug to be able to make to actually give you a better patient-centered

Page 256

outcome for sedation?

A lot of ICU trials look at survival, which is a great thing to look at when you're in the ICU, and that's a wonderful thing. A lot of stuff in the area of sedation looks at surrogate endpoints, time to extubation, extubation-free days and things like this, but those are surrogates.

What claim -- you've done all these clinical trials, FDA, introduced dexmedetomidine. What claim would a magical drug that I could give you make that you would actually study and then take to the agency, and the agency would say, yes, this is a valid claim to make for a product?

DR. COURSIN: I'd like to be 25 again, I'd like to know what I turn out, and I'd like to have a full head of hair.

DR. SHAFER: You've got it.

(Laughter.)

DR. MAZE: In the dex trial, all we set out to do was to show that it was a sedative in the ICU patients by virtue of the reduction in risk of medication. That's not a very good endpoint as we

Page 257

now know. It just demonstrated that this drug falls into a particular class but didn't tell us anything about the effectiveness versus other drugs.

For example, the lack of that placebo-controlled group -- rather, the use of a placebo-controlled group with rescue medication was what we used rather than a more comparative effectiveness type of trial, comparing it against perhaps midazolam or propofol at that time. So I don't think we did a great job at defining the endpoint. It wasn't my idea, so I can criticize it.

DR. RIKER: I'll throw something out as a straw man, and then everybody else can weigh in. Maybe it allows you to be calm and responsive so you can say I'm having pain, I want to be more deeply sedated, there's an IV sticking in my left hip that hurts a lot, and doesn't have adverse effects like hemodynamic compromise, cognitive impairment, doesn't make your platelets --

DR. COURSIN: But it's not fair -- all the

Page 258

classics; it's not fair --

DR. MAZE: It sounds great, but we had a name for that. Remember, we called that cognitive sedation.

DR. COURSIN: Yes.

DR. SHAFER: But then how do you get that to become an FDA claim on a label?

DR. FRASER: You can measure the degree of participation in care that influences outcomes like early mobility for example. I think that's a measurable metric.

DR. COURSIN: Avery, do you have a comment?

DR. TUNG: Taking a page from the anesthesia playbook where most patients would prefer general anesthesia if you gave them a choice -- and in fact, we're finding in our hospital satisfaction with general greater than satisfaction with regional.

Here's a claim: allows deep sedation without any of the length of stay, long intubation, delirium, and outcome drawbacks of deep sedation. There's a claim you could make.

Page 259

DR. RIKER: But do you think families would buy into that? Would families want you to be in a box for 3 days or 4 days, and not awake and not responsive?

FEMALE VOICE: I'm not sure all patients would want that either. I wouldn't.

DR. TUNG: If your entire ICU stay would pass by and you wouldn't even know it was there, then that might not be so bad. There have been daily sedation interruption trials stopped because the families didn't like them.

FEMALE VOICE: It depends on the stress.

DR. VAN CLIEF: I just want to make a comment about the indication that goes into the label. It really is a description of what the drug does. And if you go beyond that and say, well, it provides a deep level of sedation and x, Y and Z happened, we won't necessarily accept that because those are promotional claims, and we're not there for promoting, but we want to describe what the drug does.

DR. SHAFER That's what I'm sort of asking,

Page 260

for the outcome. What outcome that they can define and you can falsify; it either happened or didn't happen in the trial.

DR. VAN CLIEF: Right.

DR. COURSIN: Claudia, did you have something?

DR. SPIES: Yes, I have several comments. The point is I fully agree with what Timothy said also about the scores and all those things. I think that's validated, it's globally used, and I think in many settings it's validated. That's the first part.

The second part is I think it's not so easy to say that the scores are really those that, at the end, are the relevant thing because you haven't -- if you aim a RASS score, that doesn't mean you achieve it. This is one of the points. Even if you try to achieve it, it's context sensitive. That means all the nurses, all the staff, 24-7 has to agree on that.

That means, also, if you have a sedation procedure that's really adapted to awake,

Page 261

cooperative, calm, not anxious, whatever people, I think if you're really take that serious, I think you also have to do other things. That means other groups need to be involved, like physiotherapy for example. So if you don't use your muscles in the first 3 to 4 days, you also lose muscle strength.

So it's a lot of composites that need to be defined, and I think what we need is a protocol violation of all studies. I think that's something I will try to have in all of the studies, how many protocol violations do you have due to all that noise you have, and then you make a decision how you can improve that. That's nothing that's bad for the study. I think that's very good if you do that.

I think I'm probably trying to convince my colleagues to do it. It's not so easy, but I think it's the way to be honest to the patients, and then not to get reimbursed at the end for the outcome. The outcome is a measurement. I think what's better is if we really stick to that what we believe in and what is evidence based, what we

Page 262

researched. Then at the end, if we really do that, what we think we should do, and then at the end, we measure an outcome, and then we see if the patients really have these outcomes, and then we need to change the studies.

But I think that's something -- the majority of the studies, at least what I read from all IPEC journals is that the point is that the protocol violations are not given. I know from my studies at least that it's not so easy to do it, and I can tell you I'm fighting with that all the years, and I need help for that.

DR. COURSIN: Yoanna?

DR. SKROBIK: I think there's a dichotomy in what I'm hearing over the last several points that were made between the wish to find one dichotomous administration and the need to individualize, and to individualize not only based on patient preferences. I would consider being sedated deeply a violation of my personal rights, and I know 50 percent of my patients would disagree with that. How do you then choose one answer if the

Page 263

population's that different?

The technology assessment unit at McGill has just gone through the exercise of asking the question, what should determine what you consider standard or the best, as decreed by the technology assessment unit, but it also applies to drugs. And they've come up with a very interesting model that doesn't actually look at the evidence in specific populations but integrates the contextual elements that you talk about.

You have a donor in one institution that wants you to study fear and anxiety. Well, maybe you're going to add that to your questionnaire in that institution because then it will be reliable because you're going to have an extra $19 million to do it.

So with the adaptability, considering the inter-individual variability between the patients receiving the intervention, the carers giving it, and the specific institution -- I can't get my head around the dichotomy between the one model, what would the FDA recommend, as if there were one

Page 264

model, and what I'm hearing about there being many -- the personalized approach, whether for the individual recipient or the individual place.

DR. COURSIN: Mervyn, you had a comment?

DR. MAZE: I actually have a new question if you don't mind.

DR. COURSIN: All right.

DR. MAZE: We've spoken exclusively, really, about symptom mitigation versus disease modification. I presume in the ICU that is a problem because you're dealing with a plethora of diseases. But I would hazard a guess that inflammation is consistently present in your ICU patients, and I'd like to hear an ICU patient that doesn't have that.

So to what extent are their attempts to modify the disease in order to mitigate the symptoms?

DR. COURSIN: Well, there's a huge trail of tears of failed therapies that have attempted. And one of the major problems was everyone was a single magic bullet, anti-tumor necrosis factor;

Page 265

interleukin 1; complement this, complement that. It also speaks to the fact we've had one drug in my lifetime approved primarily for ICU use, and I'm still waiting for the first therapy that we can absolutely say was developed in the ICU that made a bit of difference. But older, sicker people survive in the ICU. We don't know why.

DR. MAZE: Right. I think my view of where the immunology, inflammatory response field is going is that this magic bullet, this anti-inflammatory, whether it be anti-TNFL, or Cox inhibitors, or whatever it is, that that approach is in fact not the correct approach because it interferes with some of the repair processes that have to occur. And what's more, you often don't know where the patient is in the inflammatory response at any one time.

So I think the problems with that TNF-alpha sepsis study could be that there was such a heterogeneity of the patients at where they were in their SIRS or non-SIRS. But the field now has gone to inflammation resolution rather than

Page 266

anti-inflammatory, and that's a big difference because what you're saying is we're going to trigger a new response in the patient or we're going to enhance the existing responses in the patient's recovery from that inflammatory process.

DR. COURSIN: A lot of food for thought. I'm getting a signal from the boss that we are at a break time, and we will have to get to the question I have about controls later. But thank you, everyone.

(Applause.)

(Whereupon, at 2:37 p.m., a recess was taken.)

DR. WARD: The last session, for lack of a better term, will be kind of a deeper dive into the clinical trial design, both for drugs but also for protocols. That's why I wanted Leanne to participate, because it's not just about trials for new drugs. Protocols in the ICU are an important part of improving care. It may not be something that ends up at the FDA, but it is something that is important to having a repertoire.

Page 267

We'll start out with Dan.

Presentation - Daniel Sessler

DR. SESSLER: My assignment was to talk about protocol design or trial design. Of course most of you do trials, so my challenge was to think of something that wasn't completely obvious to everyone in the room.

What I'd like to talk about is five major trends in clinical trials. One of them is towards large size, and this is a recognition that small studies give fragile results that often prove to be wrong. The second is towards composite outcomes rather than having a single outcome, and there are two reasons for this. One is that it reduces sample size, and perhaps a better reason is that a composite can better characterize the totality of an intervention's effect.

Third is factorial design, which is an efficient way to do studies and allows you sometimes to do two or even three things at the same time at very little additional cost. Then adoptive designs, which are essentially ways to

Page 268

incorporate information that becomes available during this study, either externally or from the trial itself, into the trial design, and therefore to make sure that the trial fully addresses all available information rather than following a protocol that might have been designed years ago. Then finally, I want to talk a little bit about novel trial designs that require altered or waived consent.

Let's start with large trials. How big a trial is really matters. I'm going to give you two examples here. These are only slightly disguised real studies. They were both published in the New England Journal of Medicine granted 20 years apart, and these were both studies of interventions to reduce myocardial infarction after non-cardiac surgery.

The first study had 200 patients. There was one infarction in the treatment group, 9 in the placebo group that gave a relative risk of 0.11, and the p-value was 0.02. The second trial had 4,000 patients. There were 200 events in the

Page 269

treatment group; 250 events in the control group for a relative risk of 0.8. The p-value was exactly the same, 0.02.

Now, which do you believe? Well, of course you believe the second one, and intuitively you think this makes more sense, first because everyone believes in the law of large numbers, but also keep in mind that a relative risk reduction of 90 percent is biologically implausible. There's not conceivably any single intervention that reduces the risk of something as complicated as a heart attack by 90 percent. It's just an unbelievable result.

The first result is fragile; the second is not. And what I mean by fragile is that if you add a couple of positive outcomes to the treatment to group, does it change anything? Well, in the first study, if you add 2 outcomes to the treatment group, the result is no longer statistically significant. You add to 2 outcomes to the second study, it doesn't change the p-value out to about the fifth decimal; it has no effect whatsoever. So

Page 270

the first one is fragile; the second is robust.

