FOOD AND DRUG ADMINISTRATION

_____

FDA-IDSA-NIH-Pew Public Workshop

Enhancing the Clinical Trial Enterprise for Antibacterial Drug Development in the United States

_____

DATE: Day 2: November 19, 2019

TIME: 8:30 a.m.

LOCATION: FDA White Oak Campus
10903 New Hampshire Avenue
Building 31 Great Room
Silver Spring, MD 20993

REPORTED BY: Michael Farkas, Notary Public

JOB No.: 3472551

PARTICIPANTS:

JOHN FARLEY
AARON DANE
ROGER LEWIS
DAVID MELNICK
RIENK PYPSTRA
CHIBUZOR UCHEA
VANCE FOWLER
CYNTHIA SEARS
SARA COSGROVE
REBECCA REINDEL
SCOTT EVANS
LINDSEY BADEN
WES KIM
JOHN REX
HELEN BOUCHER
SUMATHI NAMBIAR
AMANDA JEZEK
DAN RUBIN
JANE KNISELY
PAMELA TENAERTS

MARK ALBRECHT

AMY LEITMAN

NICK KARTSONIS

RYAN CIRZ

SUE CAMMARATA

MANOS PERROS

P R O C E E D I N G S

DAN RUBIN: Good morning, everyone. Please take your seats. I'd like to get started. Welcome to day two of the workshop. We're now beginning our session, strategies to better support antibacterial drug trials. I think we won't go around the table and redo introductions, except I should probably introduce myself since I was in the chairs yesterday and I've jumped all the way to moderating the session along with it. Jan Knisely, who's going to be the co-moderator for today.

So I'm Dan Rubin, and I'm a statistician at Cedars. Professor Evans, could you introduce yourself since you weren't here yesterday?

SCOTT EVANS: Sure. Good morning, everyone. My name is Scott Evans. I'm a professor and the chair of biostatistics (indiscernible) at George Washington University. And I'm the Director of the SDMC, Statistical Data Metric Center.

DAN RUBIN: And do we have any other new people who weren't at the table yesterday? Okay, great. Well, then let's get started. Our first

speaker is Dr. John Rex. John Rex is the Chief Medical Officer at F2G''s Manchester, UK based firm and brings 30 plus years of development and policy experienced, focused on antimicrobial agents. His experience puts moving antifungal and antibacterial agents from pre-clinical development through all development phases and various roles.

JOHN REX: Thanks, Dan, and thanks to you all of you for being here. And this theme came up yesterday. This is a pivotal moment in time. What do we need, what should we be doing, who should be doing it? And the title of this talk is antibiotic R&D 3.0, and I'd like us to think about rowing the boat together and in the same direction.

These slides are readily available. I actually think the slides from the whole meeting are going to be available, going to be posted, and obviously, I'll share them via my newsletter.

So the pivotal point in time is fasting. We have poured enormous resources into pre-clinical and phase one, CARB-X and so forth, and there is really signs of progress. We are seeing some

innovative things deep down in the pre-clinical pipelines, things that nobody had ever discovered before. They won't all be drugs, but some of them will be.

Unfortunately, this whole enterprise, hundreds of millions of dollars is going to blow apart due to two intertwined issues: the first is that new antibiotics are kind of like a bridge to nowhere. Actually, that bridge on the left, it actually is, at low tide, there is a path to and from it, but at high tide, it's just out in the middle of the water. And right now, it's high and we're drowning.

New antibiotics don't get used. Some of the agents truthfully aren't that interesting. But even the good ones are perceived as only non-inferior. The guidelines are out of date by years and stewardship is based on cost and utility.

And then on the flip side is that the payer model is broken. Antibiotics were the fire extinguishers of medicine, as you've heard, and we need to stop paying for them on a per-fire basis.

We cannot fix it all at this workshop.

The things that are in scope are the types of trials we can conduct, the types of data we can realistically get, how those data should be reported in labeling, how the ID community talks about this data, how the guidelines and the guidance's should handle this data. What's out of scope is payer models. And also, I'm going to say that personally out of scope is, oh, I just don't like it. Okay? If you don't like it, you either have to suggest something else that's better or get with it because you're otherwise keeping us from going anywhere; it's why the boat is spinning in circles.

So with that rant over -- sorry -- this thing, big picture. Antibiotic R&D 1.0 was sort of the dawn of antibiotics to the mid-2000s. It was generally very easy to see the value in the drugs. But over time, we were learning about weaknesses in our pivotal designs, particularly related to upwards of 20 infections.

R&D 2.0 was the moment in time when it all kind of went back to ground zero. The Ketek hearings in Congress in 2007 was a period of time when

we didn't know how to develop anything; we weren't sure about it. When we got busy, the f'ing NIH, the whole process led to rapid refinement of not a very orderly science.

And we now have very clear, very sound scientific roadmaps for major indications. Single pivotal trials gradually became acceptable for approval, and there was a lot of harmonization between EMA and FDA, and also increasingly recently Japan, and that's all really good stuff.

So what's R&D 3.0 going to look like? Well, these are the ideas. It's how do we use LPAD as a springboard, the problems we have to solve, a little bit about superiority again, and the notion that this is a community-level activity and some suggestions.

So LPAD, we heard that mentioned yesterday, and Dr. Nambiar talked about three levels of approvals: standard, limited use, and LPAD. In truth, there is no distinct labeling for -- there's no distinct pathway, it doesn't have a name for limited use; it's just -- it's the agreement that we can approve with single trials.

But LPAD is a special beast. It was created by Congress as part of the 21st Centuries Cares Act, and it really ought to be called LPAAAD because it's limited population of antibacterial and antifungal drug. You'll notice there are least two As missing in that.

The concept behind LPAD is that physicians and patients will accept greater uncertainty for serious diseases with unmet needs, a very standard concept, but now codified with respect to antibiotics. And it says that, in brief, streamlined approaches based on severity, rarity and prevalence are what you're going to have to do.

And this means single trials, widen inferiority margins, basically being creative within certain boundaries because it's not a license to run riot. You still have to meet the standard of substantial evidence of efficacy based on adequate and well-controlled clinical data.

Also, the labeling must make it clear that the limited population -- that there is a limited population for which this drug is addressed. And this

is -- it's important for us that LPAD gives us two gifts: gift one is the phrase itself, LPAD. This drug is special. Has the logo been designed yet? I don't know, but there's going to be something on the box in a little triangular thing that says this is different; don't use this unless you understand it and you understand the risk/benefit around it.

Don't use it in just anybody; use it only in these people. And the package labeling -- I'm sorry, the label data actually has this phrase right here, limited population in at least six places on page 1 of the label, some of them in big multipoint type.

And you combine that with robust stewardship programs and CDC's ongoing surveillance, LPAD agents are something that we can pretty much say to the community, you have to use this right. And we can -- we should leverage that notion that this is something where we can tell people, be careful here, use it appropriately.

But that leaves us with some problems there. R&D 3.0 needs to address two big groups of

problems: the first one is communicating the value of standard NI trials, and I put that one separately because that's a community-level thing that we can take on. We have to educate ourselves and our peers on the notion that a cUTI trial is not stupid.

And then B, C, and D all tangle together; developing for rare or resistant pathogens, developing for less common infections. And what's the adequate quanta of data for labeling for B and for C -- B, C and D are reduced to study size -- and how you think about that phrase again, "Substantial evidence of efficacy based on adequate and well-controlled trials."

Nowhere in any order or reg does it say that that means alpha means .05 and a margin equals 10 percent and a specific end point. None of that is defined. What we do -- it's up to us to define stuff that is solid enough, and we can consider risk benefit.

A little bit of a sidebar about superiority, and you just have to bring this up and remind everybody about this. Superiority is not a

generalized answer. If you happen to find superiority for a new drug, great, but that instantly resets the playing field for every subsequent drug. You have to keep that in mind. If it's -- the core problem is that antibiotic responses are dichotomous -- cured/not cured. Very few other diseases are that way. I cure your pneumonia. There's nothing better than cured. You go home and have your next 60 years of life.

And if it's easy to run a superiority trial, actually something bad has happened in public health resistance must be common enough that I can put people in an arm where there wasn't a good choice. And except for the mildest of infections, superiority means that not only somebody got hurt, somebody probably died and did not really need to die if we had an appropriate drug.

So we actually want superiority trials to be impossible. If it's briefly possible due to a gap, the first successful drug closes that gap and makes it hard to go back to doing more superiority. So we all have to know this and we have to teach it to ourselves and teach it to our colleagues: non-

inferiority is our main tool; they're sensitive to drug effects. Modern designs work really well, and we have to be very clear about this.

And this segues into the notion this is not just a regulatory problem. It's easy to be critical and ask, why is it you say it's nothing; oh, it's just a cUTI study. What it was, because you actually want more data. We talked a lot about a lot this, I wish for more stuff. Academia and the journal letters say, well, you did this resistant pathogen study, but it's too small. And the payers say, I expected superiority data in the label. And physicians say, I'll wait for the guidelines to change. And the patients say, non-inferiority sounds so dodgy; I don't know what it really means, but it really sounds awful and my doctor doesn't like it.

This is all a communication and an education problem as we've just been discussing. People don't understand the core scientific principles. Why is it that some non-inferiority trials are rubbish? We have done NI trials and published them in the past and they were rubbish. The

ones we do now are not rubbish. There's a reason for that change.

Non-traditional agents. We got somebody at the table today who's interested in bacterial phase. There are a variety of other non-traditionals we'd like to have developed. They face the same problems. We actually -- if we're going to have any of these new toys, we're have to deal with those, so here are my suggestions.

The first one is: we, as a community, have to become cognizant of the labeling regulations. You need to take the words on this slide and laser etch them on the inside of your eyeglasses so that you're constantly looking at them and are aware of the fact that there are things the FDA can do and things the FDA can't do. These are the regulations. An act of Congress would change them, but it doesn't seem terribly likely. Let's work with this.

So within -- with that knowledge, supplement as we can. So I would like to think about creating working groups to create a credible way to work with the available data. We do need to talk

about the, in particular, talk about the idea of adequate and the words adequate and well. The patient with the word adequate, remember that patients and physicians will accept different tradeoffs. The word well, I'm not arguing about control, but what's the definition of well.

Remember that 100 patients equals $10 million more or less and several years' worth of work. It's not trivial data and it's hard to get all the information you want. Can we supplement with external controls? We're going to hear a talk today about sharing across -- same as yesterday, sharing across body sites. All of this stuff needs to get brought together to talk about what can we do inside the label to supplement the label, and at the same time, we need to do this. Agencies, societies and journals need to be spreading the words.

Superiority trials. I've already talked about that, and just last rant about superiority. Now, you have people who say if it's a superior trial, it would be so much smaller. Yeah, yeah, I know that, right? We all know that. But it's

not a path forward because we try to do those studies, it's lasagnas law. The patients go away because the hospital epidemiologists are making the resistant infection go away. And nobody wants to go to the hospital to test, there's a sign in front that says, world's leading center for resistant bacteria. Come and get your transplant, right? No, no, no, no, you don't want that at all. You want that to not be the case. And this is not a migraine. We're talking about a place where superiority means something bad happened.

This is the next message: non-inferiority is not a synonym for worthless, and guidelines need to be -- or guidances need to be continuously updated. And colistin, please, if I get sick, do not give me that poison; give me a real drug.

Finally, industry needs to be aware of the constraints on the first previous three slides and to focus on novelty and unmet need. You know, Kevin talked yesterday about the need for incentives and different models and that's not today's discussion,

but there's a group of us working very hard on pull incentives. But I will tell you this: I think some of those are going to come to be, but they're not going to come to be for every drug.

If you're working in this area, it's your job to pick something that you think moves the needle. QIDP alone is not going to move the needle. You've got to actually be doing more than that.

So to close at this, you know, at heart, I'm an ID doc. I moved into industry in 2003 because I spent -- a couple of things: antifungal pipelines were coming to a grinding halt, there was nothing to work on really; and I had spent five years as an epidemiologist, and I was starting to see infections at our hospital that can only be treated with (indiscernible).

I once closed an ICU for about a week because we had an outbreak of a then-untreated infection, and I shut down all the ORs. I said you can't do anything elective, you know, which made a lot of people unhappy. We stomped it out. It was an (indiscernible) bacteria outbreaks. You know, I'm

seriously hoping these nasty (indiscernible) will work. But that's the kind of thing that drives us all to be here.

Since then, I have walked all sides of this -- large pharma, small pharma, VC, philanthropic funders, you know, all that. And my deepest message is that tradeoff-free solutions to antimicrobial resistance do not exist; if they did, we'd all use them. Since they don't, we, as a community, it's the time for us to move forward with making the best available tradeoffs, preferably by 1:00 p.m. today. Okay? Thank you.

JANE KNISELY: Our next speaker is Dr. Vance Fowler. He's a Professor of Medicine at Duke University School of Medicine. His research interest focuses on staphoreious clinical epidemiology and pathogenesis. And he's led important clinical trials, testing new therapies for staphoreious bacteremia, including a randomized controlled trial comparing daptomycin to standard therapy. And I'm a little surprised that not in his bio, he is also one of the PIs of the Antibacterial Resistance Leadership Group.

VANCE FOWLER: Thanks. These are my disclosures. I'm going to pick up on the point that John made that there are some things that the FDA can do and some things the FDA can't do. And I'm going to pick up from primarily from the perspective of the clinical need. Because the bottom line is, I feel like we in the scientific community are going to have to ultimately generate the data that we need to manage the patients that we see.

My points of this talk are outlined here. Registrational trials necessary, not sufficient. Strategy trials, and I'm going to emphasize that hard because I honestly feel like that's giving the people what they want, and in clinical networks to help both approaches.

So what do clinicians want? Well, they've told us what they want. This is a paper from 2013 in which a large group of Australian ID physicians were asked what they actually need in clinical trials. And there are some pretty familiar faces here for those of you practicing, per se; joint infection, osteoarticular infection, uncomplicated

staphoreious bacteremia, diabetic foot infection, and treating ESDLs and then there were a variation on that.

And if you actually took these 13 top prioritized items and broke them down into four categories, they were fundamentally listed here: duration of ID antibiotics, how long do we treat these patients; combination drugs, are two drugs better than one; specifically, how do we treat MDR pathogens; and then root of administration, can we use PO antibiotics and abbreviate the time a patient has a line in.

These are the four categories, and I'm going to come back to this because this is going to be the benchmark against which I'm going to compare the registrational trials that have been employed and the strategy trials.

So in terms of registrational trials and agents that have been approved from 2013 to '18, and I'll limit my discussion to 2018, so there were the four agents. Again, this is through 2018 for skin and soft tissue, complicated UTI, intrabdominal infection; also acquired pneumonia, the

(indiscernible) from way back in the day, and the (indiscernible).

So if you look at the actual trials that were published, phase three registration trials published in 2018, there were basically five. There was the American and Vabor vacuum study, two ABSSSI studies for delafloxacin, and two ABSSSI studies for (indiscernible).

So I'll bring you back to that metric of the studies that ID clinicians actually want, none of which were addressed by the registrational trials from 2013 to 2018, so that's my point. And that's not being disrespectful to the FDA; these are just the facts. So now, and the problem brought this up yesterday -- Helen brought it up, I brought it up -- you still have to make a decision. So these are three new drugs -- well, cefazoline's not quite so new anymore, but new enough -- taz-avi and (indiscernible).

So what I did was I went back for the first quarter of 2018 at Duke, and I asked the question, are these drugs being used, and, if so,

where are they being used. And the data shown here is they're about somewhere around 45 or 50 patients, unique patients, who have been treated; 87 percent were being treated off-label, so this is an N1 47 times.

And this is the problem right here, it's an inconvenient truth, but here we are. Clinical trials don't equal clinical practice. Erythema not receiving in 72 hours alone; to me, that doesn't mean failure. Not obtaining a blood culture six weeks after stopping antibiotics for a patient with staphoreious bacteremia, guess what, that's not a failure. And if a patient gets four days of drug B after getting drug A, not a failure.

So what's the point then? The point is, you know, that we as clinical trialists and we as clinicians really are tasked with importantly different responsibilities, both of them essential, but really significantly different. And so, because of that, I'd like to argue that strategy trials really go a long way towards addressing the trials clinicians want.

And I'll do that by demonstrating the strategy trials that were undertaken in 2018, so there were six of them. There was Stephen Harbath's paper on comparing nitrofurantoin for a single dose phosphor for uncomplicated UTI. Colistin monotherapy versus combination therapy for treatment of carbon resistant gram-negatives from Israel. The Danish POET study about using partial oral antibiotics for treatment of endocarditis. The ARREST study which tested the hypothesis that rifampin added to standard therapy improved outcomes in patients with staphoreious bacteremia. The comparison of piperacillin versus meropenem in the Marino trial. And the untesting of an algorithm versus standard care for staphylococcal bacteria. These were all published in 2018.

And if you go back to the original grid that I showed about the metrics of what patients -- what clinicians want and what they're provided, the registrational trials are shown there. All of those categories were addressed by strategy trials published in 2018. So there are differences in these things. We've alluded to it, but sort of to nail it down a bit

more refined, the purposes are different.

As I see the registrational trial, the primary objective is to make drugs available to practitioners and their patients. Strategy trials are primarily tasked with identifying the best means to use those drugs. The audience is different. At the end of the day, registrational trials, the audience of registrational trials are for regulatory agencies, with good reason.

By contrast, strategy trials are primarily tasked with an audience of the scientific community. The study design is different. I talked at length yesterday about the prototypical sort of ABSSSI and the complicated UTI designs and the reason that those well-established protocols are so important; test the drug, not the test.

Strategy trials, you're tasked with answering the clinical question. The consequences of failure are dramatically different. If you fail in a registrational trial, the compound is scrapped and the company may close. If you fail in your strategy trial, you publish in class one. The cost is

different and the complexity is different.

So what about networks, how do we fold this in, you know? And one of the challenges I think is U.S. site enrollment. Those were really nice presentations yesterday morning with some really clear data on the challenges of U.S. site enrollment. But it's interesting to remember, despite that fact, most clinical patients in clinical trials are enrolled from the U.S. -- this actually came from the FDA website and from a document from 2015 -- and it's not just a little, it's a lot. And that actually is not limited to other diseases, but is specific to infectious disease, as well, so that's point one. Most clinical patients in trials remain enrolling from the U.S.

The second point, though, is the U.S. site performance in ID clinical trials is variable. It really depends on what we're talking about and need to get into the details. So here's an example from complicated UTI -- these are data that were provided to me by a sponsor years ago -- with regards to the number of sites, the location of sites and the number of patients they enroll. This was for greater than a

thousand patients who were enrolled into two phase three identical design complicated UTI trials. Over 96 percent of the patients were enrolled outside of the U.S.; out of the 1,060 odd patients, 23 were enrolled from the United States.

Now, let's compare that to another study. So this is an adjunct therapy for staphoreious bacteremia that I presented at ECCMID last spring. Over a hundred patients in the phase two; 80 percent of those patients were enrolled within the United States. So then the question is, you know, what makes a study enrollable in the United States?

These are some criteria that I've kind of put together that seems to make sense. First of all, the patient's got to be there, and the sort of case in point in that instance is the ABSSSI experience. We saw yesterday that ABSSSI was the one indication in which the U.S. continues to increase its numbers, and that's largely due to -- I would argue, largely due to a handful of investigators out in the Southwestern United States in the San Diego region. And by contrast, in the R pathogens, unless it's ESBR

or MRSA, I'm sorry, but we don't have enough numbers in the U.S. to conduct a U.S. limited trial in those pathogens. And that's a good thing, we want to keep it that way. It's got to integrate with clinical practice.

The reason these single dose adjuncts were primarily useful or enrollable in the U.S. had to do with the fact that it didn't delay discharge. By contract, a complicated UTI, even in the modified guidance now, more or less mandates if it's an IV drug, the patient's got to be in the hospital for five days. That's not going to happen in 2019 in most medical practices, not going to happen in the United States. They're going out, and you see that in the reality that the numbers bear out.

The patient need to own the disease. Somebody alluded to that yesterday in terms of HABP/VABP, and I couldn't agree more. You don't talk to the ID physicians if you want to -- if you're going after a HABP/VABP trial; you talk to folks like (indiscernible). You talk to intensivists because they're the ones who have access to the trial. Now,

let's compare that to a complicated UTI. Who owns that? The ambulatory care guy, the ER person, the ID guy, the urologist; where do you get them? You don't, and so we don't. The pathogens need to be there, and we talk about that. ESBL and MRSA, yes, maybe; MDR, not yet.

And then I think patient comorbidity needs to be reasonable limit. The VRE example is, I think, a case in point to that from, you know, a decade ago.

So some other obstacles to U.S. site-based research. And, you know, insufficient trial volume to maintain site infrastructure. You get a cool study. It's, like, wow, this is a really cool study. Where's the coordinator? Oh, you got to go hire a coordinator. Okay. And then I got to assume that I'm going to be enroll enough to be able to pay that coordinator just to break even so that the division chief doesn't come back, you know, banging on your head.

So second point: the site work is academically undervalued. There are a lot of

academics in the area, myself included. I know we've all written a lot of letters of promotion. Ask yourself this question: how many times have you ever written in your promotion letter, this person should be promoted to tenure because they were a great site enroller. Don't answer it out loud, but think about it and let me understand. And I think we all probably know the answer: bureaucracy, crushing bureaucracy. And it's both, it's with local and broader.

All right. So what are some partial -- emphasis on partial -- solutions? I totally agree with John's point that this is a -- we're in a significant problem here and there's not going to be any single panacea. I think networks can help. Broadly speaking, networks really fall into three flavors: there's an observational type, and I'll use in the case in point, the International Collaboration of Endocarditis. This is one that's very near and dear to my heart but, unfortunately, is not relevant to today's discussion, so this is the closest I'm gong to get to talk about ICE today.

Registrational one, and I put the CTTI

initiative up here because I believe this is the first public or published forum in which John Rex and others were able to describe what I think was a really cool idea. And then strategy-based networks; I'm using RLG as an example. The targets are different: observational, describe the disease; registrational, new drugs; strategy, best management.

So in terms of the registrational trial, this is, again, this is the paper that John really led, along with Aaron and others, in terms of a vision for how to address the issue of getting new drugs to market. And I'm very pleased that it may have informed somewhat the RFP that Wellcome Trust put out, and I'm really looking forward to hearing more information about that from my colleague.

But suffice it to say, the concept would be these are a series of sites focused on a particular disease entity, the sort of three or four that get drugs to market, and that there would be ongoing enrollment with warm-based maintenance such that control patients can be enrolled during that time and you support your infrastructure.

The advantages to that is it gives you the right sites for the right trials and that it maintains a trial infrastructure. The disadvantage is critical mass, and then the sponsors may not use it. We're assuming that they will, but at least some of the meetings that I've attended in the past surprisingly, you know, at least gave me reason to suspect that they may not always use these networks. Strategies Trials Network, ARG is one example, there are others. So the purpose of this: design, implement, manage, clinical research in clinical practice.

The renewal is I hope will be starting in December. And the three emphasis areas are diagnostics, diagnostics, clinical trials, and relevant science. It's not exclusively focused upon interventional drug studies; and, in fact, I think diagnostics would be an important way to inform the practice as well. There's been -- you know, we've been productive in the last couple of years, and I'm incredibly proud of the efforts of our collective team.

So in summary then, the strategy registrational trial networks I see fundamental differ in regards to different functions. Registration, make new drugs, treatment; strategy, identify new treatment, best treatment, different audiences, registrational, FDA/EMA; strategy trials; the scientific community, different costs and complexity, so a lot of differences. But there's a similar need for high quality sites both in the U.S. and abroad, and I think that's probably an area where alignment can occur.

So I've sought to make the points that registrational trials are necessary, non-sufficient strategy trials, giving the people what they want, and clinical networks to improve U.S. participation in both, acknowledge funding. And happy to jump into questions, and I've been grateful for the time to speak today. Thank you.

DAN RUBIN: Thank you. Our next speaker is Pam Tenaerts. She's the Executive Director at the Clinical Trials Transformation Initiative, where she works closely with its executive committee

to develop and implement strategies to accomplish CTTI's mission.

PAM AENAERTS: Good morning, everyone. Happy to be here and thanks for inviting me. So today, I'll be talking about what you can do to enhance enrollment strategies. We heard a lot about enrollment being an issue and clinical trials have that specifically, and I'm here to represent the work of the groups that have worked on this. These are my disclaimers.

And so, just a little bit about CTTI. So we're a public/private partnership, co-founded by Duke University and the Food and Drug Administration, and we have a mission to develop a drug of optional practices that will increase quality and efficiency of clinical trials. We basically come up with recommendations on how to do clinical trials better, and that's what we've always done until 2012, and I'll explain a little bit more about that.

Importantly, we involve all the stakeholders, and we're a membership-driven organization. But we involve a lot more organizations

than just our members; it's whoever is needed for the activity that we're working on. We're evidence-based in that we use a large social science team to help us go from opinion, because by gosh, when you get a group together on whatever the topic is that they're passionate about, everybody on that group typically knows how to make it better. It's not necessarily evidence-based, so that's what we use our social science team. And we've been lucky enough to have impact in policy documents and things like that.

I wanted to highlight a little bit about some of the work we do. We split that up in six areas of focus: quality, talked a little bit about quality by design yesterday about there's resources there; patient engagement is also a big focus of CTTI, which you might have suspected based on my comments yesterday. And then we also help make sure that there's investigators and sites to do the work, even though as you move forward, there'll still always be sites, but that model might be changing a little bit as we go to site lists and things like that.

Mobile clinical trials really could

have fit into a novel clinical trial design, but that's just where we use mobile technologies. In our case, it was after consenting, so not to recruit, not to consent, but to capture end points. And then within all the clinical trial designs, you have the antibacterial drug development program; that really is the topic of discussion today. We also have a lot of recommendations on the use of single IRB. IRBs came up yesterday too as, you know, the variability in approval processes, getting different consents, things like that.