Let me put this another way. Sticking with something like a heart attack, heart attacks after non-cardiac surgery in patients over 45 has something of a 10 percent incidence; they're surprisingly common. You don't know about this because they're mostly silent, but they happen.

So let's consider an intervention that reduces the risk by 50 percent, reduces the risk by a factor of 2; the relative risk is 0.5 that's set here. This simply shows the 95 percent confidence intervals as a function of trial size. These are all statistically significant results.

In the first lowest one, N equals 500. This is a statistically significant result. 500 is a large trial. I suspect there are not many people in this room who have done a 500-patient trial. Yet, the confidence intervals range from about 0.25, which is a factor of 4 reduction -- this is biologically implausible -- to almost 1, which is no effect whatsoever.

This trial while statistically significant

Page 271

has not provided a lot of guidance to clinicians. To shrink those confidence intervals to a range that gives clinicians good guidance, you need to increase sample size by a factor of 10. You need to go to N equals 5,000, and that's why trials of myocardial injury are 5 to 10,000 patients these days. So you need to have very large studies.

Almost everyone believes that a p-value of 0.05 means that there is a 95 percent chance of replicating the study. That is not at all what it means. What it means is that there is only a 5 percent chance that by pure random motion, you've got the observed distribution of values. It doesn't directly tell you about replication.

So let's talk about replication. Let's say we're testing a drug that is completely ineffective. This is essentially placebo versus another placebo. You expect to have no treatment effect. They're both placebos. I'm giving you that. The relative risk should be zero -- should be 1 or the treatment at absolute risk should be zero.

Page 272

So I'm giving you that, and then you go repeat the trial. But if you repeat the trial, you're not going to get exactly the same result each time. You're going to get things around a zero treatment effect but not exactly treatment effect. In fact, if you repeat this thousands of times, what you will get is a normal Gaussian distribution. It's going to look like that.

Equal to 0.05 means that the distribution is in the most extreme, 2.5 percent on each end, because you don't know in which direction you're going to go.

Now let's change the paradigm, so now I'm giving you an effective drug, you do a trial, and the p-value turns out to be 0.05. What does that tell us about replication? Well, if you start repeating this study, you will on average get the effect that you got the first time. That's now your best estimate of the treatment effect. But of course you won't exactly get that every time. You will again get a normal distribution around that.

Okay. Well, let's look at that then. Half

Page 273

of the values will be more extreme, that is the p-value will be smaller, and you will consider those to be replications. But half the time, you will have less extreme values and a larger p-value. So a p-value of 0.05 means that you have a 50 percent chance of replicating the study. That is a coin flip. That's not actually very helpful.

A reasonable question then is how extreme a p-value do you need to actually have a 95 percent chance of replicating the study? You get that answer by sliding this bottom curve to the right until only 5 percent is less than your original observation. Then what you do is you take the peak of that and you trace it back up to your original, and you read off the p-value. It turns out to be p is equal to 0.0003. It's really small.

So why on earth do we use a p-value of 0.05 as our criteria for significance? It's a mistake of history. It came from a misunderstanding of what p-values really mean. It never should have been the p-value. The p-value probably should have been 0.001, and if that were the p-value, it'd be a

Page 274

lot harder to get a positive result, it'd be a lot harder to publish papers, and our literature would not be crammed with rubbish the way it currently is.

(Laughter.)

DR. SESSLER: Next, composite outcomes. Composite outcome is any group of outcomes; for example, a cardiac death, myocardial infarction, nonfatal cardiac arrest. These are usually used for dichotomous outcomes, and the reason people use them is that it allows a smaller sample size. The reason it allows a smaller sample size is that the number of patients you need for a study with dichotomous outcome depends mostly on the treatment effect -- but that of course is beyond your control -- and partially on the baseline incidence of whenever you're looking at.

So if you have a composite outcome and you're looking at lots of things, the incidence goes up. The incidence of a composite is always higher than the incidence of the components of a composite.

Page 275

Now, that's not actually the best reason to use a composite. The real reason to use a composite is that it better characterizes some intervention. Take for example a drug treatment for diabetes. It doesn't really make sense to say I'm going to do a study of an intervention for diabetes, and I'm going to make blindness my primary outcome and amputation secondary, and renal disease tertiary.

These are all important outcomes, and anybody who had diabetes would be interested in all of them. This is a perfect example of when it makes sense to have a composite of blindness, amputation, renal disease, and heart attack, the four major things maybe that diabetics worry about because it characterizes the disease well.

Now, one thing you have to be careful of with composite outcomes is heterogeneous results. A perfect example of this was the original POISE trial of beta blockers, which had a composite of myocardial infarction, and stroke, and death. Well, myocardial infarction went down with beta

Page 276

blockers, significantly; stroke went up. So the two components of the composite were going in opposite directions.

When you have that, you have an interaction term, and it doesn't make any sense to average them together. It doesn't make any sense at all to average an increase in stroke with a decrease in myocardial infarction. So when you have that, you have to split it apart. And the trouble is that most trials are not powered at that point because one of the reasons you used the composite was to reduce your power. So if you have heterogeneous results, it's very likely that you'll end up with an underpowered trial.

The most common way of dealing with a composite is a so-called collapsed composite, which is a fancy way of saying all are none; that is if any one component is positive, one or more components is positive, you say the composite is positive. If you take that approach -- and it's by far and away the most common approach -- then there are two rules you have to worry about.

Page 277

The first is that the incidence of each component has to be at least roughly comparable, because if you have one component that, say, is 10 times as common as all the others combined, effectively that becomes your outcome. That's all you're looking at, so you can't do that. The second thing is that the severity of the components has to be roughly comparable. So it does not make sense to have a composite of, say, sternal wound infection, abdominal abscess, wound dehiscence, and urinary tract infection.

You see this all the time. This has been published lots and lots of times, but it makes no sense. Urinary tract infections are 10 times as common as the others, and they're about a hundred times less serious. That essentially is saying a urinary tract infection is the outcome, but that's not what people care about, so that's a bad composite.

Now, you don't have to use a collapsed composite or an all or nothing composite. You can evaluate the number of components that are

Page 278

positive. It's not a common approach but it's one that you can use. A better approach, at least to take care of different incidents, is to use something called the average relative effect, which was popularized by our statistician in Cleveland at MASHA [ph], and that's a way of looking at the average effect of each component independent of incidence.

You can also weight the components. So if you have some components that are far more serious than others, you can essentially clinically weight them and say, I'm going to conclude urinary tract infections, but I'm going to count them as 100th of a deep sternal wound infection because I don't think it's very serious.

Third trend is towards factorial randomization. Factorial studies are really powerful because they allow you to evaluate two or more outcomes with only slightly more effort and patients than you would have for a single one. It also allows you to evaluate the interactions between different interventions.

Page 279

Suppose you're looking at two different sedatives. You would like to know if each sedative is effective, but suppose you show that each sedative is affective? Any reasonable clinician would turn around and say, "Okay, what about if I combine them? Do I get better efficacy with less toxicity?"

Well, let's say you did a 500-patient trial of one sedative, and it shows efficacy and not too much toxicity, then you do a 500-patient trial of the second sedative; again, efficacy and not too many complications. The clinician asks you, what if I combine them? Do you have any information?

You have no information whatsoever because these are separate trials, but suppose instead you had done a factorial trial where patients were randomized to the first sedative, the second sedative, to the combination of the two sedatives, or to nothing? Then you could evaluate independently what each one does and what the combination does. If you have enough patients, you can evaluate the type of interaction; specifically,

Page 280

are the effects additive, are they synergistic, or are they antagonistic?

Now, fair warning; you need a lot more patients to evaluate the interaction term, about 4 times as many. But if you're just looking at the marginal effects, that is a one drug, second drug, and the two combined, you can do that with an increase in sample size of only about 10 percent, so it's very efficient, and we're seeing more and more of these.

Let me just very quickly show you how this works. The example is the POISE 2 trial. In this trial, we randomized 10,000 patients to clonidine or placebo and to aspirin or placebo. Now, suppose you want to evaluate the clonidine effects. You get to the end of the trial and say, okay, what did clonidine do? Well, the most obvious thing would be to evaluate clonidine plus placebo aspirin versus placebo-placebo. These are drugs, patients who only got clonidine or only got placebo.

The trouble with that is that you can only use half of the patients, so this is 2500 patients

Page 281

in each group. But in fact there's absolutely nothing wrong with looking at the clonidine plus aspirin versus placebo plus aspirin. Aspirin drops out of the equation here. It's like being over 60. It just drops out of the equation. And by definition, by the way it's randomized, you have exactly the same number of people with aspirin in each group. So in fact, you can do your analysis across all clonidine patients and all placebo patients, 5,000 of each.

Exactly the same thing applies for aspirin. Again, the most logical thing would be to do aspirin plus placebo, but there's absolutely nothing wrong with doing aspirin plus clonidine or aspirin and placebo, and that allows you then to look at clonidine plus placebo -- aspirin with or without clonidine versus placebo with or without clonidine. You don't care about the clonidine; it drops out. It's a baseline factor. So you can use all 5,000 patients for your analysis.

The trial with the most factors that I know of was Christian Apfel's study of PONV. In this

Page 282

trial, we actually had 6 different factors, but I'll present just three of them here, the three drug antiemetics.

This is an example of how you can study the interactions. On the top, you have the amount of nausea and vomiting with no intervention, and then you have the effect of any one intervention, any one antiemetic, and it turns out that they all provide a 25 percent risk reduction. But then you can go on and look at the combinations. You can look at all three combinations of the antiemetics, and again, we had almost exactly a 25 percent risk reduction, and then you can look at all three, and again, it's a 25 percent risk reduction from the previous condition.

So large factorial randomized trials are powerful, not only because you can look at multiple things simultaneously without much increasing sample size, but you can look at the interactions and determine whether they are additive, antagonistic, or synergistic.

DR. SHAFER: What was on the left?

Page 283

FEMALE VOICE: I was going to say, why Marilyn Monroe?

DR. SESSLER: Oh, yeah. The risk factors for nausea and vomiting are female gender, opioids, nonsmoking, and a history of motion sickness.

(Laughter.)

MALE VOICE: She under-fits in that picture.

DR. SESSLER: Next up is adoptive designs. Adoptive designs are relatively new, and there's been a shift in thinking. Until fairly recently, the thought was that you should design a protocol and it was essentially written in stone. You registered the protocol, and even if the trial took 9 years -- and I would hate to tell you how many of our studies have taken 9 years -- you couldn't change anything. You had keep everything exactly the same.