So this is our antibacterial drug development. So before this, I would have said CTTI does process improvement on clinical trials, and then, oh, and we also do antibacterial drug development. FDA asked us to work on this in 2012, and we developed a couple of projects. There's a lot of you in the room here that have helped on these projects, so I want to thank you for that. We could not do this without the people that help us, free most of the time, so we really appreciate it.

We did three big topics,

(indiscernible) HABP/VABP trials, and the reason we picked that one is the reason we're here today, and it's how difficult it is to do these trials. And if we tackle the hardest issues, maybe some of that work could then apply to the other trials as well. Pediatric trials and areas of unmet need.

So today, I'm going to talk about the HABP/VABP studies we did, but we also did work on streamlining protocol elements and data collections. But the overwhelming thought was it's all great that we can streamline our protocols; that'll help us nothing if we don't have patients in the trial. So that's why we went back to HABP/VABP studies and try to figure out how to do these better, which included two portions, a risk factor study and some formative research.

We would really, really, really, really like to do an early enrollment of clinical trial, which is where we test the methodology. And Vance is laughing because we've shopped this around to a lot of people, but we would really like to do this. Haven't been able to do it yet, but it seems really if you

want to be evidence-based, you need to have evidence on your methodology, and this would prove it for early enrollment.

So why did we talk about early consent, is because there's an early need to treat HABP/VABP; we've learned that today. There are few ongoing or planned HABP/VABP. But this must be cyclical because when we started in 2014, there were very few. When we came out with our work on early consent in 2017-'18, all of a sudden, there were a couple, and we were, like, oh, we don't want to compete for patients that should go into real treatment trials with our methodology trials, so we put it on hold a little bit. And then now, there's like there's no -- almost no HABP/VABP trials again.

So anyway, so what we talked about is -- and I have that project. And the other thing is, you know, we all know that enrollment rates are, you know, abysmal, they're really low, and the cost of enrolling a patient is really high. We also did work with Tufts University on the cost of trials, patients in HABP/VABP trials and it's about $100,000 per

patient, this is not trivial, and this is for a patient enrolled.

So we looked at that and we had a lot of discussions about this. And at one of the meetings, the patient said, well, why don't you ask me to be part of an HABP/VABP study when I'm still there, when I still can consent, and we tossed that around. But Vance really turned it into something that was manageable where we could figure out risk factors. Because if you want to consent a lot of people up front, you probably don't want to consent a hundred to get one patient that ends up with a pneumonia, but something reasonable that somebody could pay for as far as creating consenting and that. And so, we needed to figure out how to get there.

So we wanted to do this before they're critically ill because we hope that the patient, the participant could actually participate in the discussion. The family could be there. So by the time the patient down the road has HABP/VABP, is potentially on a ventilator, nobody has to sort of worry, like, would grandpa really want to do this or

not. You kind of know because you were part of the discussion.

And we then figured out how to actually do this. So we did ask to do a demonstration study; we haven't been able to do that. So the reason we went to early consent is in our streamlining HABP/VABP work, we did a lot of project team discussions and focus groups who'd experienced court leaders, and it really was a challenge to enrolling. The 24-hour timeline is the big bugaboo here. And we sort of appealed, like, well, can't we make that 48 hours; for a while we were on the track of let's extend it out to 48. And there's been scientific reasons why you shouldn't be doing that, so the 24 really is sort of something. But then this early enrollment could really help with that.

So even when the patients -- you know, this is prior to effective antibiotic therapy. So if the patient is identified before 24 hours, it's difficult to conduct all these things that you need to do -- consent, labs, study, drug availability -- in that 24 hours. So how can you do this by beginning

consent before the HABP/VABP develops? So you would approach and consent patients at high risk before the 24 hours, before the antibiotics are started, and before the symptoms develop.

So you would do all that ahead of time, and then you would enroll the patient the minute they would actually develop the diagnosis of -- when you would normally start your planned antibiotics. This is where now, if you had consented early, this is where you can now start them in the study instead of starting them on the regular antibiotics that you would normally starting them.

And, like I said, we had planned to conduct a study, which we have not done. If anybody wants to talk about that, I'll be here until 1:00.

So if you want to think about doing this, you have to first find the patients that you would want to consent early, and then figure out whether this is even acceptable or feasible. Because when this plan started coming together, there were a lot of people that said, oh, the IRBs are never going to go for this; we heard that from a lot of people.

So we wanted to do both: can we find a set of patients that you could consent that makes economic sense for a sponsor to pay for those consenting and screening procedures; and then, how do you get that done, like, how do you make that acceptable and feasible both for the participants, the science, and the IRBs.

So we did some preliminary research; it's really determining the population that you should approach early. And after a lot of planning, we sort of came up with following the oxygen in the ICU, because we figure that running -- there's plenty of patients that develop pneumonia outside of the ICU, we're very well aware of that. But, I mean, logistically, that's just not an easy thing to start, you know, go after. The ICU is contained, so we figured if we could find a population there, that would make sense for the coordinator to do that.

So we identified, we did a risk factor study in the U.S.; the EU participated with a comeback network. And I heard somebody talk about they were participating on our studies as well. And then we

also did pediatric sites because they also have that issue.

So we looked at -- we required invasive or non-invasive ventilation, so this following the oxygen idea, and you had to be receiving antibiotics for suspected pneumonia. So in total, we enrolled 7500 patients. This is -- I'm going to talk a little bit more about the U.S. one because we're a little further along. Europe took a little longer, and then the pediatrics people are working on their section.

So this study in the U.S. lasted about eight months in 2016, so this is a pretty short enrollment period. And we identified patients that were high risk, so they had to receive more than 12 hours of treatment with invasive and non-invasive mechanical ventilation or high levels of oxygen within the past seven days. And of the 7500, 4632 were high risk, identified as high risk; 1400 -- so 1500 patients were actually treated for pneumonia; and of those, 539 met the FDA guidance criteria for HABP/VABP.

So let that sink in a little bit. How

many of you have gotten, you know, those cards about my hospital has not had any hospital-acquired infections? So this kind of shows that HABP/VABP is alive and well, so this is another thing to think about. But what we basically found is, you can identify patients that basically if you'd consented 10, one of them would have pneumonia.

So like I said, we could quasi-predict who might get pneumonia, but you could enrich that sample with other things based on the ICU admission, what had they received. So you can enrich that sample and make it a little richer; documented aspiration risk, admissions source have received of systematic, and the bacterial in the last 90 days were additional risk factors. So if you combine that, you could really identify prospectively patients for an early enrollment strategy.

So, okay, that part is done. We can find the patients reasonably that a coordinator could work with to see if they were to develop pneumonia.

How about the other thing; is this acceptable and feasible? What concerns would an IRB

have about this early enrollment strategy? How burdensome would it be to the trial investigators and the study coordinators, because you don't want to make it a whole lot harder. And how would patients and caregivers feel about enrolling in a clinical trial, and they don't even have the condition for what's going to happen. Not only will they not have the condition, but they're pretty sick already.

So here is somebody who is really sick coming into an ICU and somebody's going to come to talk to them about, oh, and by the way, not only do you have everything you have, we think you're at high risk for pneumonia. So this is something that people thought may not go over very well. And if we did that, what would patients want to know about this approach so they can make an informed decision and not sort of a screaming and run away, so we did formative research.

So we did in-depth interviews, and then those people were then also asked to participate in two surveys. The interviews dealt with the acceptability and preferences for components of that

strategy and kind of figuring out what topics we should describe in the early consent. And then with our online survey, we delved deeper into that into really getting at the sentences we should be using. And then the last survey was final agreement of the language to include in the consent.

And the people that participated in this research, in this formative research were patients, caregivers, investigators, study coordinators, and IRBs, so we kind of tried to get everybody that might have an opinion on this.

So what the patients and legal representatives thought is that this is really not that hard. I mean, they could accept this very readily; that early consent and enrollment strategy was overwhelmingly accepted. They found it acceptable that their charts would be monitored, because that would come with that obviously before the acquired the pneumonia; they could understand the consent information before the -- you know, before they would be diagnosed with HABP/VABP, in this case, and they would be willing participated under early enrollment

trial using approved antibiotics. They felt a little more iffy about the newly ones, but, I mean, we could probably get there, but some others don't work.

When we asked investigators and IRB, they kind of felt that this early enrollment strategy, you know, may work. The investigators thought, like, oh, this could maybe work. And it may improve efficiency of the trial conduct for HABP/VABP. And we hope to think that this also shows that we might potentially be able to do this in other conditions where, you know, there's sort of acute things that happen in a chronic patient.

None of the IRB members raised concerns about the early enrollment strategy. So, honestly, they said this sounds pretty straightforward, it doesn't sound like it's going to cause a great deal of concern. And, you know, there would have to be a discussion about the possibility, the percentages, the chances that might happen, but not looking at this as something unusually concerning.

What needs to be in that consent ahead of time? So they would want to rationale for the

early enrollment strategy; they would have to have explanations about non-inferiority. As you can imagine, this was not the easiest concept for anyone to understand. And so this was something that you could actually take your time with and close the patient or the perspective participant potentially. And their family were there, so this is not like, you know, your family member is on a ventilator and you're freaking out a little bit.

Reassurances as to what would happen if the studied drug might not be working. We asked them how they would explain that information, and then we used that information to develop and obtain agreement on the text that should be used in the consent, and we have finalized texts for all of the issues that they mentioned in the consent.

This is available in a publication, and the survey data and the final consent is also going to be published later this year. This got caught up a little bit and, you know, we talked yesterday about pre-prints a little bit. And we've starting doing pre-print articles because feel that it takes too long

to publish, and we are not the hot topic in many cases. You know, we've tried to, you know, go to ID, you know, all these -- we just find it hard -- it's hard to publish these things. We finally found a place, but we're going to go more to pre-print. We haven't done it with this project yet, but if we had, it would be available already.

We also think that you could use this in other diseases and applications and, you know, other ICU acquired infections or infections that tend to occur, UTIs, (indiscernible), you know, sepsis, other chronic conditions with frequent exasperations like I mentioned earlier, or conditions in which patients have periods of decisional incapacity where for times, they can't make decision maybe so you can talk about this ahead of time.

I would say, though, that when we talked to patients and IRBs about this, while they happily wanted to consent ahead of time, they would like a little heads up when it actually would start. Like, it wouldn't be -- you know, you wouldn't know when you were getting into the research study. So

there would be some type of, you know, hey, you remember what we talked about; we're going to start and this is the specific study we're going to put you in.

So this may improve efficiency of clinical trial conduct for HABP/VABP, and it was overwhelmingly accepted. Prospectively identifying these patients requires high level of following the oxygen basically, and you could enrich it. And we are developing tools to assist with this trial planning; we're going to come out with the consent language. We'll publicly share the risk factor data.

We're actually going to put all that clinical data online with levers that if you want to look at your inclusion/exclusion criteria, you could sort of move them around on whatever the topic is you're measuring to kind of see that in our population, how that would have affected enrollment, whether you would exclude patients or not. And we're also creating a trial planning -- that's the trial planning tool I talked about, and that's it. Thank you.

JANE KNISELY: Our next speaker is Chibuzor Uchea, a Senior Officer in Wellcome Trust Drug-Resistant Infection Priority Program, working on a range of projects focused on the development of new treatments and diagnostics and improving the efficiency of clinical trials. He joined Wellcome in June of 2019 after three years of working in healthcare consultancy.

CHIBUZOR UCHEA: Good morning, everybody. Thank you for the introduction. I'd like to also thank the organizers for inviting me here to speak. It's a pleasure to be here.

So as was mentioned, I'm from the Wellcome Trust. And for those of you who are familiar with us, we're a charitable foundation that supports research to improve health around the world. We have a mission of taking on eight global health challenges, making an impact by meeting the response through our priority program.

The drug-resistant infections program is one of these, and it aims to use Wellcome's funding, convening, advocacy, and influencing power to

help lead the global response to antimicrobial resistance. Our program has a strategy that's based on four pillars, which represent key areas of unmet need, but also opportunities in which Wellcome are well placed to be able to make an impact.

The four pillars of this strategy can be seen here. I would like to call your attention to the two on the right-hand side, which are most relevant to this workshop: that's the development of new therapeutics, and the acceleration of clinical trials.

We have a vision of developing a pipeline that's sustainable to develop antibiotics, diagnostics, and vaccines for infectious diseases to protect local public health. This is reflected by our commitment to CARB-X. As Aaron mentioned yesterday, we're a funder and it remains our greatest investment of $155 million across five years.

And also our work with the (indiscernible) program in collaboration with the Innovative Medicines Institute. As the pipeline strengthens, clinical development will become an even

tighter bottleneck, and we're committed to developing initiatives that will help accelerate (indiscernible).

We also have a strong policy and advocacy focus, and we interact with key decision makers in all of our activities as we recognize that antimicrobial resistance has an urgent need for global action.