There is now increasing recognition that things happen during trials. Things could be external, for example, other people publish relevant work. Maybe somebody else publishes a trial that's almost identical to yours, or it's

Page 284

similar to yours in a different population, and they get some answers, and the answers might be about efficacy, but they might be about toxicity also, and it might be about toxicity in a specific population.

Well, if you now know that a certain subset of the population of your trial is especially sensitive and especially likely to have complication, it would be unethical to keep enrolling them, so you have to make changes.

But similarly, suppose you know that a certain subset is more likely to benefit? You might well say, okay, I did start with something different five years ago, but now I know more. Now I'm going to change my trial to target a group that seems to especially benefit from whatever intervention I'm evaluating.

So you could alter the study population. You could restrict enrollment, or perhaps broaden enrollment, or somehow change the enrollment criteria to enrich the population for efficacy and reduce the risk of complications. You can also do

Page 285

things like adoptive randomization. You can change the treatment ratio. You could give more people the drug; fewer people placebo. But if you're testing two different drugs, you also could say, I'm going to focus on the drug that's looking best, and it might be data for internal for your trial.

From an interim analysis, you can say, okay, one of these treatments seems to be far better than the other one. I'm going to play the winner, and that might be just dropping one of them, but it might also be saying I'm going 2 to 1 randomization. So instead of having 1 to 1 to 1, you might have 2 to 1 to 1 type of randomization.

An example of adoptive design that's common in anesthesia is the Dixon up-and-down method for determining volatile anesthetic potency. The way those studies are done is that you start with some essentially random dose. You give it to the first patient, and at skin incision, you see whether the patient moves or not. The movement is unconscious. It doesn't hurt the patient, although it looks spectacular. If the patient moves, then you

Page 286

increase the concentration. If the patient doesn't move, you decrease the concentration.

So it doesn't matter whether you started too high or too low, you very quickly move down to about the average anesthetic potency and then you start bouncing around there. This is classic adoptive design, and it's been used in anesthesia for a half century, but we're beginning to use it in other contexts as well.

Another thing you might change is sample size. When you start a trial, you do a sample size estimate, and you use best available information in estimating sample size. But the most important contributor to sample size is treatment effect, which of course you don't know because the whole point of the study is to determine the treatment effect.

Very often you're wrong, and mostly people are overly optimistic about guessing what the treatment effect is or they adjust the treatment effect to get a sample size that they can do before the end of their fellowship or what have you. The

Page 287

trouble is that biology doesn't care. Treatment effect is whatever the treatment effect is going to be, so it's not uncommon to get most of the way through the trial, and it's absolutely obvious that your trial is underpowered.

It is not really very logical to sort of slavishly go ahead and say, okay, well I said I was going to study 239 patients; that's what I'm going to do. There's a certain logic in getting to 150 patients, picking what data you have, re-estimating sample size, and saying I'm going to go to 325 patients, which is what I'm actually going to need to make a reasonable conclusion.

Now, of course it has to be transparent and you have to disclose this. Ideally, your protocol would have this in the statistical plan. So right from the beginning you would say we are going to do interim analyses. We will re-estimate the sample size as necessary and increase treatment effect, and should be done somewhat independently from the investigators. We always do this on a group A/group B basis. We do it without knowing which

Page 288

group is which.

Then finally, you can change the drug or the drug dose. It might be that you're halfway through a trial and obviously you picked the wrong dose. You're giving half as much of this drug as you should, or you're getting complications and lots of efficacy, and maybe you should use half the dose that you started with; or maybe it's just the wrong drug. The drug isn't working; pick something else. That could be in the context of the same trial, which would be sort of a platform type of design, or you could kill that trial and start a new one.

Finally, novel designs. I think everybody knows about cluster randomization and a randomized step-wedge, which is a type of cluster design. There's a new type of trial. The first one is actually being done now, which is called opt-out in the routine care design.

If there's a clear standard for routine care, you can enroll those patients in a trial but not actually get consent because you're not doing anything different to those patients. They are

Page 289

getting routine care. You only get consent in patients who are randomized to the experimental treatment. The danger of course is that some won't consent, and they may consent non-randomly and with bias.

The final type of novel design, which I think we developed, so I'm fond of this, is an alternating cohort study. This is like a clustered trial, except that the clusters instead of being randomized in space are randomized in time. And basically what you do is you do some treatment for a period of time, like 2 weeks, and then you switch to the alternate treatment, and then you switch back again, and you keep doing this for, say, a year.

Since there's no reason that patients would be in any particular 2-week block, it is a controlled trial; you're controlling the exposure. Even though the exposure periods are not randomized and certainly the individuals are not randomized, it's a trial design that's easy to implement. It's inexpensive. It allows you to enroll very large

Page 290

numbers of patients. We've done a bunch of these now with thousands of patients. It costs almost nothing, and they have a lot of the protections of a randomized trial at a tiny fraction of the cost.

I'm going to skip the rest of this, so thank you much.

(Applause.)

DR. WARD: We've been talking a lot about drugs, but protocols are very important. So Leanne's going to fill us in a little bit more on design for protocols.

Presentation - Leanne Aitken

DR. AITKEN: Thank you.

Yes, so my thought is very much that we have spent all this time talking about drugs, and absolutely we need to find the right drugs, but we need to look at how we're giving the drugs because the best drug in world, if we're giving it in the wrong way, we're not going to achieve the outcome that we want to achieve.

Largely, what I was asked to do was to talk about some of my experience in doing predominantly

Page 291

these Cochrane reviews, although I have done a couple of studies in the same area, so I'm obviously informed by that, and I'm informed by some of the more recent work that I'm doing in looking at some of this sedation as well, and I'll bring that in later in the time.

Bearing in mind that the first of these Cochrane reviews was done six or seven years ago, so the protocol was written eight years ago. And I look at it now and think I'd write it very differently now to what we did back then. We just did the revision, which was published last year.

My learning from that is that if you end up in the situation where most of your studies are individual patient randomized studies, and then there's one cluster randomization trial that needs to be included, run as far as you can. Don't hang around or pay a statistician a large amount of money because it becomes a nightmare when you've got one cluster randomized study to go in the review, which was the situation in this case.

When I'm talking about protocols directed

Page 292

sedation, what I'm talking about is where the sedation has been ordered by a physician and is implemented by nurses, pharmacists, or others. That was our provision, but the reality is all of the sedation protocols are implemented by nurses in the review that we've included.

The protocol should contain information on the sedative agent or agents to use, and when to commence increase, decrease, or cease sedative agents. It should be in some way based on patient assessment, and it might include other interventions such as daily sedation interruption.

It's similar to but distinct from a weaning protocol, so there are other studies that look very specifically at weaning protocols that are not included in here. The likely mechanism for improvement of a sedation protocol is simply through reducing the individual variations, so getting people to work more consistently towards a target.

This is the bit that I now look at and think, yes, I'd probably choose some different

Page 293

outcomes if we were starting afresh at this point, but these are the outcomes that we identified about eight years ago based on what was available in the literature at that point and where our thinking was at that point. So some of them are still not consistently in the literature, but this was the drain list at that point, where the primary outcomes were either duration of mechanical ventilation or mortality, either within the ICU or within the hospital.

The secondary outcomes -- and I've got no idea why that's appearing in both. Oh, no, that's length of stay, not mortality; sorry, I'm reading wrongly. The secondary outcomes were length of stay, total dose of sedation, adverse events within the ICU, incidence of delirium, incidence of tracheostomy, some post-hospital outcomes along the lines of memory, psychological, or cognitive function, and quality of life. And I'll talk just a little bit about how often we found those outcomes in the studies.

In the review that we published last year,

Page 294

we included four studies, and in those four studies were a total of just over 3,000 patients. The study that bumped up the numbers, because that's a fair size patient number for four studies, was the pediatric cluster randomized protocol study that Martha Curley led that was published about three years or so ago, so that is a big study sitting in the middle of this.

But you can see that all of the studies had measured duration of mechanical ventilation in some form, and I'll talk about that in a moment. Two of the studies had ICU mortality; three had hospital mortality. All of them had ICU length of stay; three of them had hospital length of stay. Two had self-extubation and one had reintubation. And obviously, they're getting at the same concept but are slightly different. And then one had traecheostomies in there.

When I said duration of mechanical ventilation, one of the challenges that we had to deal with was that in the various studies -- and there were only four, but duration of mechanical

Page 295

ventilation was defined either as duration of mechanical ventilation, or time to extubation, or ventilator-free days in the first 28 days, and that obviously created a huge problem for us.

Now, fortunately we were able to get from the authors some consistent data that we could then do a meta-analysis, but it wasn't necessarily the format that was published in the study in the first place. So I think we do need to think about what's the right version.

To this point in time, there's been no sedation protocols that have studied, that have looked at, total dose of sedation or any of the risks of that list that's there. Obviously, those outcomes have been measured in lots of other studies, but not many studies that's been comparing different versions of sedation protocols. It's worth reminding you that this is a Cochrane review, so it was only RCTs. There are some other observational studies that do have some of these in there but not much.

Total dose of sedation is an interesting one

Page 296

that I'm not sure I would put in there now, and I'm not sure it's of value. I think we need to think more carefully about that, and I've got some notes for the [indiscernible]. The other thing that I've said there is these four included studies were conducted or published back in 1999 and then more recently than that. The 1999 one was -- his name's just gone.

MALE VOICE: Brook.

DR. AITKEN: Brook. Thank you. I was going to say the wrong name. It was Brook, and that's a mid-nineties study when it was designed, so we have moved on quite some distance of time since then.

Now, as you can see -- and it doesn't matter that you can't see the detail particularly well, but what we've done there is just included a few different studies, the top three studies of individual patient randomized studies and then the fourth study is the cluster randomized study. But you can see with the individual study results that there really is quite a lot of variation between the studies in terms of some of them providing

Page 297

absolutely no benefit and if anything harm, whereas other studies are a long way on the benefit side.

This was the original Brook study. Now, particularly given my background is the one study that does go on the harm side, is the study done by Trace Bucknell in Australia, and we have a setting that is very well known for having 1 to 1 nursing at every bedside, having 70 to 80 percent of our nursing staff with post-graduate qualifications in critical care, probably a different environment to the other three studies that are done in the North American setting. So it raises the question a lot about context, which I'll speak about in a moment; so certainly inconsistent results across those contexts.

Some of the factors that we think affect this are things like what's the usual practice; how much implementation was there of the intervention? In other words, it's all very well and good that we've set out what the intervention is meant to consist of, but was that actually achieved. And as I said, what were the staffing types and levels.

Page 298

This is where I've become conscious of the lesson today that my use of language is probably heavily influenced by the UK environment now, rather than something that necessarily is used internationally as language. But we've spent a lot of time thinking about process measures or process evaluation.