So the current funding model of clinical development in infectious diseases is complex and burdened with a range of inefficiencies that slow down the commercialization of antibiotics. This is especially true in low- and middle-income countries, and is reflected by the disparity between the number of trials that were run in these countries and the burden of antimicrobial resistance.

A key driver of this is the ad hoc funding nature of individual trials, whereas sponsors custom build a single use network of trial sites, spending considerable amounts of money building capacity and infrastructure and training staff on individual protocols. At the end of these trials, the capacity is disbanded, and there's a loss of

infrastructure and expertise within sites as well. Subsequent funders then come in and spend considerable amounts of money reactivating sites, redeveloping infrastructure, and training new members of staff on subsequent protocols.

Another key issue, as mentioned yesterday, are issues with patients recruitment. And because of the clinical nature of infectious diseases, there's a very narrow window for recruitment. And also, as described yesterday, it's not only impossible to be able to move patients to various different sites.

Also, if the focus on studies in high income countries, there's difficulty in finding patients with -- some of them more difficult to treat infections, especially if they're multi-drug resistant and extensively drug-resistant indications. Whereas, there were a much wider pool of patients in low- and middle-income countries with these extensive drug-resistant infections because of the increased burden of antimicrobial resistance.

And in low- and middle-income

countries, there were also issues with the (indiscernible) of trial site quality, which affects the quality of the data that it produced, and also the requirement for individual studies for each indication exacerbates the issues with our funding model.

So what are the proposed solutions for these problems? We have a vision of accelerating clinical development in this space using a two-pronged approach that will address the various different barriers.

The first of these is the development of international clinical trial networks, which will strengthen the clinical trial capabilities within low- and middle-income countries, alleviating the inefficiencies of the current ad hoc funding model and vitally provide access to large populations of patients with key drug-resistant infections.

We have a vision of these networks being able to run multiple studies from different funders simultaneously.

Our second approach is the development of platforms for innovative trial design, especially

the use of continuous master protocols. This will allow control group sharing between trials and reduce costs and burdens.

These two aspects are independent, but also complementary, and joint implementation will provide even greater efficiencies.

So the first of these is the international clinical trial networks. And we're working towards developing a pilot network that's going to be entered in Southeast Asia. It will start as a flexible scalable regional network of high-quality sites, which it will be building on existing capacity, and vitally will provide rapid access to the Southeast Asian population with a high burden of antimicrobial resistance.

Our model will be (indiscernible) with other regional networks, which will be vital for us to be able to broaden our base, and also will provide access to key expertise in infectious diseases internationally. Another key aspect of this is that we would be able to leverage key learnings from the development of other clinical trial networks.

Our business model will be orientated to be able to engage with the private sector, which will allow us the support by investigating initiated studies and registrational studies. We've identified our nucleus of sites who will be our founding members, and we aim to launch the network around an initial trial, which will be used to test and help develop the network and also inform scale.

An effective clinical trial network will provide a range of benefits to a number of stakeholders, including sponsors, investigators, and ultimately patients. For sponsors, there will be reduced costs of conducting trials through warm-based benefits, and also will facilitate power around follow on studies and optimization studies for strategy trials as (indiscernible)'s ability.

For investigators, crucially will provide access to keep populations with high burden of disease, and also will improve the quality of trial sites in the region, improving the data that are produced and helping to produce better studies.

And ultimately for patients, it will

increase the speed at which treatments have become available, and also potentially make them cheaper through reduced developmental costs.

We've recently commissioned a piece of work, which is looking at how best to structure the network to maximize our possibilities of success. And the process we engaged in with a range of stakeholders from industry, clinical trial experts, as well as other members from other clinical trial networks.

Our own grading model will be with the network secretariat, which is responsible for managing and coordinating central functions, and also it will be guided by a steering committee which will provide strategic oversight. The responsibilities and roles of the secretariat include pipeline management and regulatory engagement and market access, which will be vital for the long-term sustainability of the network, as well as administrative functions.

We've also identified key areas which are vital for startup, and also others which are more important for the long-term vision of the network. These include geography; we've identified our founding

members, but we are also looking to expand, and we can add additional sites that match the site quality criteria to the network. We're currently engaged in conversations with sites and institutions in India, which will further increase our geography and vitally access to either ratable patients.

The network will start by looking at, investigate and initiate treatment trials within drug-resistant infections, but we'll scale up to include studies from sponsors in a commercial sector and also looking at infectious diseases in general.

We're currently working to finalize our governance structure, which will be ready at the end of the year, and then we will start to recruit for our secretariat function, as well as identifying and developing the initial clinical trial, which will be across next year.

There are a number of potential challenges of developing international clinical trial networks, and these include a strong and robust and transparent governance structure, as well as alignment on direction within the network.

Another key aspect is potential for underuse of the network between studies. This is why to prevent a lot of times sustainability of the network. And the work we've commissioned has looked at the pipeline. And although there are considerable studies that can come into the network, there's also the opportunity in which there will be down time during -- when registrational studies are not available. For this reason, we will work for our network to be able to have the capacity and the expertise to run additional studies, especially more complicated studies, including pediatric studies and smaller optimization studies.

We're currently working to build out our secretariat function, which will provide support for the network to be able to address these challenges.

So the second approach is the use of innovative trial design. In 2016, we held a workshop assessing the benefits or potential benefits of using clinical trial networks for antibiotic development, which a number of you would have attended. An

innovative trial design through the use of continuous master protocols was highlighted as a key tool that could further accelerate the efficiencies that are provided through clinical trials, and a recommendation was made that this should be further explored.

Since then, we've commissioned some work to assess the feasibility of using continuous master protocols within clinical trial networks. And in the process, we engaged with key stakeholders in the industry, regulatory bodies, and John Rex has also been a great guide in our thinking. The continuous master protocol will provide even greater efficiencies by allowing the sharing of control groups within clinical trials with the same indications, which will reduce the enrollment periods and reduce the costs as well.

There's the possibility to use concurrent control groups within the same time periods have also potentially non-concurrent controls and possibly control data, historic control data from previous trials. And the use of adaptive randomization allows the flexibility of entering and

exit of compounds throughout the process, mitigating the irregular startup of regulatory studies.

Our feasibility assessment has shown that such a platform would be especially beneficial for undertaking regulatory science studies and later, expansion studies. We have identified sufficient demand for this among the stakeholders which we've engaged with.

The use of concurrent controls has been supported in principal by regulators. However, there remains considerable concerns and issues around the use of long concurrent controls as we discussed yesterday.

We have identified adult HAP/VAP and CAP studies, as well as pediatric studies as those that will benefit most from the use of the continuous master protocol, and we're currently exploring with other funders whether this type of platform is investible.

So in conclusion, we believe that the reason strengthening of the pipeline really demands a more efficient clinical development process. And the

initiatives that we've mentioned here today around the use of clinical trial networks and continuous master protocols will provide initiatives -- will provide additional benefits from a scientific, financial, and developmental perspective.

Importantly, conducing these in low- and middle-income countries will provide access to key populations of patients with drug-resistant infections.

DAN RUBIN: Thank you very much for the introduction. I'll try to keep my remarks brief to leave time for other speakers and for Q&A and to keep us on schedule.

One thing I might say is that after this stat session, we have some Q&A scheduled for all those talks. We didn't have any on the agenda for Q&A for the talks we just heard this morning. But maybe after all the stat speakers, we can have a Q&A for everything so far today and then go into the panel discussion later this afternoon.

What I'd like to do is provide some high-level comments about three important statistical

issues in anti-infective registration trials, which I hope will frame the discussion for this session. These issues are endpoints, borrowing information across body sites, and considerations for carbapenem-resistant pathogen studies.

With respect to endpoints, anti-infective registration trials have most commonly used binary endpoints in which each outcome is classified as a success or failure. The definitions recommended in our guidance document vary across the disease types. And hospital-acquired and ventilator-associated bacterial pneumonia, the main endpoint recommended in the guidance is all-cause mortality at a fixed time point following randomization between day 14 and day 28. In trials of intra-abdominal infections, failure is based on a clinical determination of failure. And in complicated urinary tract infections, the primary endpoint is a composite based on both clinical and microbiological components.

Professor Evans will describe in more detail how binary endpoints could potentially be improved by moving to endpoints with finer gradations

of success and failure. For instance, one could consider an ordinal endpoint with the three levels of death, survival with major morbidity, and survival without major morbidity. The main advantage is that this may lead to more informative comparisons when reasons for failure are not all lumped together. And another advantage of this approach is that it can increase statistical power because it's making use of more information.

One consideration is that levels should be chosen so that differences will not be solely driven by effects on components with minor importance or solely by safety.

As a hypothetical example, if you wanted to show that a new drug was superior to say colistin and included reversible renal toxicity in an ordinal outcome, one thing to check would be whether a new drug could be superior to colistin using this type of outcome scale even if it was say increasing mortality. And this issue can be addressed to some extent by assigning weights or utilities to different categories.

My understanding is that these forms of DOOR and RADAR methods have been used by ARLG, have mainly been proposed and implemented for superiority trials.

Dr. Lewis will discuss borrowing information across body sites of infection. Some statistical methods such as Bayesian hierarchical models have been widely used for combining information from different sources, both in clinical trials and many other types of applications, but largely not in the anti-infective space.

The models would attempt to provide an integrated or synthesized analysis of patients with different infection types.

One way to think about information borrowing is that it tries to reduce noise for estimating treatment effects in any one body site by bringing in other relevant data. So for instance, suppose you have a very small number of patients with complicated intra-abdominal infections so that the estimated treatment effect will have a lot of variability. If there's a larger number of patients

with complicated urinary tract infections who have been treated with the same drug, then incorporating them into the analysis could in theory reduce this variability.

Of course the issue with doing this is that there is a reliance on modeling assumptions about the degree of similarity between infection types. So for estimating treatment effects in cIAI, you would no longer have the protection of a fully unbiased, randomized comparison.

One remark to point out is that FDA has at least informally used some forms of information borrowing in assessing anti-infective drugs even without use the full Bayesian machinery. For instance, an NDA can be based on a single successful Phase 3 trial for patients at one body site with supportive evidence from a related disease.

Registration trials do combing heterogenous patients, such as studies of both hospital-acquired and ventilator-associated bacterial pneumonia. And our unmet need guidance also accepts pooling body sites for superiority trials, and

registration trials in resistant pathogen studies have combined nosocomial pneumonia and bloodstream infections.

There are a few issues that FDA is likely to take into account when reviewing proposals for an integrated analysis of body sites. One issue would be statistical operating characteristics if treatment effects differ.

For instance, if there is a proposal to combine HABP/VABP, cIAI, and cUTI, and in truth the drug doesn't work in HABP/VABP, how inflated is the chance that the design will lead to a false conclusion of efficacy in HABP/VABP?

We may also consider previous history of discordant results of cross-body sites, such as with daptomycin not working well in pneumonia due to inactivation by pulmonary surfactant. And for deciding which data to integrate, this wouldn't just be a statistical decision; there would of course also need to be a clinical judgement that the infection types, pathogens, and endpoints make sense to combine in an integrated analysis.

The last topic I'll touch upon are carbapenem-resistant pathogen studies, which are related to Dr. Dane's upcoming presentation. These studies are commendable because they do directly address questions most closely related to unmet medical needs.

The main challenge is that enrollment in randomized trials has been very difficult, as you have heard over the course of the workshop, due in part to the rarity of the pathogens.

I will just briefly note that there are some government or academic randomized trials that have compared combination therapy to colistin monotherapy in patients with carbapenem-resistant Acinetobacter baumanni infections, which have had some degree of success with enrollment and together have enrolled about 650 patients. And that's not including the OVERCOME study we heard about yesterday from Dr. Dixon. But I've referenced a metanalysis on this second bullet.

If moving beyond randomized comparisons to external controls or non-randomized comparisons,

the key challenge is being able to control for confounding in patient populations with many comorbid conditions. Dr. Dan will go into more details about some of these considerations surrounding randomized and non-randomized evidence.

Our unmet need guidance also discusses pathways that have now been used by several sponsors based on another type of information borrowing. And this is where a non-inferiority trial is conducted in patients with susceptible pathogens, potentially with a wider than normal noninferiority margin. The labeling is then for patients with more limited treatment options with the constraints of the data package communicated in the labeling.

The main uncertainty of this approach relates to differences between patients with susceptible pathogens and the less well-characterized group with limited treatment options.

One question to consider about this form of extrapolation is whether the discordant directionality of numerical results in the carbapenem-resistant CARE trial of plazomicin and the CREDIBLE-CR

study of cefiderocol could have been predicted or explained based on all the preclinical information or from the successful non-inferiority trials conducted for each drug in carbapenem susceptible complicated UTI.

The caveat is that the numerical results in these two trials of carbapenem-resistant infections had statistical noise from the sample sizes. But the references in this bullet are to our advisory committee materials if you'd like to think more about this question.

The last point I'll make about carbapenem-resistant pathogen studies is that folding these patients into standard noninferiority trials but with more flexibility in the active comparator would follow the template used for previous types of resistance.

For instance, if evaluating a gram-positive drug, there wouldn't be a separate trial for MRSA. Patients with MRSA instead would most likely simply be included in a noninferiority trial of skin infections or pneumonia or bacteremia and there would

be provisions for patients in the control group to receive an agent with activity against MRSA.

The uncertainty or one of the main uncertainties that I see with using this approach for carbapenem-resistant infections is that any recently-approved active comparator might itself have been studied using relatively streamlined data. And this could limit the interpretability of noninferiority conclusions if there were remaining unanswered questions about the active control. And one term that has been used for this type of risk is biocreep.

Here are references for several of the topics that I've mentioned, and thank you.

JANE KNISELY: Thanks. Our next speaker is Dr. Roger Lewis. He is a Professor and Chair of the Department of Emergency Medicine at Harbor-UCLA and a Senior Medical Scientist at Berry Consultants. And he will be speaking about Bayesian adaptive clinical trials.

ROGER LEWIS: Great. Thank you very much. It's a pleasure to be here. Thank you to the organizers and the audience.

So I'm going to be talking about statistical approaches for antibiotic trials and posing the question, are we answering the wrong questions? But in posing that question, I hope to follow Dr. Rex's recommendation to actually give an alternative.

These are my disclosures.

So the take-home points are that the statistical and overall design strategies that are commonly used for antibiotic trials really should directly inform the clinical uses of products if they are approved. In light of how antimicrobials are actually used, I believe they should support antibiotic stewardship to the degree available or accessible within the regulations, and also inform regulatory approvals and labeling.

I put these points in this order because I believe that if you have trials that directly address the appropriate clinical use of an agent, that will naturally inform regulatory decision-making. But I believe as currently deigned, many trials in antibiotics do not meet these goals. They

are overly-narrowly focused. We heard amazing ratios of screen-to-enrollment numbers yesterday. They may predictably undermine some stewardship efforts, and they risk missing important benefits due to delayed initiation of agents and contamination of patients 'outcomes by prior treatment, but they could address the goals.

So let's briefly go through an anatomy of an antibiotic trial as sometimes carried out. And I hope that I don't step on anybody's toes this way.

The patient presents commonly to an emergency department, which is where I work clinically. They have signs and symptoms of a serious bacterial infection. They are hopefully briefly evaluated, and empiric antibiotic therapy is started very early because there is an important clinical and compliance imperative to do so. There is a small loss of patients if it's determined that treating the disease is not in the patient's interest and consistent with goals of care.

Some time after that, they are admitted to the hospital. And it's commonly in the in-hospital

setting in which they are first evaluated for an antibiotic trial. At some point it's confirmed that the intended single site of infection for the trial is in fact the infected site. Culture results come back that meet the criteria. At some point they're actually randomized. And if they are successfully enrolled and randomized, the investigational medical product is begun, and at some point there's an outcome assessment.

There's a long period of time where they are receiving empiric therapy prior to randomization, and lots of things happen during that time that may dilute the true treatment benefit of the investigational product. So the limitations of this approach include little alignment with clinical practice, a narrowly-defined study population, relative late initiation of the investigational agent, and the evidence that is gathered and therefore the labeling that can result from the trial addresses a single infection site. So how can we address some of these challenges?

Another way of pointing out the

differences between the trial structure and clinical care is to look at the timing of the treatment, the populations that are included, the motivating event that leads to the use of either the investigational product or the approved product in clinical use, the types of infections that are treated, and whether we're looking at non-inferiority or superiority. All of those things are relatively different between the clinical use and the way we study these agents. And those can increase the risk that we produce data that doesn't allow us to make the best-possible regulatory decisions.

So clearly new agents are needed to treat challenging organisms across multiple sites of infection. And in clinical practice, antibiotics with demonstrated penetration into the infected site and an appropriate antibacterial coverage are routinely used for infections at those sites, independent of any site-specific supporting data, and certainly independent of labeling.

Now, there has been, as we all know, problems with a surprising lack of antibiotic efficacy

at specific sites, most notably lung. And that's generated concern regarding sharing efficacy data across anatomic sites. When I talked to people who are deeply embedded in this area, there is almost an air of PTSD regarding the trials that resulted in these surprising findings. And that experience should not cause us to abandon the common clinical reasoning that an antibiotic that works well in multiple sites is more likely to work well at another site than an antibiotic that does not work well at multiple sites, but we also need to do so with our eyes wide open.

So the proposed strategy I'm going to discuss is a platform trial with enrollment timing and antibiotic initiation designed to match clinical use, careful integration of information across body sites -- and we'll spend quite a bit of time talking about what careful means -- that simultaneously addresses both noninferiority and superiority, because I think that choice is a false dichotomy, and may achieve additional efficiencies through the platform trial structure.

So the proposed trial would look

something like this in terms of flow. The patient presents with the signs and symptoms of a serious bacterial infection and is at risk for but does not have a documented infection with a highly-resistant organism. The patient is evaluated for participation in the trial as part of their emergency department evaluation, and the first antibiotic they receive is the randomized investigational product. At some point after that, they're admitted. And at some point later we get the results of the cultures, additional diagnostic tests that verify the actual site of infection and the resistance pattern of the organism or the fact that there is no isolate identified, and then the outcome is assessed.

And the question is how could we use data from this type of simpler structure to appropriately inform both regulatory decision-making and clinical decision-making regarding use of approved products. The advantages is that this aligns with clinical practice, there's a broad study population, there's early initiation of the investigational agent, and you will get information that informs you about

evidence across resistance patterns and across clinical sites of infection.

So first, what do we mean when we talk about a platform trial? So a platform trial is an experimental infrastructure that's intended to evaluate multiple treatments, often for a group of diseases. For a group of diseases here, you can think of different infection sites or different underlying pathogens. And the platform trial is intended to continue beyond the evaluation of any individual treatment. The treatments are often used in combinations. The trial explicitly incorporates the idea that the diseases that are being studied are similar to each other but not identical, and there's a dynamic list of treatments that are available.

So this terminology is borrowed from the oncology world. And I first want to make a distinction between a master protocol and a platform trial. The term master protocol is a very broad term that may include simply the use of very standardized clinical processes to increase the efficiency of a clinical trial. But the separate experiments within

the master protocol may be inferentially separate. You may be asking separate questions about completely separable populations or treatments.

A platform trial in contrast in my view incorporates some sort of statistical sharing of information to increase the efficiency of the inference given any particular set of data.

An umbrella trial is a term used then you're testing multiple different drugs for diseases that occur at the same site. And you may think of that as multiple studies of a pneumonia with different underlying etiologies or organisms.

A basket trial is a trial in which the disease shares a commonality that is a target of the treatment, but is in different anatomic sites. That would be for example sharing of information across body sites when you're using a single drug that's active against a particular class of organism.

There's a very nice review article written by Drs. Woodcoock and LaVange that goes through this terminology and provides a very nice background.

So what is the proposed strategy? The patient presents with an infection in either the lung, the urinary tract, or the abdomen. But if you're interested in gram-positives, you could replace this figure with the places where gram-positives are of most interest and challenging. And from each of these sites, we obtain bacterial isolates, and we either identify them as being sensitive to the standard of care, resistant, or unable to be cultured. We want to demonstrate superiority to the standard of care among those patients who have resistant isolates. Non-inferiority to the standard of care control arm in those that are sensitive to the standard of care. And I put a question in there what you would do with those who have missing bacterial isolates. And we want to take advantage of the multiple body sites.

Over time, the agents that are being considered could also change. One could start with a control arm and maybe one, two, or three active arms. And over time, additional arms can become available. You may decide that based on additional safety data or other preclinical information, it becomes appropriate

to combine arms into a single arm. That's the A plus D. Arms may be discontinued for harm or futility at any point. And if one of the arms is demonstrated to be superior, that can seamlessly become the new standard of care in a platform trial without redesigning the trial. This addresses a point that John made about what happens if you identify a new standard of care; do you suddenly undermine your ongoing research efforts?

This approach has been modeled numerically through work that was supposed by the ARLG and others and involved a wide variety of collaborators. And in general if you consider this approach and you make reasonable but particular assumptions about the number of drugs, the fraction that are positive or negative, et cetera, you can save an average of something between 40 and 60 percent in sample size per answer you get regarding the superiority of the drug and resistance, isolates, or non-inferiority in sensitive isolates. And this efficiency comes from the shared control, the sharing of information across body sites, and the fact that

the study can seamlessly drop an arm for inferiority or for futility. And the assumption is that another drug is available to be tested. So that's a very particular assumption.

I want to specifically comment about the sharing of information between body sites. So we all know that antimicrobial agents are likely to have different effects at different body sites, for a wide variety of reasons. The question is whether the degree of variability in the treatment effect is clinically unimportant and we should think about the treatment effects as similar across body sites, or it's clinically important so we should think differently about the use of that drug across body sites.

So the question is how can we address both the possibilities that the treatment effects will be largely similar and that the treatment effects will be highly disparate in a single statistical model.

Clinicals do this sort of borrowing all the time. We take data that have been obtained in particular types of patients and we apply them to

others. But we do this in a way that is not well-documented and is not quantitative. Here we need a strategy that is completely prespecified, that is statistically rigorous, and whose operating characteristics can be evaluated numerically.

It is very tempting to try to avoid the reality that similarity is not an all-or-none thing. We tend to behave as if it is. So we will group all types of infections together if they were included within the inclusion criteria for a trial, or we will exclude them if they were not. We will include, for example, all different sorts of isolates if they were at least represented at all within the trail. Although we certainly don't have enough evidence to make independent estimates of the efficacy of the agent across each species of bacterium. For example, it was isolated.

The all-or-none approach puts us at tremendous risk of failing to identify subgroups that do experience different treatment effects of complications if we start by combining them all together. So we miss real differences that exist.

And we also fail when we separate them to recognize compelling circumstantial evidence of treatment efficacy. For example, if the drug works across four or five different infection sites, the chance it will work against the fifth or the sixth infection site is greatly increased.

So the approach that we'll talk about is a Bayesian hierarchical model. And I've underlined the key conceptual strength of this approach. The Bayesian hierarchical model shares information across subgroups, or in this case, body types, to the degree that is justified by the consistency of the information. So, contrary to something that is commonly stated, that you have to make an assumption about how similar the treatment effect is across body sites, instead what you do is you create a model that can learn how consistent the treatment effect is across body sites and use a greater degree of pooling if you see consistency and a lesser degree of pooling if you see heterogeneity of the treatment effect. So how does that work?

So consider a data set in which you're

enrolling three subtypes of infection and you get three separate point estimates if you simply looked at the treatment efficacy that appears to be present in each one of those three sites. Each one of those point estimates has some uncertainty around it. And the hierarchical model assumes that those treatment effects are themselves drawn from a population of treatment effects among lots of different body sites that might exist. And the hierarchy is an assumed distribution of body sites or subgroups, and there's uncertainty in how similar those subgroups will be to each other. That's the prior variability at the third level of the hierarchy. So you actually allow the model to realize you don't know how similar the treatment effect will be in the three body sites or the four body sites.

When the hierarchical model fits the data -- and this is sometimes called a shrinkage estimate -- it creates new estimates for the treatment effect in each of the body sites individually that learns from the consistency of the effect across body sites.

There is a very general statistical fact called the James Stein effect -- I put the 1961 publication there just to show it's not new -- that says that the best estimate of a true treatment effect in a subgroup is actually not the treatment effect you get by simply looking at the data within the subgroup if there are three or more subgroups. This is a very general, non-Bayesian result that is horribly, horribly inconvenient because it is so counterintuitive. It means that when you look at the treatment effect in a type of disease for a drug, you should think about how well that drug works in all kinds of other similar diseases, just like clinicians do every day.

So how about the non-inferiority and superiority issue? The data that one would get from this proposed approach to a platform trial would include data when the infecting organism is isolated and found to be highly resistant, when it's found not to be highly resistant, and also you're going to have some patients who clinically appear infected for which you are unable to identify a specific etiologic

isolate.

The first data set from the patients who have isolates that are highly resistant can be used primarily to address a superiority hypothesis. It can secondarily be used to address a non-inferiority hypothesis if there is value in approving a drug simply because it's non-inferior to the comparator in that setting.

When the infecting organism is isolated but found not to be highly resistant, that can be used to address a non-inferiority hypothesis and also to address safety and PK and other goals.

When the patient has no isolate obtained, that can be used to address the non-inferiority hypothesis, but there are some very specific considerations there depending on the site and what a non -- a positive culture means in that setting. That would be a great thing to discuss at length another day.

But the point here is that the membership of the patient into each of these subgroups is based on a pre-treatment assessment, which is the

culture. So these are valid subgroups for which you can draw valid statistical inferences. This approach allows the early initiation of the investigational product before you know the subgroup into which the patient will fall, and more of the patients that you screen contribute data that informs clinically-important questions.