Earlier today, I mentioned that I'm a co-app on an RCT for dexmedetomidine versus clonidine versus usual care. My role in that is to lead the work strain for process evaluation. So even in an RCT of a drug, we have a whole work strain that's looking at how are we implementing this drug, how are we actually achieving what we think we're achieving? I guess on reflection, I realized that that's very UK oriented language in thinking about process evaluation, but it's in essence how well implemented was the intervention.

So I don't think of things like total dose of sedation as being an outcome measure. I think of it as being a process measure. I don't really think of a percentage of time target sedation as

Page 299

being an outcome measure. I think of that as a process measure now. So particularly for something that is a behavioral intervention like the sedation protocols are, I think it's vital that we have some detail process measures about what the context is and what the intervention fidelity is; in other words, how well implemented was the intervention.

What was the dose of sedation that we achieved? One of the things that we've noticed in looking at some other work around depth of sedation is that there's no agreement on how we should be measuring depth of sedation. So is it the average daily dose of various drugs? Is it what sedation measure was achieved? Or is it some sort of calculated measure? And there are a couple of variations on sedation index that you can find.

I'm not sure at something like percentage of time at sedation target because achieving a sedation target of a RASS of minus 4/minus 5 versus achieving a sedation target of zero to minus 1, both of them are completely achieving the target but very different sedation states. So I'm not

Page 300

sure it tells us much about the depth or the dose. Certainly, talking about coverage or rate, how many of our patients got the intervention that was intended? Did we get to all of our patients, and did we get to them in a timely manner?

Just recently, Lydia Emerson, she's about two seconds of finishing her PhD, but she's developed a model for process evaluation in critical care studies, including RCTs of drugs, but critical care studies more broadly. I know that's a bit difficult to see from the size, but she's talked about there being elements that you need to look at during the baseline period of the study, the exploration period, and then during the study. And then to clarify at the end of the trial with the thought being that these data will help us better implement the study as we go, but perhaps more importantly, help us to explain the results at the end of the study.

The elements included in her model are context, attitudes and perceptions, fidelity, dose, reach, recruitment, and then level of

Page 301

implementation. In one of the recent ICU studies that's just been finished in the UK, which was the POPPI study, which was a nurse-delivered psychological intervention within the ICU.

They applied this model to that, and on first analysis, which is all that's available at this stage, it looks like those sites that had a higher level of implementation had more effective benefit, even though the study as a whole didn't find benefit on the straight RCT. So they're going to do some more analysis to see if that measure of implementation is valuable. We're applying it to the A to B dexmedetomidine versus clonidine study to see if that can help us there. So I raise that as many of the elements that particularly in a behavioral intervention like a sedation protocol I think is absolutely essential.

My thoughts in moving for forward -- I've raised a lot of the questions as I've gone through, but I think in thinking about the patient-centered outcomes, that we need to be obviously thinking those that are ICU focus but then those that are

Page 302

hospital focused and those that are long term. And in sedation studies, we're going to be thinking across all of those.

My strong emphasis is that whatever outcomes we have, we also need what I've referred to as process measures to help us explain the variation in outcomes that we get to at the end. That's interesting that Lydia said in one of the ventilation studies, where she was leading the process evaluation, most of the co-apps on the study couldn't work out what the process evaluation was all about and couldn't really see the benefit until they got to the end and got no difference in the statistical analysis and said, "Oh, now we need to look at the process evaluation" and work out what was going on. So it wasn't quite the right way around, but that's certainly what she's found in getting to the point where those measures became important. So I'll leave it there.

(Applause.)

DR. WARD: Before the panel and we get to ask all the questions, my reading of the literature

Page 303

is one of controversial statistical measures that people use, particularly when you get to things like composite outcomes. Hopefully, they're going to enlighten us.

Presentation - Elizabeth Colantuoni

DR. COLANTUONI: I hope so. Do you guys want to stretch, a 4 o'clock stretch before the statistics talk? Highly recommended. Feel free to stand while I'm talking.

I should just start by saying that sedation trials is somewhat out of my wheelhouse. I've been involved much more with long-term observational studies and randomized trials within ARDS populations, and now getting a little bit more into trial setting within the context of delirium. Leanne gave such a nice summary of the literature. I was reading up into the published trials, so I'm going to highlight some of the outcomes that she just mentioned.

But here's just a schematic of the standard design. Intubated and mechanically ventilated patients are enrolled and randomized to receive one

Page 304

of two pharmacologic agents representing sedatives and then administered those drugs through extubation, and typically followed through ICU discharge and perhaps through hospital discharge, at least accumulating length of stay.

The whole time that the patients, then in the ICU and moving through hospital discharge, death is a potential competing risk. In my reading of the literature in these sedation trials, it looks like death is 30-day mortality, ranging from anywhere from 15 to 30 percent, so a pretty high rate of mortality in these populations.

Identified endpoints from my quick look -- and many of these just popped up on in the prior presentation -- is that primary and secondary endpoints are highly variable. They range from proportion of time; reaching the sedation target and goal; duration of mechanical ventilation; ICU and hospital length of stay; and mortality and delirium. But there's a lot of inconsistency even in just primary endpoint definition across trials, let alone a wide range of variation in secondary

Page 305

endpoints.
Today I'm going to talk about how to operationalize delirium as an endpoint within this setting, so that will be the first part of the talk. Secondly, in reviewing some of the protocols and ongoing trials, you see some additional duration of follow-up in the sedation trials, maybe perhaps extending to 3 months or 6 months post-randomization, where we're looking at longer term mortality, but we're also starting to measure functional outcomes similar to what Dale described earlier today.
These could be measures of physical function, either self-reported measures of physical function or actual, like hand-grip strength -- those sorts of things could be included here -- mental illness or mental health measures, and then quality of life.
So my talk is going to talk a little bit about how we operationalize delirium as an endpoint and the statistical challenges there, and then separately I'm going to talk about the challenges

Page 306

in evaluating these longer term functional outcomes, particularly within the context of this competing risk of death.
I want to highlight here before I move on, the competing risk of death is not just affecting delirium and these longer term functional outcomes. Our evaluation of duration of mechanical ventilation, ICU, and hospital length of stay are also endpoints for which mortality has to be considered.
This paper that I'm highlighting here is a paper from a bunch of colleagues at the School of Public Health at Hopkins. It's just a nice review of the differences in the statistical methodology available to compare relative hazards versus relative risks when there's a competing risk of death. I find myself going back over and over again to this manuscript to remind me of all the definitions.
Delirium as an endpoint, up until a few years ago, my primary exposure with delirium was thinking about delirium as an exposure and

Page 307

correlating delirium with subsequent outcomes in patients. Dale approached me a few years ago and said I need you to write a statistical analysis plan. The endpoint is delirium, and I had no idea what to do with proposal. So we were evaluating an ancillary study to the SAILS trial, which was a multicenter randomized trial evaluating the use of rosuvastatin versus placebo, looking at patient mortality and duration of mechanical ventilation in patients with sepsis-associated ARDS.
The data we had was an ancillary study, so within a small number of sites. Delirium was measured daily up to death, ICU discharge, or 28 days. Our goal was to try to operationalize delirium as an endpoint, and then make a comparison between delirium as an endpoint across the two treatment groups.
I'm going to walk through my thinking around developing this statistical analysis plan. We utilized statistical approach that was different than what was the predominant approach in the literature at the time. Our paper appeared, the

Page 308

actual analysis appeared in Lancet Respiratory Medicine in 2016, in January, and then there was a subsequent series of commentaries, for which I responded to one where I just had a highlight of some of the statistical challenges.
Delirium, as many of you in this room are experts in delirium, so talk correct me where I get off course. But this is a state that's in constant flux of change. Your delirium outcome can change over the course of hours or days. Here I have hypothetical patient. Time zero is enrollment, randomization, and then we're following the patient for 28 days. The zeros and 1's above the time scale are just indicators of when the patient was evaluated and whether they were observed to be in a delirious state versus not.
I highlight here at the bottom that this kind of change over time also applies to sedation -- that's most interesting to most of you in the audience -- with the potential for maybe greater variation and more rapid changes over time.
Second, delirium occurs along a continuum of

Page 309

severity, and you cannot assess delirium when a patient is severely impaired. When a patient is comatose, we're not able to do a delirium assessment. For this particular patient, we see the first 2 days, the patient is comatose and unable to be assessed for delirium. Once the patient is not in a comatose state anymore, we have 0-1 indicators for their delirium state, so that's a challenge.

Third, delirium evaluation is often stopped when patients are transferred out of the ICU, so stepping down from the ICU to the hospital ward, but delirium may persist. Some of the data that we have available when patients are evaluated during the last day of their ICU stay, anywhere from 15 to about 50 percent of the patients are positive for delirium at that time. So how do we treat delirium as an endpoint where we're only observing it, a half of it or a potential small portion of the delirium process? And lastly, death, death is a common occurrence in these ICU studies. The whole delirium process is truncated once the patient

Page 310

dies.

The approach that had been taken in the literature and continues to be used as calculation of delirium-free days to X days. This statistic or composite is based on ventilator-free days to X days variable that's used commonly in studies of mechanical ventilation. This composite endpoint is composed by assigning zero to patients that die prior to day X. Among survivors through day X, you count up the number of days where the patient is off the ventilator; take that composite variable, and you compare it across treatment groups typically using a rank-based test and/or present prespecified quantiles.

Over the years since this was proposed in 2002 by David Schoenfeld and others, there have been a lot of publications trying to identify and just bring to attention that there are some challenges with using this endpoint. Recently, a French group from Inserm last year published another paper on some of the drawbacks to using this as an endpoint.

Page 311

How has this endpoint been translated into sedation trials? Well, first thing is how do we define X. Just in my reading of sedation trial literature, there's quite a bit of variation in how we're defining X. It's 7 days, 12 days, 28 days. Ideally, you want X to be specified such that the vast majority of the patients would either have died or have been extubated prior to your time point. That would be a target to try and figure out how to set X.

How do you deal with coma days? You can change the endpoint from delirium-free days, to X days, to coma and delirium-free days to include coma within the continuum of the delirium process. I'm sure there would be a heated argument here about whether that's part of the process or not. In the ABC trial, they counted days of CAM-ICU positive but when non-comatose. So there are alternative ways to treat coma.

In death, do we set delirium-free days to zero if a patient dies? In the protocol for the SPICE 3 trial I was reading, they're counting the

Page 312

days free of delirium prior to death as part of the composite, so there is another twist to the variable definition. But universally, most when we're defining this as an endpoint, almost everyone assumes that once the patient leaves the ICU that they're delirium free.