Agents should only be included in this sort of approach when there is strong learn-phase rationale, demonstrated penetration, PK, lack of inactivation at the site -- that's the PTSD from the lung example -- that makes it reasonable to test that drug in that site against the likely spectrum of organisms. But with this approach, each enrolling site -- meaning clinical study site, hospital, or whatever -- can contribute a larger number of patients per month because there are multiple infection types and resistance patters included. And that decreases the per-patient cost and helps you support your network. The efficiency of the platform is increased when there is more than one investigational agent available at the same time for an indication because

of the shared control arm, but that is not necessary to achieve some savings. And the current environment with multiple smaller companies may be particularly conducive to a platform strategy.

I have a couple of slides which I've included in the site that have to do with labeling and stewardship. The point simply stated is that when we approve drugs that are most valuable from a public health point of view when their use is restricted to highly-resistant organisms but we label them for broad clinical indications, that that potentially undermines stewardship efforts. I don't want to dwell on this, because first of all, I'm not an expert in this area. But I do think that the statistical design drives the data, the data drives the labeling, the labeling drives the marketing, and we want to make sure that to the extent possible that that supports stewardship from a public health perspective.

So in conclusion, I think the most common structure in statistical design of confirmatory trials of antimicrobials risk failing to answer the questions that are of most direct clinical urgency and

impact. They fail to address the likely subsequent use of approved products across multiple infection sites, demonstrated only on the presence or risk of highly-resistant pathogens and hopefully PK and penetration data, but that a multi-infection site, multi-drug platform trial addressing both non-inferiority and superiority simultaneously could address those challenges. Thank you very much.

DAN RUBIN: Thank you. Our next speaker is Scott Evans. He is a Professor and Founding Chair of the Department of Biostatistics and Bioinformatics at George Washington University and the Director of the George Washington Biostatistics Center. He is also the Director of the Statistical and Data Management Center for the Antibacterial Resistance Leadership Group.

SCOTT EVANS: Thank you very much, Dan. And good morning, everyone. Thank you for the opportunity to talk with you today. I'm going to move quickly.

A couple of years ago, I had a leaky roof in my house. It created a water bubble in my

wall. It was a very strange-looking thing, but I had a water bubble in my wall. And in addition to a new roof, I had to repaper the wall. And my neighbor had recently papered a similar-size room in his house. And so I asked him how much paper did you buy. And he replied six rolls.

Upon finishing papering the wall, I had only used four rolls. And went to my neighbor and I said, listen, I had two rolls left. What happened? And he replied, oh, that happened to you, too?

Now, I tell you this story because I asked the wrong question. And I find in clinical trials we're often asking the wrong question as well. And as a matter of fact, the two things I've learned about antibiotic clinical trials since I've been involved with them. First of all, they're rigorously conducted by experts in the field, they closely adhere to highest standards and fundamental principles of randomized trials. And secondly, they're essentially useless for helping clinicians make treatment decisions. And so we've been working on ways to try to figure out how to rectify this.

This was said perhaps more eloquently by the former FDA Commissioner Rob Califf. "Most clinical trials fail to provide evidence needed to inform medical decision-making. However, the implications of this deficit are largely absent from discourse."

So the example uses we have in antibiotic trials is drugs are compared in susceptible disease, but susceptibility is not known until we actually start treating, after we've started treating the patient. Patients are considered failure when they change therapy, though they may not actually fail. We lose interest in patients that change therapy, despite therapeutic adjustments that could effectively treat the patient. Populations studied are not the same as the population applied. In non-inferiority trials much of the time we exclude patients with recent prior therapy. These drugs are then used in these patients, possibly representing a majority of these patients.

There are further issues. We often define analysis populations in trials. Efficacy

analysis we define an intention to treat population. Safety analysis we define as safety population. Those populations are not the same.

We then combine these two analyses into what we call a benefit-risk analysis. But to whom does this benefit-risk analysis apply? We're estimating a parameter from a population that doesn't exist. Nobody seems to mind.

Another question. We measure duration of hospitalization, duration of the ICU stay in clinical trials. Shorter duration is better. The faster the patient dies, the shorter the duration. So you give me a summary statistic of duration of hospitalization, I don't even know what it means. Part of that's self-inflicted wounds by the way we design and analyze studies. Outcome interpretation needs context of other outcomes for that same patient. Once you tell me whether the patient lived or died, the number makes sense.

Question three. Trials typically use binary endpoints. For example, cure. The patient has to survive, their symptoms resolve, they have

microbiological eradication, no changes to therapy. But consider the following. One patient fails because they die. Another patient fails because they have a lack of micro-eradication. It would seem reasonable that now our primary analysis doesn't distinguish these two patients. Shouldn't a primary analysis recognize the difference of this? That would seem important enough to recognize.

This would seem to be particularly important recent FDA Advisory Committee for evaluating plazomicin in complicated UTI. The composite cure rate, 81 percent for plazomicin, 70 percent for meropenem. But if you just look at the clinical cure rate, very close. Right around 89, 90 percent for both. It turns out that the advantage for plazomicin is in the micro-eradication. That would be particularly important to know particularly when you start to look at the safety data that suggests, well, there's more safety concerns with plazomicin than perhaps with meropenem.

The last lesson I'd like to sort of motivate this is one plus two times three is not nine.

And grade school children know this, but the clinical trial community have missed this course. Let me explain what I mean by this. So here's a question for you. Supposed a loved one is diagnosed with a terrible infectious disease. You get to elect treatment. We have three treatment options; A, B, and C. All right?

Now, let's suppose there's two outcomes for simplicity. Let's assume they're equally important. You have treatment efficacy, treatment success outcome. Yes or no. The patient gets it or they don't. They also have a safety event. Patient experiences it, yes or no. Now, luckily enough, we had a randomized trial that compares A, B, and C that help guide our decision about which one we should choose. Had a hundred patients in each arm, A, B, and C. The treatment success rate is 50 percent in A, 50 percent in B, and 50 percent in C. Now, the safety event rate is 30 percent in A, 50 percent in B and C. Which treatment do you want?

Well, they all have the same success rate. A has got the lowest safety event rate. B and

C are indistinguishable. Can't tell the difference between them. Clearly, we choose A. We're all reasonable people.

But instead, what we've done here if you evaluate what we've done here, we've taken the patients in the trial and analyzed the outcomes, the two outcomes. What I'd like to do is flip that upside down. Take the outcomes in the trial and tell me what happened to the patient; that the purpose of measuring outcomes in clinical trials is to tell you how the patients are doing. We seem to have gotten it backwards.

So if I take the outcomes and analyze what happens to the patients, there are four possible outcomes for what happens to patients. They get the treatment efficacy, yes or no, and they get the safety problem, yes or no, in combination. So let's look at the combination.

Well, it turns out in treatment A, the efficacy and the safety were uncorrelated. So there were 35 patients that experienced the treatment success and avoided the safety problem. But in

treatment B, they were positively correlated. So there's zero patients that had treatment efficacy without the safety problem. In treatment C, they're negatively correlated. So I've got 50 patients that experienced the efficacy without the safety outcome.

Now, one slide ago I couldn't tell you the difference between B and C. Now, we're supposed to be in an area of personalized medicine where we're doing personalized -- and I can't tell the difference between B and C. So our culture has been to use the patients to analyze the outcomes. Shouldn't we use the outcomes to analyze the patients? That's the purpose.

So as my father told me many years ago, the order of operations is important, and we haven't got the order right. And we may be missing things without realizing it.

So William Osler, the well-known clinician, said years ago, "The good physician treats the disease. The great physician treats the patient." And maybe we should be analyzing things that way.

So Dean Follmann at NIAID and I wrote a

paper describing some of these issues and how we might attack things from a different angle in doing so. A couple of years ago we wrote this paper, DOOR, Desirability of Outcome Ranking. Think about what those words mean; desirability of the outcome. And it's going to be a ranking. And what we do is we end up computing what we call the DOOR probability, which is a probability of a more-desirable outcome when assigned to one therapy relative to another therapy.

Now, if you're a clinician treating a patient or you're a patient and asking I've got to treatment options, you can ask for differences in means and difference in proportions and hazard ratios and relative risk all you want. But what is a more natural question to ask when I'm trying to figure out whether I should take one therapy over another? Well, how about the global probability that one therapy is better than another? Wouldn't this be intuitive? So although it may be foreign at the moment, that perhaps this may be intuitive in the long run.

So here's an example of an application of this. Should we use ceftazidime, avibactam, or

colistin for the initial treatment of infections due to CRE? And this was a paper published by David van Duin using some observational data from an ARLG study.

Now, the DOOR that was set up, a desirability of outcome ranking, an ordinal outcome. You see most desirable at the top. The patient lives, they're discharged home, everything's back to normal, they avoided major adverse events. Least desirable is at the bottom; the patient dies. But there's layers in between where some things go right, but not everything goes right. And the idea is if colistin, whatever colistin produces, patients fall into these categories wherever they fall. Can I get patients to migrate northward using CAZ-AVI to more desirable places and evaluate that. And that's exactly what happened in this particular evaluation. The DOOR probability is 64 percent that you're going to have a better outcome using CAZ-AVI than you are in colistin.

Now, how do you summarize this sort of patient journey, which is the next question if you're going to use DOOR? Now, before analyzing several hundred patients, you have to figure out how you're

going to analyze one. Now, clinicians are supposed to be doing this when they're treating patients, but we've got to get into the head a little bit more.

So there was an example strategy about how we might create this ordinal outcome. This was led by Sarah Doernberg at UCSF through another ARLG study. And we call this BAC DOOR, because this was for bacteremia. And what we did was we were envisioning we're going to do a staph aureus bacteremia trial. And we did a pretrial sub-study to try to figure out how we might create a DOOR in this particular area.

So what we did is we took 20 representative patient profiles from a prior study that had benefits and harms and quality of life, and in a single paragraph we wrote down what happened to those 20 patients.

We then sent those profiles to 43 expert clinicians that treat this disease and said rank them in terms of the desirability of their outcome. We didn't tell them how to rank them; the idea was to figure out what are they valuing in terms

of how they rank patients.

And then we examined the components that drive the clinician rankings and used that to try to create a DOOR outcome. And we came up in this particular case with six different levels.

Now, one thing we learned from this particular exercise. First of all, you for the first time begin to evaluate the cumulative nature of things that happen on patients. Now, fi you examine one outcome at a time, you never see anything cumulative. We never look at it. But if a patient is having multiple bad things happening to them, then that should be recognized. That's how patients are experiencing these outcomes, and that we should be thinking about how to do that.

So some natural questions arise when you evaluate things like this. There are potentially unequal steps between these categories. We like to be able to recognize that. Perhaps the step to the bottom category, to mortality, is bigger than the steps above it, is larger than the steps above it. We can recognize that. There's also varying perspectives

among patients and clinicians regarding the desirability of these different levels, and could we recognize that. And we have done this sort of analysis. We proposed what's called a partial credit analysis, which if you've got these ordinal outcomes -- these are the four ordinal outcomes from the CAZ-AVI colistin example -- what you do is you say, well, if you have the most desirable outcome, you get a perfect score; you get a hundred. If you die, you get a zero. If you get in the middle, then you get partial credit. And there are some very natural ideas about how we -- next natural question is what are you going to give me for partial credit. And we've been working on that particular problem and we've got some ideas for that.

The other thing that can happen with this is there's been this debate in infectious disease, this area of infectious disease for quite some time that says, well, who do I want to enroll in trials. And one theory says, well, don't enroll the very sick patients, because they're going to die anyway and you're not going to have any sensitivity to detect any sort of effects.

On the other hand, people say, well, don't enroll the very more healthy patients, because they're going to recover anyway, and you're not going to detect any effects, there's not going to be any sensitivity. Well, in this particular case if you look at things using a DOOR type outcome, you can evaluate which patients are actually benefiting from say CAZ-AVI over colistin. And in this particular case, the most severe patients were the ones who were benefiting over CAZ-AVI -- CAZ-AVI over colistin.

Let me show you this idea in a little bit in another example. This is the PROVIDE study, another ARLG study, which was a prospective, multi-center, observational study evaluating among hospitalized patients with MRSA bloodstream infections.

The research question; what's the vancomycin PK exposure target that's associated with an optimal treatment outcome. All right? 265 patients in this study.

Now, what we did is we set up a DOOR outcome. Most desirable at the top says treatment

success, and you avoid major toxicity, acute kidney injury. Least desirable at the bottom is the patient dies. But there are gradations of patient response along the way.

So what we did is looked at DOOR outcomes by dosing quintiles. So the top dose there is the highest bar. And what you see here is a distribution of the DOOR outcomes by dosing quintiles. Highest dose at the top, lowest dose on the bottom. All right?

So the blue on the left-hand side is that most-desirable category, treatment success without acute kidney injury. The purple on the far right is mortality. So what are you getting as you increase dosing? What you get is toxicity. You're not getting more efficacy; you're getting more toxicity. So higher doses bring higher toxicity, but not greater treatment success. All right. So actually this particular paper came out today in the Annals of Internal Medicine using ideas like this to actually monitor patients during the course of say DMC monitoring.

So in ARLG 2.0 one thing we're hoping to do is to develop standardized, syndrome-specific DOORs for the major infections in this particular area.