As an alternative approach, we're going to suggest that you can directly model both the delirium and the competing event process by using a joint model sometimes referred to as a shared frailty model in statistics. In the first model, you would build a survival model for being positive or absent of delirium on any given day. This is like a recurrent event survival model. The second model is a survival model for your competing events, ICU discharge or death.

The two models are linked by a random effect or what's referred to as a frailty term in the survival analysis literature. The frailty term appears in the first model as a way that we can link the repeated daily observations of delirium within a person over time, and then that frailty

Page 313

term appears in the second model as a way to link the risk of delirium with the risk of the competing event.

How we apply these models is that we allow the coma days to be days for which the patients were not at risk of delirium. Within the recurrent event model 1 there, patients were only in the denominator of that survival analysis when they were comatose free. The treatment effect is estimated by having a main term for treatment, and the recurrent event survival model in that term can be interpreted as on any non-comatose day in the ICU, the relative hazard of delirium comparing the treatment to the control group.

How all these analyses played out in the SAILS trial ended up not mattering, really, how we evaluated the endpoints, so we compared ever and never delirious across the treatment groups, days alive without delirium and coma where essentially the number of days were identical -- the median number of days were identical across the two arms, and from the joint model, we estimated a hazard

Page 314

ratio of 1.4, but our confidence interval's quite wide. Here we would say on any non-comatose day in the ICU, the hazard of delirium is 14 percent greater for patients receiving rosuvastatin compared to placebo.

There are many challenges within this setting. If you're going to go with a composite endpoint approach, there needs to be a consistent definition applied across the trials, both with the duration of follow-up, how you're going to account for death and coma in the ICU discharge.

If you're taking the joint modeling approach, there are limitations here as well. In the current implementation, the joint modeling approach only allows for a single model, a single competing risk, whereas we really have the competing risk of discharge and death, which are two separate processes and have two different relationships with delirium. So patients who have a higher risk of delirium are at higher risk of death, and patients with lower risk of delirium have higher risk of ICU discharge.

Page 315

There could be alternatives to both of these approaches that we haven't thought of. One thing that I didn't talk about along the way is the complications introduced by missing data, so missed delirium assessments on any given day add another layer of challenge.

I just started an NIA funded R01 that is specifically looking at delirium as an endpoint within preventative and therapeutic delirium RCTs. I'm going to be doing some systematic reviews of the methodology applied across delirium trials and then also a series of extensive simulation studies and try to identify where these endpoints can work and where they can't. Then there includes a whole aim for statistical methods development, so try to improve the joint model by allowing for separate models for the competing events, and hopefully make some good recommendations for use of these approaches.

Now I'm going to shift from thinking about delirium to talking about the functional outcomes. When I mean functional outcomes, I'm thinking of

Page 316

something that's not defined as a survival endpoint, something that you evaluate the patient and you get a measure of their physical function or their quality of life; so something that's a scaled or quantitative variable.

Everything I'm going to discuss here you can find in this BMJ paper. This was with Tim and Dale as co-authors. This was a culmination of the third aim of the R24 that Dale described earlier today. I'm going to have a little bit of mind games at 4:30 in the afternoon. I'm going to introduce this idea of potential outcomes to the group just as a way for us to organize our thinking around how we can identify the causal effect or identify a treatment effect.

First, I want you to imagine you're in a setting. Your goal is to evaluate 90-day cognitive function in patients, and there's no mortality. There are two interventions, an intervention and a control. Under the potential outcomes framework, you're imagining, or we organize our thinking to say, that any given patient would have a measure of

Page 317

cognitive function if they had received the intervention. Similarly, they would have a measure of cognitive function if they had received the control.

The individual causal effect is the difference, then, between those two potential outcomes of cognitive function, one under intervention and control, and the marginal or the average treatment effect is the average of all those individual causal effects over the population of interest.

How does this change when we have mortality as a complicating factor? Now I'm going to imagine, first, that I have potential mortality experiences in each two groups, so I'm going to imagine that I can know the time of death in days and an indicator of whether a patient survived to 90 days both under the intervention and the control arm.

In addition to knowing this information, then I can also start to categorize people into their potential survival experiences. Always

Page 318

survivors would be a subset of patients that would survive to 90 days regardless of the treatment they received. These are likely the most resilient patients in the trial. Mortality benefitters would be those that would survive under intervention but would experience death by 90 days if they received control, so these would be less resilient patients.

The always diers, these would be our pretty severe patients. These are patients that would experience mortality regardless of the treatment they received. And then there's this category called the specials. These would be patients that might die under the intervention but survive under the control group, and I'll talk a little bit more about these in a couple of slides.

Now we can think about when we actually get to observe cognitive function at 90 days based on these potential outcomes. I only have cognitive function declined if a patient would survive to 90 days. Once you die, the cognitive functions no longer are evaluable.

Here now I see that under intervention, I

Page 319

only get to see 90-day cognitive function for the always survivors and the mortality benefitters. Similarly, I only get to observe 90-day cognition under always survivors or specials. The only group of patients for which I can identify or even define an individual causal effect is the always survivors.

In the statistics literature, the survivor average causal effect, which is also known as the SACE, is the average of these individual causal effects but only among the very specialized subset of the population, and the specialized subset is those who would survive regardless of what intervention they received. You immediately think that there are some problems here because in practice, we don't get to observe those states; we only get to observe one. So in the end when we're analyzing the data, we don't know who's an always survivor or not.

There are ways in which we can estimate the SACE by making additional assumptions. We can get an upper and lower bound for the survivor average

Page 320

causal effect if we're willing to assume there are no specials, so that there would be no one who would survive under control but die under the intervention. If you want to get a point estimate for this causal effect, you have to make additional, more restrictive assumptions, and none of the assumptions are verifiable by any observed data that you have in the trial.

In practice, that survivor average causal effect is very rarely reported in the literature. What's more often reported is just the survivors only analysis. There you should just take all the survivors data, take the average of your cognitive function measure under intervention, and compare that to the average under your survivors in the control arm.

The only time in which the survivors only analysis reduces to an actual estimate of a causal effect is when the mortality is not different across the treatment groups. So if there's no mortality difference across the groups, there's no mortality benefitters or specials, so the survivors

Page 321

only analysis reduces to the actual causal effect.
The problem with the survivors only analysis is if there is a mortality benefit for the intervention, then you basically have a mixed bag of patients. Your survivors under intervention are always survivors and mortality benefitters whom could be inherently quite different from one another, and that can introduce a bias.
Both of these approaches are what would be referred to as conditional methods because they condition on a particular subset of the patient population in order to make a treatment comparison. They suffer from a disadvantage in terms of evaluating randomized trials that they don't satisfy the intention-to-treat principle. There are other advantages and disadvantages of them, but that's kind of a primary one.
What could we do as an alternative to these approaches where we might be able to utilize all the patients that were randomized? One approach is to utilize a composite endpoint. Most of the composite endpoint approaches require that we've

Page 322

ranked the patients in terms of severity. One example is a proposal by Lachin in 1999 that utilizes a ranking of patients that incorporates the timing of death, not just an indicator of when patients die, and then information about the scale of interest or the functional outcome.
Let's imagine that we all agree that earlier death is worse than later death, and remember, these are longer term outcomes, so it might make sense for us to be willing to compare survival 3 months post-randomization is worse than survival 180 days post-randomization. Also among survivors, poor functional outcomes are worse than good functional outcome.
Then we define a new variable -- I'm just calling it W -- which would be equal to the time of death for those who died prior to the time of interest, 90 days, and then is equal to the functional outcome plus some constant just to allow us to differentiate times of death from the functional endpoint.
So now we have this composite variable,

Page 323

which is just a variable that's happening on continuum with higher values indicating better function. It doesn't make sense here to compare the means across the treatment arms in this composite; it would be better to compare the distribution of a composite endpoint like this, so you could do like a rank-sum test or you could compete various quantiles from the distribution of this composite.
Just as an example, I just made these numbers up, if you targeted the median of this composite endpoint, you could compare the interventions like this. So you could say under the intervention, 50 percent of the patients receiving the intervention survived to 90 days with cognitive function scores that were less than 30, compared to under the control group, 50 percent of the patients had experienced death by 50 days. This is a useful metric as a way to rank experiences across the two intervention arms.
In terms of recommendations, when mortality is involved, there's no real solution that doesn't

Page 324

have a disadvantage. The approach that you choose is going to depend on the assumptions that you're willing to make within the context of the problem.
There are a couple of recommendations I would make. If it's biologically unlikely that the intervention is going to impact mortality, then you're safe with the survivors only analysis. The survivors of the intervention to a particular time will represent a random sample of the original randomized patients, so you should be fine.
When mortality is a primary endpoint, as it is in many of the trials that we do in critical illness, you're hypothesizing that there is a difference, so you should build into your statistical approaches the potential that there is a difference.
You should have some step-down approach or specification that any analyses of functional outcomes would consider mortality, so by using the composite endpoint approach and/or one of the causal inferential approach the SACE.
Here, this is the two parts of the talk that

Page 325

are most familiar to me, and then here are some other observations I made while I was reading through the sedation trial literature. It looks like there's limited use of group sequential designs within this setting. I found one trial, the NONSEDA trial that performed a single interim analysis after 350 patients were recruited. Choice to use a group sequential design depends on a lot of things, but mainly on your projected rate of recruitment and the duration of follow-up.

There also was no mention of utilizing baseline covariate adjustment. If you're collecting baseline variables that are prognostic for your outcome of interest, you can include those variables in your analysis to improve precision and to estimate your average treatment effect, so get this in power. There's a whole host of adaptive enrichment designs, which were alluded to in the prior talk, and then other novel designs.

One of which that came to mind today in our discussion particularly around how patients are changing rapidly over time, there are these micro

Page 326

randomization trials that are developed by Susan Murphy at Harvard beyond my scope, because it requires a lot of interesting optimization problems.

The idea is that patients our originally randomized to a treatment. If the patient responds to that treatment, they remain on the treatment, but if the patient doesn't, then they're re-randomized again to other different conditions. Then that happens sequentially in time until the patient ends up in an optimal treatment category.

Along the way of those micro randomized designs, there's constant assessment of the patients. So when the patient is identified to not be performing well under the current randomized treatment, you can do the randomization again to move the patient into a more optimal condition. There's also POP [ph] trials and pragmatic trials that have been used in other critical illness settings that might work in this setting as well, so that's all I have.

(Applause.)

Page 327

Q&A and Panel Discussion

DR. WARD: Before I have a couple of other people to join us, [inaudible - off mic] to get a deeper dive into how we should be designing those clinical trials.