My last point I would like to make. There's been some talk today about platform trials and so forth. And in the ARLG we actually evaluated a platform to try to get at more pragmatic answers to the questions that are most important for us. We called this SMART-COMPASS. And the idea is -- so SMART-COMPASS stands for Sequential, Multiple-Assignment, Randomized Trials for Comparing Personalized Antibiotic Strategies. So this is consistent with the strategy theme that Vance and others had talked about earlier. But it's actually quite flexible and very consistent with the way in which patients are treated. It addresses several types of research questions, including identifying optimal treatment strategies, it can evaluate empiric therapies, and it can evaluate definitive therapies, those that are likely to be licensure-type questions, and provide efficiencies compared to traditional

multi-arm trials. And it's very pragmatic in the sense that it really mirrors clinical decision-making. You can think of it as personalized medicine.

So I'm going to end with a quote from NBA coach Frank Layden, who had a player that was not producing. And Layden asked the player, so what is it with you, son? Is it ignorance or is it apathy? And the player looked at Layden and said, Coach, I don't know and I don't care.

So I say to you today that if people don't know, let's educate them. And if they don't care, then let's motivate them. And I want to thank my collaborators, Dean Follmann at NIAID, Dan, and Chip and Vance from the ARLG, and all the ARLG team. I have no doubt that you will enthusiastically applaud now because you're so relieved that it's all over. Thank you.

JANE KNISLEY: Okay. Our next speaker is Aaron Dane. He is a consultant statistician who has been investigating how to make the design and interpretation of antibiotic trials more feasible. He has been a consultant since April 2016 and has

continued to make such development more feasible.

AARON DANE: Thank you. So as has just been said, all I've been looking at is for a while now is seeing more and more that we talk about doing rare pathogen trials. So this is a bit different from some of the other presenters in that this is in that area where it's very hard to get very much information. But we seem to be stuck in that if we can get 50 to 100 patients with a resistant pathogen in a pretty long timeframe, we've got nothing we can do with the data. Because we can't get to traditional statistical criteria -- so we run the trial and then we can't do anything with it.

So as a result of that, I've been doing some work with Professor Nigel Stallard at Warwick University where we've been looking at -- and he's taking ideas from the orphan drug area where how can we use that information a bit more readily and get something from the data, particularly in areas where very specifically we can't generate much more information.

So we have a technical problem.

Well, maybe while they're sorting it out what I'll do is just -- so what I'm doing here -- the idea is can we get more from a small study when we haven't got the option of a bigger study? And what I'm looking at here is very specifically we're still talking about doing a randomized trial. So this is still randomized, it's still controlled. This isn't about external controls, but it's having done that because we still feel that's the best way to get some useful information; what can we do with the information we have got? Because as we talked about yesterday, at the moment we do the study and then no one can use the information. So it's just a way of doing something a bit differently.

Okay. Is it working? All right, okay. Try number two.

So these are my disclosures. And so this was a collaboration with Nigel Stallard in Warwick. He has done some work in the orphan drug area. So basically that was the idea here, was to try and take some of that thinking when you've got that limited population and see if we could apply some of

that. And Paul Newell and John Rex have also helped as we've worked that through to put some clinical perspective on that as well.

So I'll skip through this quickly. So in terms of the superiority studies, this has been covered a number of times. What I would say is the idea is if we could do superiority studies, we would. It makes things much easier and much clearer. So it's not that we'd rather do non-inferiority studies, it's just that they often not really stick. For the reasons that have been spoken about, we can't study ineffective comparators. The numbers are small just by definition, which makes it more challenging. And there's often at least one therapy with some degree of efficacy. It may be toxic, there may be various other reasons you don't want to use it. But it's not that we're often dealing with something that's completely ineffective.

And also from the societal point of view, we don't want to just be developing new antibiotics when we can show superiority; we want to do it before that time.

So this was really just to illustrate how quickly the numbers fall away when we're talking about resistant pathogen studies. This is just an example, it's made up. But it's this illustration that from 300 patients if you're trying to then focus on a very specific pathogen, your numbers fall away very quickly. And that's before you even talk about that pathogen being resistant to all other therapies and all the other comorbidities that may cause confusion. So this is why we still need to find ways of developing trials for non-inferiority setting as well as superiority, even when we're talking about resistant pathogens.

And I think the key thing is that a lot of the time we're developing this for tomorrow's patients. So it's for a future unmet need as much as a current unmet need. And that's why we have to do it this way.

And I think then what we need is find a way of saying, well, how can we develop something when we can only get a hundred patients? So what do we do with that information rather than saying, well, that

doesn't meet traditional statistical criteria, so we just can't do anything. So it's finding a way of getting some information -- the conversation yesterday was, well, these drugs are still going to be used, so it must be better that we provide some framework for whether that drug's effective even if it's not the way we would do that traditionally.

Something I'm not going to mention in this talk is the idea of the safety database. It's obviously an important element as well. And that would have to be considered. But I'm focusing on the efficacy aspect for this talk, but that would clearly be an important aspect as well.

So a possible approach to design of rare pathogens. So some of this is standard for any trial; what are we interested in when we run a trial? So we want to be confident that we can show an effective treatment works, and we want to be confident that we won't approve ineffective treatments. That's the premise of any trials, and that's a Type I error in power.

But the question is can we look at that

differently for rare pathogens and can we do that for example using some of the questions in the orphan drug area about understanding what extra information and value we gain from making the study bigger.

I think one of the key factors for me here is -- because often this discussion is saying, okay, it would be acceptable to do a small study, and that's all right. But actually we still need a framework for the decision-making. So as a sponsor, you need to understand what you need to show so you can understand the risks. These studies may be small, but they take a long time and they're very expensive. So we still need to understand how likely we are to be successful. And also just having clarity at the outset on what the decision-making is going to be means that it's much clearer to everybody what's needed.

So the aim here was to provide a framework in that setting where you've got other therapies that may be suboptimal, but they still have some efficacy. And what I'm going to go through, I'm going to step through some examples. And this is not

about performing an interim analysis or not about continuing on after the trial in a single-arm setting. What this is about is saying if you've got a limited population, you run a trial in part of that population, and then everybody else goes on to one of the two treatments. Either they go on to the test if it had a successful outcome, or they continue on the standard of care if your new treatment failed in the study.

The other thing, I've stuck to frequentist statistics here purely from the point of view that I'm talking about a different idea anyway, so I didn't want to get into the Bayesian side as well. But you could equally apply this with Bayesian methods, some of the things I'm looking at.

So the first question is large versus small trials with rare pathogens. So larger trials lead to higher power, so that's great. Right? You'd always want that. But what we have here is that if the trial is too large or it takes too long, what it deprives patients of a more effective therapy, and it also means it might not be feasible to develop the

drug in the first place.

But on the flip side, if the trial was too small, it may be more feasible, but then you end up with much higher chances that you're going to make the wrong decision, which clearly we don't want to do, either.

But the theme for all of this is how do we work with a small data set? Because that's all that's going to be possible. We're not going to be able to generate any more. And that part isn't really much of a choice. So can we show there's a sweet spot for the sample size where we get sufficient information and we're not going to gain a lot more if we keep going on with the trial? And really that's the idea of the diminishing returns outside of that sweet spot.

So what we're aiming for here, so if test is worse than control, then in that situation, every patient randomized to test in a trial risks a worst outcome. So if test is approved, we perpetuate that problem, and there's even more patients that are having a worse outcome than they would otherwise. So

the mitigation here is that we have a small trial, but we avoid or reduce that chance of incorrect approval, which is what we normally talk about as Type I error.

Equally, if a test is better than control, then every patient randomized to control risks a worst outcome. So in this case if test is not approved, then that's perpetuated. So you're still risking everybody having a worse outcome. So here within that small trial we want to keep the power high.

And then the third scenario I'll present is when test is similar to control. So in this case, we still want to make additional therapies available. So again, in this case within that small trial, we'd want to keep the power high.

The important part with all of this though, this all sounds fine and easy. But the fact is we don't know which of these three scenarios is true when we're designing a trial and when we're interpreting a trial. So this is where we need to understand the risks we have of all these things happening. So what are the risks that we incorrectly

approve a new drug, or what are the risks we miss a good drug by the criteria we apply?

So what this is presenting -- so I'm going to present three different scenarios, one for test being better, one for it being worse, and another one where it's the same to understand some of those risks.

So in this situation imagine that we've got a non-inferiority trial which is looking at test against control. The test response rate is 60 percent and the control is 40 percent. So what we've chosen is a non-inferiority margin of 20 percent here with a 95 percent confidence interval. So we've already started with a wide margin because of the rare pathogen and the area of unmet need. So the correct outcome and what we want to say in this case is that we conclude non-inferiority.

So we run the trial. And the plot on the right shows the chances of concluding non-inferiority. And what it's showing is that from about 40 or 50 patients per group, you've got a reasonable chance that you're going to show non-inferiority. So

80, 90 percent power. So you might need more for the safety database, but in this scenario what that's showing you is that you'd have a decent chance of success given that situation.

And then on the flip side, if the test was 20 percent worse than control, so if the response rates here were 20 percent and 40 percent, what would happen then? So we're still using the same NI margin and confidence interval, but in this case you don't want to conclude non-inferiority. And what that's showing is that the chance of concluding non-inferiority on the plot now is very low. So in this case, again, 40 or 50 per group would be okay for this situation. Because what it's showing is you rarely make the error of concluding non-inferiority.

And finally, when the test and control are the same, we have all the same setup. So this time they've both got a 40 percent response rate. Now, as you would expect, as the sample size gets bigger, the chances of concluding non-inferiority goes up. But what this is showing is that it would take a bigger study in this case to do that. So at 50 per

group, you'd only have a 50/50 chance of success even when you're actually truly the same as the comparator.

So around 80 per group may give you what you need. And it takes a long time to get beyond that. But I think the key question here is you'd need a substantial bigger study, and otherwise, you'd end up with a 50/50 chance for success.

And the reason I raise this is in some situations, that might be fine. It might be quite easy to get to 200 patients in one of these studies. But if it's not and you could only realistically get to about a hundred, a sponsor isn't going to put the time and money into a study if they've only got a 50/50 chance of it succeeding.

So what I've gone through so far is sort of the more standard aspects of Type I error in power. But the other part that's been pulled through from the orphan drug area is what does this mean for patients after the trial. So in this case, suppose we had an overall population of patients with a rare pathogen. There's a thousand of them in total. And then if we include a hundred of patients in the

clinical trial, so 50 per arm, that means we treat the remaining 900 patients with whatever drug comes out of that study. So that means we have 50 patients on test, 50 on control, and then another 900 on one or other of the drugs depending on whether the trial concluded that the test was non-inferior or not.

And the purpose of showing this is that a bigger trial isn't always better with rare diseases when you're doing this. So if we assume we had a thousand patients with the rare pathogen and we had a hundred patients in the randomized trial -- and from what we looked at before, if we assume that we have a 60 percent response rate for the test and 40 on control, as I showed before, what that means is that you would expect 30 of your 50 responses in the trial on test, and 20 out of 50 on control to respond. So that's the 40 and the 60 percent response rates.

And then dependent on whether that trial concluded non-inferiority or not -- so if the test was successful and got carried through, from that point, those remaining 900 patients you would expect to have 60 percent of them to respond. Okay? Whereas

if test failed, so you didn't conclude non-inferiority, that would mean that the standard of care would continue to be the treatment used. And so only 40 percent of those 900 patients would now have a response.

And I won't get into all the details of the power and the probabilities of those two. But what you can then work out is the expected number that you expect to respond, which in this case would be 587 of the thousand. So that's a 58.7 percent response. So that's fine.

But now then you'd say, okay, why don't we just do a bigger study? Because then we'd get more certainty that we have the right decision. But for this situation and this scenario, if you did the same thing with a trial of 400 patients, what happens now is you've still got the same 60 percent response on test and 40 on control, only with 200 patients per arm in the study now. And what that means is overall there's only 560 patients, or 56 percent of all the patients with this pathogen respond. And the reason for that is because you're waiting longer and you're

giving more patients the ineffective therapy before you switch to one therapy or another. So this is where the consequences can be worse for patients afterwards if this is the case.

So in terms of finding the sweet spot that I mentioned before, what we need to do is find a sample size we have a good chance of success when we're effective, low chance of approval when we're ineffective, and a reasonable chance of success when the two treatments are similar. But also consider this expected number of patients who benefit so that we maximize that. So the following plot summarizes that information.

So this is plotting out that expected chance of success. So here the top half is what I presented earlier, which is to say how often do we select the test agent. So when test is 20 percent better, we pick that pretty well from 40 per arm. And what the bottom plot shows is that a trial with 40 patients, you ought to optimize that number of successes when you get to about 40 or 50 patients. And after that, it's diminishing returns because

you're keeping more patients in the study even when one of the treatments is less effective. So this is where you can then use this to look at that and say, okay, well, maybe 40 or 50 patients is reasonable.