DR. SHAFER: I'm going to be moderating this session, and I want to start off by just repeating something that Dr. Colantuoni said a second ago. Everybody stand up and stretch.

It turns out that -- we were talking about patient-centered outcomes. I'm actually in a category that hasn't been discussed so far. You've heard from my wife, Pamela Flood. I'm an ICU survivor-survivor --

(Laughter.)

DR. SHAFER: -- and when Pamela was hospitalized at UCSF, I spent literally every night sleeping there by her side. And then when she needed to have the care advanced to intubation and sedation, I quite fortunately advocated for this, and you won't be surprised to know that I advocated that she get propofol.

Page 328

She was on that for a couple of days, and then Mervyn Maze, the chair, said, "You know, what do you think about dex?" So Mervyn had a strong role in the suggestion that we move to dex, which of course since it was Mervyn's suggestion, we did. And I want to share that as an ICU survivor-survivor, it was very consequential because it was when I knew that she was going to make it.

She had been unresponsive on the propofol for about 3 days. We went to dex for about 8 hours, and Mervyn and Marty Bogetz came into the room. Pamela has been completely unresponsive this whole time -- and she knows what I'm going to say here -- and I said to her, "Pamela, two stud muffins have just walked into the room, Mervyn and Marty."

DR. FLOOD: Hoping it would wake me up.

DR. SHAFER: And Pamela goes [gestures], and that's when I knew she'd be coming back. But as a survivor and a survivor-survivor, it does continue to affect us. And one of the ways that I'm

Page 329

conscious of every day is that I have to take care of my -- I've never had to take care of somebody like that before. And I have to take care of myself so that I can take care of my wife. It changes one's perspective on these things.

We've had a wonderful discussion here, and I'd like to open this up for questions and thoughts about clinical trial design and some of the ways of moving this field forward.

DR. RIKER: Riker. Yahya, you haven't had a chance to really tell us much about what you've learned in your SPICE series of studies and what you would do today if you were designing SPICE 3. I'm eager to hear your thoughts as far as RCTs versus other alternatives or where you are.

DR. SHEHABI: Thanks, Rich. I will start by what Steven just alluded to about Pamela being unresponsive for 2 days, and then suddenly becoming awake and doing this [gestures]. I think this is really a part that will end very early in the SPICE program, that the first 2-3 days of the acute phase of critical illness is very different to the days

Page 330

from day 3, day 4, and onwards. You are kind of like in the eye of the storm in the first 2 days, and then the storm will pass, and you're now cleaning up.

I think clinical trials ought to accommodate for that, and perhaps we need clinical trials that tackle the early part of critical illness where it's very hot, very dynamic, and everything's happening, procedures, imaging, dialysis, to go to theatre and come back; all that stuff is happening and it's very different from when the dust has settled and we're now in a recovery phase.

I think trials so far has ignored that first 2 or 3 days mainly for logistic reasons because we could not consent people in time to get them into these studies. The only way you could do it with SPICE is to have a deferred consent where the patient would be randomized, and then once their legal surrogate becomes available or they wake up, then they will consent to continue part as a patient, or say, no, I don't like this. I want to get out.

Page 331

So that's one point. The second point, I think we had a paradigm that is essentially age independent in our trials. It's quite clear from what we've done within SPICE -- which we knew before but we didn't really have to know what's the impact of the trials -- is that the patient who's 75 years old is not like a patient who's 35 year old. They're both adults but very different adults. They have very pharmacokinetics, very different pharmacodynamics. They're sensitive to drugs differently. They have comorbidities. They're different, but regardless of that, we treated them both as the same patient.

I think we have to stratify going forward in clinical trials by age because we are definitely dealing with two different biological systems between a younger adult and an older adult.

The third point, which we've also realize, is that particularly early in the course of critical illness, clinicians use a combination of drugs. While we do go and study X versus Y, even in the guideline we say we're going to look at

Page 332

whether propofol is better than dex or dex is better than this. But in real life, clinicians use a combination of things. At one stage, they use propofol, then they move to dex, and then they add some midazolam. They add morphine. They add fentanyl.

That combination pharmacotherapy is what happens in real practice. For trials' conclusions and results to be generalizable, it needs to accommodate for that combination of usage.

DR. SHAFER: Other comments?

DR. SPIES: Maybe one additional, I full agree with Yahya. The point is I think one thing is vulnerability, so many patients have different -- so chronological age is difficult because usually people can be very frail when they go into that setting. For example, if they have cancer, prolonged cancer, they are much more frail to what we are doing. I think that's something that needs to be also considered, the physiological reserve of the patients.

DR. SHAFER: Anybody want to respond?

Page 333

DR. SHEHABI: If I could just add to your comment, Claudia, I think when we do clinical trials, having a large sample size would allow you to have adequate power to look into those different subgroups and make meaningful results from doing that. I think earlier Rich was talking about having mortality as a primary outcome. We use primary outcome primarily to sample size studies rather than find what it's going to show. We just want to know whether it's going to show different or not but primarily to sample size for a study.

I think if used mortality, for example, as a primary outcome, your sample size is with a large sample, but that allows you a lot about clinically relevant outcomes with a lot of power and a lot of precision.

DR. SESSLER: Absolutely I support large trials. If you know in advance that you're interested in a particular subgroup, consider stratifying so that you end up with a good balance across your groups of interest. It essentially cost nothing. With electronic randomization, you

Page 334

can add lots of stratification, and it will give you good balance for free.

DR. COLANTUONI: I agree.

DR. SHAFER: Talmage?

DR. EGAN: I don't want to derail the discussion too far afield, but at some point I think this is worth discussing, and I don't see that there's another point in the agenda where it has an obvious place to come up, and that is the question of using target-controlled infusions as part of the study design.

I think if we look at this very broadly, these kinds of studies are both pharmacodynamic studies and outcome trials, so on an hour-to-hour basis, it's a pharmacodynamic study. You're trying to make some assessments about where the depth of sedation is. And then you've got the sort of broader question of what the ultimate outcomes are, which is obviously the more important endpoint.

But in any case, at least for the pharmacodynamic part of this study, controlling the kinetic aspects of the study is so that you don't

Page 335

have drug levels jumping all over the place, so that you're giving the drug in a very precise way where you're getting some approximation of a known plasma level. And more importantly, you are locking in a relatively steady state of the drug, and I think improves the overall design to some degree.

As you'll recall, Steve, you and I collaborated on a trial that used TCI as part of the study design. So I just wondered our panel thinks and what some of the audience thinks about how that might improve these trials.

DR. SHAFER: Let me just follow up on that. Quite specifically, that was one of the registration trials for propofol.

DR. EGAN: Right.

DR. SHAFER: So propofol registration for the ICU was done using TCI, and without knowing your doses and your concentrations -- which is one of the other things TCI can do, is it can capture what you've actually done as well as allow you to target things, which otherwise is very hard to

Page 336

capture what drugs were used. You can take a trial and say, hey, drug A works better than B, but A is just 20 percent more propofol. You can't really identify it without actually getting the kinetic dynamic model involved in the outcome analysis.

Anybody want to comment on this?

DR. GIRARD: This is Tim Girard. I think that's a great idea. I think probably we need even back up further because the pharmacokinetics of most of these drugs is poorly understood, if not completely un-understood. That's not a real word, is it?

(Laughter.)

DR. GIRARD: I think you get my point. Many of these drugs have had very little, if not any pharmacokinetic studies, in this population. Our group has done some work looking at pharmacokinetics and pharmacodynamics, and found that actually plasma concentrations of many of these drugs did not correlate well, or at all, with the observed clinical response to the drugs.

DR. SHAFER: Mervyn, wasn't dex also done

Page 337

with TCI?
DR. MAZE: No, it wasn't done with TCI, but we had very specific infusion criteria.
Dan?
DR. SESSLER: Dose really matters, and it's something that we've a little bit ignored here. Talmage and Steve could speak to this better than I can, but we're often comparing two different drugs at essentially random doses, but if you use slightly different doses, you could get completely different results. It's very easy to do a trial with two drugs and conclude one is better than the other. It's not actually better than the other, you just didn't give enough of the alternate drug.
DR. SHEHABI: I think the TCI model is based on computer modeling in relatively healthy volunteers, and I think transferring that into the critical care population I think is not a straightforward phenomenon.
MALE VOICE: Where are we going to get the devices as well? Because wouldn't they be investigational? They're not approved in the U.S.

Page 338

DR. SHAFER: Interestingly, for AstraZeneca, it was just part of the trial, and the whole thing was approved. And it was ironic because the FDA device division would say, "Oh, we can't do these," but CDER was quite happy to use the data from TCI in approving propofol for ICU administration.
But let me just address that specific point because I know what Talmage and I would both say to that. That is, what TCI allows you to do is at least hold a more or less steady concentration around which one titrates. If you're just talking about giving boluses and infusions and randomly going up and down, you're assuming that the infusion rate, in and of itself, is going to instantaneously translate to what's going on in the patient's brain. At least the target has got a better shot at giving you something and holding it steady so you can then make your adjustments.
DR. SESSLER: Even if you don't know the absolute.
DR. SHAFER: Even if you don't know the absolute, because you'll measure it.

Page 339

DR. SESSLER: You get more stability.
MALE VOICE: And you're looking at an acidic end-stage renal disease population for which we have no data as to what the pump is actually going to have in the body and the multiple compartments.
DR. SHAFER: So what are you going to do; just pick dose? That's going to be better?
MALE VOICE: Titrate to an end effect.
DR. SHAFER: And if you're going to titrate, use TCI.
Other questions?
DR. BALAS: I have a question. I'm wondering if anybody in the United States has been successful at getting an IRB through with a deferred consent.
MALE VOICE: With what?
DR. BALAS: Deferred. As Yahya was saying -- is it true for the SPICE trials? They enroll the patients, randomize, start the intervention, and get consent later.
DR. GIRARD: Are you specifically asking about sedation trials or any trial?

Page 340

DR. BALAS: Anything in the ICU.
DR. GIRARD: Yes, It's been done for non-sedation related trials, yes.
MALE VOICE: I think it's been done in emergency settings like seizures out in the field and that sort of thing, but I think your point is it's almost impossible.
MALE VOICE: I mean, there's an ethic model, so I think that's with cardiac arrest and those things. And there's a community consent, and there there's a process. We've actually had three trials that we've done, none of them sedation trials, but we talked about doing it in a cardiac arrest model with early antibiotic therapy and doing ethic for that.
We didn't do the study because we didn't get funded, but ethic is a pathway where you can proceed. You try to get consent if you can. If there's no family or no surrogate around you, you enroll the patient if they haven't excluded themselves in advance, and then when the family arrives, you inform them and go from there. So

Page 341

there is a model.
MALE VOICE: There's a second thing you have to do with those patients, is when they become at a state where they can consent, you have to approach them, and they can obviously withdraw at any stage.
DR. SESSLER: Cluster randomized trials automatically have waived consent because you're randomizing an entire facility to something or something else, and they're used typically for system-wide interventions.
Let's say electronic records. Electronic records are not something you can turn on and off on a patient basis, but if you want to assess the effect, the only way to do it rigorously is either cluster randomization where you have whole facilities that start, or don't start, or a step-wedge, which is similar to cluster. Neither of those has individual patient consents.
We've also done half a dozen of these alternating cohort studies, which are good for comparative intervention studies, so when you're comparing two perfectly reasonable standard

Page 342

clinical interventions that are done all the time. For instance, isoflurane versus desflurane, lactated ringers versus saline; two different title volumes. These are examples of trials that we've done with waived consent, so there certainly is a precedent for doing some sorts of studies with an altered or a waived consent.
DR. AITKEN: I've got a feeling that I remember Martha Curley telling me that in her cluster RCT, they could obviously allocate the sites to the intervention, but they couldn't collect any data in the intervention sites until they had consent. I don't remember the details, but I remember her having a real problem.
DR. SESSLER: It's certainly possible; it's not the classical way to do a cluster --
DR. AITKEN: Yes, it seemed odd.
DR. SESSLER: -- because if you're doing something like electronic records or an enhanced recovery pathway, you can't really turn it on and off.
DR. AITKEN: No. It seemed really odd, but

Page 343

it might be worth just checking.
DR. SESSLER: So the normal way to do is just you do it.
DR. AITKEN: Yes.
DR. SESSLER: There are rules about waiving consent. In Europe and in Australia, it's very difficult. There's essentially no regulatory pathway for doing it. In the United States, there is a legal pathway for waived consent, and it requires a number of things, which include minimal risk, and the study can't be practical without waived consent. Now what defines practical is open to some dispute, but one of the things that is considered part of the practicality decision is the cost and difficulty of the trial.
DR. SHEHABI: Can I just make a comment about the clustered randomized trials? We're involved with two clustered randomized trials in Australia. One is the MIT [ph], which is the likely rapid response team, and doing the clustered studies require a huge number of sites, a huge number of -- I mean, each cluster essentially

Page 344

blocks one --
DR. SESSLER: N equals 1.
DR. SHAFER: It needs such a massive number of patients involved to get to power.
The second thing I wanted to make a comment about is the waived consent. There is a regulatory process in Australia to do that. There is a regulatory framework to do that. Essentially, it varies from state to state, and it changes by whatever the parliament thinks on the day.
In Victoria for example, if the two interventions are considered within usual accepted practice, then waived of consent or deferred consent is acceptable. In New South Wales, the same trial, which is FOSTERI [ph], the same trial, the guardianship board, which is like the body that makes the law, said, "No, no, you can't do that." And we said, "No, we disagree with you." We have to take the guardianship board to the Supreme Court to change their mind, and now they're changing the law in New South Wales to say, yes, when things are similar, yes, you can do a deferred consent.

Page 345

DR. SESSLER: All localities allow at least deferred consent for emergencies, say out of hospital cardiac arrest. That was actually on hold for about a decade worldwide to everyone's detriment. Now everyone allows that.

MALE VOICE: Not everyone. Sweden doesn't do that.

DR. SESSLER: Okay.

MALE VOICE: Be careful when you travel to Stockholm.

(Laughter.)

DR. ABSALOM: Tony Absalom. I was just going to say there was this trial of adrenaline during CPR in the UK, but they had to jump through an awful lot of hoops to do that. They had to have all these media campaigns to allow all exposed possible people to notify that they wouldn't like to be enrolled should they have a cardiac arrest.

MALE VOICE: Do you wear a bracelet?

MALE VOICE: Yes.

DR. DEVLIN: John Devlin. The one thing I just want to bring up, too, is the extent to

Page 346

exclusion and inclusion, which a lot of things are obvious, safety issues and confounders. But we end up with studies that are sort of a leading [indiscernible]. It's quite low. Maybe only 10 percent of the population is actually enrolled.

Any thoughts from the panel on -- I realize I guess a pragmatic approach could be used to eliminate some of that, but any thoughts on where we're trying to always remove all those sources of bias up front?

DR. SESSLER: It's an issue that all trials face. One of the factors that influences sample size is variability, and that's under your control, unlike treatment effect, which is not under your control. So you can reduce sample size by reducing variability, and the way to reduce variability is to have a fairly narrow homogeneous population.

The other thing is that if you're testing a new drug or device, especially if this is funded by people who have an interest in your drug and device, they want to select patients who are most likely to benefit and least likely to have

Page 347

complications. So that's why sponsored trials have these very long lists of inclusion/exclusion criteria, whereas investigator initiated trials tend to be more reasonable.

Narrow enrollment criteria reduce variability. It makes the trial results easier to interpret, but it makes it harder to enroll and less generalizable. So the broader you can make them at the expense of variability and increased sample size, you end up with a result that's more useful.

DR. SHAFER: There are regulatory implications, too. If I might ask how the FDA views going narrow to get something that's very precise. And you can say, well, gee, I can really interpret this trial because it's very narrow versus some label that's going to be used by 2 million people within a month of being approved.

DR. ROCA: This is Rigo Roca again. I think those are both very valid points. As you noted, if it has a very narrow enrollment, you're right; it's easy to interpret. We get a better assessment of

Page 348

the treatment effect, the side effect profile of the product and all that, but the ability to generalize is limited.

So one of the other things that we are certainly open to is that one trial would be narrow, and then you could have another trial to replicate the results but have that be, if you wish, all comers, or a little wider, but you can get a wider population that may be more generalizable to the public. We're very much willing to see that.

DR. SHAFER: Dr. Ward?

DR. WARD: Denham Ward. We've heard a lot of things about outcomes. The amount of time at sedation level is probably no longer an appropriate primary outcome; we've kind of moved past that. And now we've heard some things about composite outcomes as a way to improve and get more power on a clinical trial. What composite outcome should we be using? If we're going to use a composite outcome in a clinical trial for sedation, what is it?

Page 349

DR. COLANTUONI: Never ask you statistician that question.
(Laughter.)
DR. COLANTUONI: I'm just kidding. No, but I'm not.
DR. SHEHABI: I think a composite outcome may look like a solution, but it's really a very imperfect solution. I think there is a lot of issues with composite outcomes. We find that public funders in Australia, for sample, they will rank a trial that has a primary outcome as a composite outcome of multiple things, and their rank is brought down because of all the issues that you've mentioned before and I've mentioned before. So I'm not sure that we do need to invent a composite outcome for sedation trials.
Probably to go further to what you said, Dan, before, that a certified baseline, what we chose to do with a spot [indiscernible], rather than serve at baseline, is to have a better sample size is to choose the subgroups at the median level of what are you looking at, whether it's age, or

Page 350

Apache [ph], or whatever, and that would immediately give you two halves of the groups, distributed nicely between the groups.
DR. SESSLER: Composite outcomes are good for rare dichotomous outcomes. Most of the outcomes we're talking about here are not dichotomous. Death is, but the others are not. We're talking about mechanical ventilation, time in the ICU, functional outcome; those are all continuous outcome. They don't lend themselves to composites very well.
DR. COLANTUONI: I think there's a distinction between what Dan just said and the approach to summarize an outcome that's dynamic over time like ventilator-free days and delirium-free days, that incorporates the complicating factors of mortality, which is how I was defining composite outcome versus the difference between what Dan was defining as composite outcome.
So you can create a composite that is a delirium, but summarizing over a time course and

Page 351

potentially calling for competing risks is slightly different in my thinking than saying we're going to include mortality and some other adverse events that we might see -- other binary adverse events that we might see over the course.
DR. SESSLER: Well, I think it's good for complications --
DR. COLANTUONI: Oh, yeah.
DR. SESSLER: -- because very often, complications are rare. Your primary outcome is how well does a drug sedate somebody? Well, you're going to look at measures of sedation for that primarily. But if you want to know does this drug cause complications, now you're suddenly looking at a wide variety of presumably rare events, and many of these are dichotomous. Composites are a really good way to look at the complication. You're never going to be powered for individual types of complications.
DR. SHAFER: Let me point out that Dr. Ward just asked a question that was similar to the one I asked, and looking at what we are here for, a

Page 352

patient-centered outcome. I think he sort of said what is that patient-centered outcome, which I tried to ask earlier, and I did not get an answer. And you tried to ask it, and you just didn't get an answer.
Let's try again. Both to the people on the panel and people in the audience, what is a measurable, trialable, falsifiable, patient-centered outcome?
DR. AITKEN: I'll start by saying it has to be beyond hospital. I don't think that we can be having a primary outcome that's only in hospital would be my suggestion. But what it should be, I think we could argue for various things.
DR. SHAFER: Like?
DR. AITKEN: Like, certainly mortality is the obvious one, but I think more functional measures like returning to work, or that's probably the one that jumps out as an obvious one because that incorporates a whole lot of other things in there. You can't return to work if you don't have reasonable, functional health, reasonable,

Page 353

psychological health; and reasonable cognitive health.
DR. GIRARD: And we should recognize -- actually, you just described it beautifully -- return to work is a composite outcome because all of those things have to be true for you to return to work.
DR. AITKEN: That's the risk of it.
DR. FLOOD: The thing about return to work is that many people in the ICU aren't working.
DR. AITKEN: Sorry. I should say return to work or previous normal activity. Sorry. Yes, it has to be a broader definition than that; you're right.
DR. FLOOD: How about a quality-of-life outcome?
DR. TANG: Sorry. Real quickly. There's a work productivity, activity measure that -- it's WPAI. I apologize. I'm just scrambling to remember what the acronym stands for, but it does measure essentially not only work but also activity impairment that could be associated. So just a

Page 354

note that that's a regularly used one in the quality-of-life space that's typically used.
DR. DEXTER: I was going to say from a point of view, two separate issues. When it comes at least to the retrospective analysis of data, I found it to be quite challenging to do something, whether it's work or functional activity, among patients for which you don't have baseline measures. Unless you are planning to be in the ICU, at least when I try to analyze those data, I tend to find very weak baseline measures. So it's something to consider.
But I'd like to go back to address your point about patient-centered outcome. Avery has made the comment in terms of thinking about an indication. The point that you brought up is to be able to provide for those patients for whom the goal is to provide a deep level of sedation to prevent adverse events or something like that.
I think that actually there is something to be said for that. When you think about analogies in terms of endpoints of anesthesia trials, there

Page 355

are a large number of patients, large numbers of procedures, where general anesthesia is not an appropriate endpoint. You may use a drug for different applications, but initially at least you start with the subgroup of patients for whom you think you want to do something.
The other thing about it is you can also ask individual patients whether or not that would be an appropriate choice and so forth. I think that that is not an unreasonable approach. It's not going to be that trying to provide a deep level of sedation is appropriate for all patients; it's going to be a minority of patients, but you can have a patient-centered outcome for the subset of patients for whom you want to be providing that.
MALE VOICE: So an individualized patient-centered outcome.
DR. DEXTER: Well, that is how we would be doing -- if this were not a question about ICU patients, If this were a question about satisfaction of patients after general anesthesia, satisfaction with monitored anesthesia care, that

Page 356

is exactly how we would do it. You would ask the patient, so to speak, or the surrogate for the patient, what would that patient want given this particular context. That would be the patient-centered approach. And given the condition on the idea that it's going to be general anesthesia or deep so to speak, go forward in that way.
DR. NEEDHAM: This is Dale Needham. I think we're giving all our perspectives on patient-centered outcomes, but what research has been done to rigorously understand what are patient-centered outcomes, I'm not aware of it with respect to sedation. So I think until there actually is research done doing that it's just kind of everybody's opinion on that.
So I think there needs to be an agenda so that people actually do that, and I think one starting point that is great, as Leanne showed, is these are the outcomes that have been used. That's kind of like a scoping review kind of thing to think about. Those are kind of candidate things.

Page 357

Then we need to be talking to patients.
Also, we need to think about maybe the most important outcome is going to be a resource utilization one, perhaps, in terms of shortening duration, mechanical ventilation, length of stay at hospital, that might be where the strongest signal is between an intervention and an outcome, at least based on my understanding of prior studies, and we want to show that there are positive signals of benefit in other things as well and no harm, and that might be, at least from my naive perspective, the best way to be thinking about it.
Do people want to argue the opposite?
DR. SHEHABI: I just wanted to add, I think the context is quite important. And if you're looking at a patient-centered outcome that looks at function or outcome, for example, it's important to go back to the inclusion/exclusion criteria that Tim was talking about, where you would not include in a sedation trial, for example, traumatic brain injury patients, or patients who come with a green beret, or patients who are going to be intubated

Page 358

for 6 months because of a neurological disease because the outcome is going to be determined by their underlying illness rather than by the sedation that you're doing.
I think it's very important to marry the patient's outcome we're looking at with the population you're studying and pretty much like what you mentioned about the seizure population, the same for the ICU population. In our trial, John, we've excluded anybody who had any neurological problems whatsoever, whether they have weakness or brain injury of any kind.
So we want them to be completely neurologically intact on entering the study, and for that that, the patient-centered outcomes that we looked at were specific on things like -- in addition to mortality, we looked at cognitive function at 180 days, institution dependency at 180 days with basically societal resource utilization.
Then on top of that looked at their quality of life at 180 days in terms of what are they able to do, and because we knew that they entered

Page 359

intact, we could assess them at 6 months and say this is where they were at this point in time.
DR. SHAFER: Leanne?
DR. AITKEN: I was just going to pick up on Dale's comments. Certainly, I think the issue of talking more to patients and asking them what they want is absolutely essential. It still doesn't tell us what every individual is going to want, but that gives us a better sense.
My only hesitation in what you said about resource utilization is I think we have to think health system wide rather than just hospitals, so I'd be hesitant in only looking at resource utilization within the hospital because if we're shifting sick or dependent patients outside the hospital, then we're shifting resource utilization. So I do think we have to think across the system.
DR. SHAFER: I'd like to pose a question to Frank. Frank, you do a lot of work with economic analysis, and basically why should somebody invest in something? Why should they invest in a certain kind of system? Why should they undertake a

Page 360

certain study?
Let's say that we come up with a patient-centered outcome that everybody says this is a great patient-centered outcome. This will really improve the wellbeing, somehow measured, of patients who are sedated in the ICU. Somebody's got to invest in that. There's got to be some sort of return. I mean, I agree that we do this for noble reasons and for academic glory and things like this, but these trials are big. Big trials are expensive. Somebody's got to invest.
How would you put together economic argument that whatever this great thing is that we're going to measure should be studied, measured, and improved? How does one go about making that economic case?
DR. DEXTER: I don't think I can answer that question per se.
DR. SHAFER: Can you answer a different one?
(Laughter.)
DR. DEXTER: Yeah.
DR. SHAFER: Let's suppose that you've got a

Page 361

company with a hypothetic gold product, and they're thinking about actually bringing product to market. Let's suppose that you've got the following option. One is you've got resource use in the hospital, ventilator days with adjustment or something like that.

One of the challenges you have -- I'm sorry; it's a slightly long answer here. One of the challenges you have is that the dollar value associated with these resource uses will vary massively among organizations, and really this is a function of the variability in the workload within the organization.

So that's why things like ventilator days, a few primary endpoints which are measurable, works totally adequately. If you've got tons of ventilator days, you have more costs. That is easy to understand. Also, there's a difference from a regulatory point of view, you can measure it and do the trial.

In contrast, when you're thinking about -- let's take a couple of others -- long term

Page 362

from the point of view, something about the functional measure, quality recovery of the patient, or something like that, at least from the point of view of critical care and watching companies make these decisions, they freak out because you don't have the baseline measurements. You're not really randomizing patients where you have this and stuff like that.

It seems very large sample sizes compared to the consumption or something like that. That seems to be something which you would do after you have the drug approved, then you might go ahead and do it; at least that's what I tend to hear.

The costs are oftentimes the families and things like that. But again, the problem is going to be are they then going to be able to sell the drug and what is going to be the variability, and how are you going to actually randomize a patient, stratify based upon that? I think that the answer would be, typically, hospital resource use makes quite a bit of sense practically.

DR. DEXTER: How about how about a post-use

Page 363

like looking at SNF facilities. They're long-term care after ICU, so try to avoid these very expensive outcomes that are measurable.

DR. SHAFER: Yes, but I think that one of the things would be is that it's quite -- when I say straightforward, I don't mean like trivial; to be able to use a variety of different economic endpoints such as that, which is days in [indiscernible] care or after the hospital; days on the ventilator and things like that. Those things can be combined in terms of quantitatively and stuff like that.

DR. MAZE: Steve, can I ask you a question about your question to Frank, that he changed the question up.

DR. SHAFER: He really didn't.

DR. MAZE: No, he didn't. In a situation where the patient is not directly responsible for the cost of the care, and there are many countries like that, what does it matter? Where's the patient centeredness about that?

DR. SHAFER: About --

Page 364

DR. MAZE: -- about resource utilization?

DR. SHAFER: That's why I'm asking the question.

DR. DEXTER: I don't think it's patient. I going to take an extreme example. Like ventilator days, I don't see how that's patient centered at all. It completely escapes me how that would be patient centered, or maybe I'm totally missing something, and I apologize.

MALE VOICE: You probably haven't been on a vent.

(Crosstalk.)

FEMALE VOICE: The risk of respiratory infection and death is directly tied to ventilator days.

DR. AITKEN: But those say some patients describe quite vividly wanting to get the tube out, so there's that angle of it as well.

DR. KRESS: But I think it's important, this concept of patient centered, I certainly think it sounds good. You have to be careful what you ask for, though, because you ask the patient, the

Page 365

patient doesn't necessarily understand what the implications are. Put me in a coma for 4 days; wake me when it's over. It sounds good except when you actually come to realize what that entails.

So ventilator days isn't [inaudible - mic face] patient-centered from one perspective, but from another perspective, the longer you stay on the ventilator, more likely you are to have problems X, Y, and Z, that are going to affect you down the road. So maybe that's just semantics, but I would argue that ventilator days is very patient centered if you look at what it means to the patient down the road.

DR. SHAFER: So we're in our last two minutes. Go ahead, Frank, but we're going to kind of go quickly here.

DR. DEXTER: I think when I think from an anesthesia point of view of patient centered, it is in things that are -- all outcomes, death is very bad for the patient. Pneumonia is bad for the patient, but that's not what I think of. When I think of patient-centered outcome as something like

Page 366

that, it's quality of life, quality of recovery, those types of things.

FEMALE VOICE: But if you have pneumonia, your recovery is going to be awful.

DR. SHAFER: Claudia?

DR. SPIES: I think the preference of some not very valid structure established, and I think what we did last year is try to inform the patients much better. So the patient preferences need to be in fact boxes at some point.

I think it's very difficult for patients to understand what we tell them, and even as for us, it's difficult at the end because we often give the wrong information because we have not enough knowledge. This is also a problem because we don't always see the whole path.

So I think we need preference with processes of structured interviews like in the shared decision making processes, and then we can evaluate if we have the right knowledge, all of us, think. That's a major issue I have. So I think we need global knowledge in that structured patient

Page 367

preference.

DR. SHAFER: We're down to 60 seconds. Bob, you're next.

DR. DWORKIN: We were reminded this morning that the FDA approves drugs, biologics, and devices that improve the way patients feel, function, or survive. At least to me, those all sound like patient-centered outcomes: feels, functions, survives. So I think as long as we're in that big bucket, we're talking about patient-centered outcomes, and also in line with how the FDA regulates drugs, devices, and biologics.

DR. SHAFER: Anna?

DR. ZHAO-WONG: Thank you. I'm just trying to think out of the box. When you ask the question, economically, why should we do these large expensive trials? Well, one thing to think about is insurance, and insurance would pay those ways, methods, and treatments that have better patient outcomes.

DR. SHAFER: My insurance company doesn't care, but maybe yours does.

Page 368

You had one comment.

DR. SHEHABI: I just wanted to make sure that we don't really lose the baby with the bath tub. As you sit there, the patient-centered outcomes should improve the patient's survival, function, and feeding.

For that reason, I think we must not just focus on things that come outside the ICU because, for example, delirium, when we know how much it impacts patients, must really be an important patient-centered outcome. We may argue about ventilation, an extra day or an extra 6 hours, but I think we need to be quite clear that delirium is absolutely a patient-centered outcome.

DR. SHAFER: Dr. Ward, you get the last word.

Adjournment

DR. WARD: A couple of things, housekeeping. But just to comment, for example, the drug from my generation, droperidol, is a great, non-patient-centered drug. It works great as a sedative, but if you ask a patient how they felt,

they felt horrible. So you address patient centered by finding out how the patient actually felt through it all.

Thank you all.

(Applause.)

(Whereupon, at 5:04 p.m., the meeting was adjourned.)

B

C

D

F

G

H

I

M

N

O

P

Q

R

S

T

W