Similarly, when test is 20 percent worse, what happens here is -- we already mentioned that the test is not selected very often. And again, the expected number of patients who respond drops right from the start in this case. And that's because in this case you run the study longer, and that means you're continuing to treat more patients on the test agent even when it's not effective.

So finally, the largest sample size for test is similar, as I mentioned before. So that goes up slowly. And you'd need to get to about 80 or a hundred patients pre arm before you'd get to a reasonable chance of success. In this case, the number you'd expect to respond stays the same. And that's because both treatments have got the same level of efficacy. So whichever one you choose, your expected response would still be 40 percent in this case.

So the reason I've gone through all of that is because what this shows is that 40 to 80 patients per arm has reasonable power in these settings that I've shown. And that's assuming that the control has 40 percent efficacy and then the test has either 20 percent better or worse response.

But the fact is if the product has similar efficacy to the control at the moment, that needs more patients, a bigger study, which in some settings might be possible. But from what we've been talking about, that's not always going to be possible. So getting to 200 patients with rare pathogens is not always the case.

So if that's the case, how could we provide criteria when it's only feasible to recruit 50 patients per arm, for example? And that could well be taking five years plus to get to that.

So one approach then would be to look at the plots that I've just mentioned, but use something like 80 percent confidence intervals rather than the traditional 95 percent confidence intervals.

Now, what I would specify here is that

this is in areas of large unmet need, and it would be very specific. So this wouldn't be a blanket approach to this. They would have to be specifically agreed that this is the idea that we understand the risk, but we do it in situations where the alternative is we have no data.

So I won't go through these plots in detail again. I'll save you all that pain. But what I'll summarize is that what shifts here is the observations that in this case is that when test is 20 percent worse, the risk of approval goes up, the incorrect approval. So now it's about a ten percent chance that you'd conclude non-inferiority for a less-effective treatment. But when the test is similar or better, using the 80 percent confidence interval gives you a higher chance of success than using the 95 percent confidence interval. And the pattern for the expected number of responses is similar to before.

So the important thing here is that this is example framework. So I'll give an example of using test and control being 20 percent better, 20 percent worse. And I've given examples of 80 percent

confidence intervals. But the idea is this is a framework to say at the moment what we do is we say, well, we run a resistant pathogen study, and if it doesn't meet standard criteria, then we can't use it, and we know we can't get to those criteria. So this is a way of discussing what the unmet need is, what the feasibility is, and trying to get some idea of where we're going to be with some of these aspects and change some of the success criteria that we might have for these very specific situations.

So in summary, this is a framework to display the tradeoffs. So this is a situation when only a small trial is possible. So we need to understand the false positive and negative error rates, but we should think about whether we can change what those error rates are compared to a traditional area when we can study hundreds of patients. And the data on 100 to 200 patients can be very informative. And it feels like it's still better to generate that data and understand what it's telling us even if we have to acknowledge that the levels of risk and the level of information is not the same as we are

traditionally used to.

So in conclusion, I would say that the use of the power Type I error and the overall number of patients benefitting could be used to agree the sorts of criteria for these treatments of rare pathogens so that we get some information where everybody understands what we're going to see as a successful trial at the outset rather than us having to do the trials and then not really knowing what success is going to be until we get to the end of the study. Okay, thank you.

DAN RUBIN: Thanks, Aaron. We now have scheduled 15 minutes of Q&A and we'll go until 11:20, and I'd like to open up the Q&A for speakers for both sessions from this morning. If you'd like to make a comment, please just turn your card up, as was done yesterday. Try to get my attention or Jan's attention and we'll try to keep track of the order. Please, go ahead.

MANOS PERROS: Thank you. Just a comment and then a question. I find it shocking but not surprising that I believe, Vance, 87 percent of

the prescriptions in your institution for those new drugs were off-label. It's shocking. And I think we all need to decide whether we can settle for something like that or whether we want the labels to be more reflective of how the drugs are actually used. I would be in the latter camp but that's a discussion to have.

On the more specifics, I like the idea of registration trials, strategic trials. Let's remember that we also have other groups than prescribers and regulators that need to be satisfied in order for us to be in this business, and that includes pharmacists, P&T committees, payers, and who is going to -- what kind of data and what kind of information is (indiscernible) those groups.

And the bottom line of that is more trials take time and money that we don't have. Maybe Merck could do that. I don't know why they would but they could run. But we only launch a drug once. The price is set at launch. The trajectory in the first quarters is important. And while I love the idea, I think we need to think how we can do those kind of

trials before the drugs I launched.

VANCE FOWLER: Well, yeah, we can think about that all the -- you know, all the livelong day. I think that -- I mean, I think you... You know, what's that saying about don't let the perfect be the enemy of the good. I think we simply -- what I'm thinking about from a clinical standpoint is we need these drugs. You know, we need them today. And that 87 percent, that's -- that is essentially -- if anything, it's maybe a little low. I mean, that's been... Mike Rybak published -- you know, you talk about ceftaroline a little bit more, for an example.

So, we've been -- you know, that compound, incredible compound in my opinion, you know, it's been handed off -- it's been owned by no less than five different entities so far. So, Cerexa, then it was Faris, then it was something that started with an A, then it was Allergan. And now Allergan's going away. I mean, so, you know -- and the likelihood of that -- of meaningful trials taking place with each of those handoffs plummets.

You know, Mike Rybak published a thing

from 2014, that's the best data we have to deal with it. It was a retrospective series of, I don't know, about 400 patients. Over 80 percent of those patients he described, and that was in 2015, were off-label.

So, that's the status -- I mean, that's the state of affairs in clinical care, treating drug-resistant pathogens in the United States. That's just the way it is. And so I don't see that, you know... And if we can go in and do trials beforehand, totally agree. I don't care when the trials get done, honestly, from a clinical standpoint, as long as the clinicians have the -- you know, can confuse the issue with facts about what we're actually doing, it's a win for the patients, right?

So, I'm just -- you know, what I'm trying to really propose is I guess maybe I'm saying I don't think it is going to happen, you know? Maybe I'm a pessimist. I usually am. I think that what can happen is we can get, you know -- that companies can get compounds through trials and complicated UTI and ABSSI or what have you, because it minimizes risk.

And then what I would like to see

happen in the next ten years, I do think clinical trial networks are the way forward. I feel what I would like to see is a means by which there -- that funding can be provided not just even from a single country but on a multinational perspective whereby everyone puts in a little bit, there's a means by which decisions -- responsible implementation of that precious resource can be applied to do the trials that have to be done.

So, yeah, maybe that's a little -- you know, big but, you know, someone's got to think big, you know, or we're just going to stay right where we are. It's the clinical -- it's the clinical community that's going to have to take on these tough-to-treat trials, you know, tough-to-complete trials because it just doesn't... You know, you guys have already got so much stuff just trying to stay alive that adding a -- putting your entire future on a single trial that is too -- that's never been done before, to me, is a bridge too far.

MANOS PERROS: Yeah. And, Vance, just to be clear, my point is not that -- I don't disagree

with anything you said. I'm not arguing that we should change the way those drugs are used. This is the right way to use them. But the point I'm trying to make is getting the drug approved is no longer enough for us to survive as an industry.

DAN RUBIN: So, we have a question online from John Tamico -- from John Tamico to Roger Lewis. We'll do that and then I've got Dr. Rex. So, the question, Roger, is "CIAI is partly a surgical disease, even resistant organisms can be managed with source control to the point host defense can clear infection confounding antibiotic effect. How is CIAI reliably informative across body site analysis?"

ROGER LEWIS: Okay, well, as a non-expert in this area, I think I'm going to take this as the general question of if there is reason to believe that a treatment is likely to be less effective in one setting than another, then that violates an assumption of exchangeability of the body site. So, one of the assumptions of a hierarchical model is that a priori, you're not sure which site the treatment effect is likely to be largest if you couldn't order them a

priori.

So, if you have reason to believe that an antibiotic is likely to be relatively less effective because of other things that affect outcome, then you have to adjust for that in some way. And perhaps pooling it doesn't make sense.

That said, what I would want to do is look at trials of treatments of antibiotics for complex intra-abdominal infection in which the isolates were sensitive and look at the magnitudes of the treatment effects based on the endpoints that were used, which presumably were compromised by the same effects, and see if it's really true that the treatment effect that we have observed form previously demonstrated efficacious antibiotics was really smaller to make that assessment.

DAN RUBIN: Thank you. Dr. Rex?

JOHN REX: Those were great talks. And back to you, Roger, this Stein-Lewis concept is almost causing my head to explode. And if I say it back to you it's that if I have a point estimate of some measurement, and then I have three point estimates of

three measurements, I'm actually -- the most valuable measure of my first point estimate isn't the actual number; it's actually a combination of the three point estimates. That seems to be the concept here that I can -- and the error in one direction might compensate for error in another direction, if I'm catching the drift of this.

And it seems to me that has some very interesting translations for us when we use -- you know, we focus on the one-off data -- that's my point estimate in CUTI. That's my... And it almost seems to suggest we're better off doing studies where we actually deliberately grab several different body sites as part of it and get -- and combine them. Or am I misunderstanding this? Talk -- talk a little bit more about James, Stein and their very counterintuitive concept.

ROGER LEWIS: So, the first has to do with understanding the difference between bias and estimation in treatment effect. So, the James-Stein theorem says that if what you're interested in is the error in estimating the treatment effect -- and I mean

error like mean squared error -- you get a lower mean squared error if you use information from all of the related things.

The -- it is not an unbiased estimate, okay, but the bias -- what you lose in the unbiasedness, in my view, in terms of the clinical utility estimate, is far outweighed by the benefit you get in the reduced noise that Dan mentioned.

So, the -- and you can think about it in the following way: Let's suppose you do a clinical trial and -- for a treatment in osteomyelitis, and there is a forest plot presented which shows the estimated treatment effect across all the different bones. You can't get to 206 but some large number of bones.

There is going to be some bone where the estimated treatment effect is surprisingly high or surprisingly low compared to all the other bones. And you as a clinician will look at that and know that that bone might be a little harder or easier to treat but it's probably not as much harder or easier to treat as the data suggests because there is the -- the

estimates that are particularly high and low are likely to be a combination of the true treatment effect being high or low and random fluctuation.

And what the James-Stein estimate or a shrinkage estimator does is it accounts for the expected variability in multiple measurements to give you an estimate that appropriately balances the evidence that one bone or subgroup is different than the others against the additional likely variability. That's what the shrinkage estimate does. When it does that, the distance between the estimated treatment effect and the truth, on average, is smaller.

So, what this means -- and I know there's a lot of people who do editorial work here -- is that if you look at the forest plot or the table of subgroup effects in a clinical trial, that is fundamentally the wrong estimate if your goal is to give you the most accurate treatment effect estimate in each subgroup. It's fundamentally wrong.

LINDSEY BADEN: Yeah, but that's how we do it. I mean -- that's so interesting. Wow.

ROGER LEWIS: I wasn't debating what's

in print.

MAN 1: But that has its own assumptions.

LINDSEY BADEN: But people spin around --

ROGER LEWIS: So, the interesting thing about it is that the assumptions are extremely loose. And so the reason that this is not adopted as a standard approach I think has to do with a limitation in humans to understand situations where statistics work in a way that's counterintuitive. Just as if most people's experience doesn't include the fact that time slows down when things go really fast, okay? It just turns out to be true.

DAN RUBIN: I would just add on the James-Stein theory -- that the theory is that your mean squared error must improve if there's three or more subgroups. If you're talking about your average error or the sum of the errors across the groups, if there's any, you know, one outlying group, there's no guarantee that that, you know, within subgroup estimation must improve by bringing in other data --

although, you're right, that on average that shrinkage is going to be really helping you.

So, next we have Dr. Melnick and then Dr. Farley.

DAVID MELNICK: I just wanted to go back to a point that Dennis raised. You know, in a world where trial networks exist and there's a capability for early initiative of strategy-type trials, is there anticipated flexibility about the dissemination of that information?

You know, we need a mechanism for communicating that strategic information to the clinical community. And whether it's resistant pathogen trials or the application of a novel agent to bacteremia or something else, what -- how do we get that information out, you know, without incorporating that into the label?

MAN 2: (indiscernible) guidelines again.

CYNTHIA SEARS: Well, I don't know. I mean, guidelines is one approach but, Vance, you presented several examples where strategic trials are

published in the literature and disseminated. You know, it certainly would be helpful to clinicians to have more concentrated information so you don't have to go to individual facts. And also from the FDA documents, it sounds like there's readily available data early on that could be summarized while we wait for publication. Sorry I didn't leap up. I was thinking about it, but... I welcome other input, though.

DAN RUBIN: John?

JOHN FARLEY: Anything more on this thread? Because I'm going to change the subject to comparators. Anybody? So, Dr. Ucheo -- Uchea -- is that correct? Dr. Uchea?

CHIBUZOR UCHEA: Yeah.

JOHN FARLEY: So, we really appreciate you being here and appreciate Welcome's leadership in really focusing on clinical trial capacity in the areas where the pathogens of interest really are. So, both sort of where you're headed as well as actually some of the statistical methodologies which might allow for smaller studies -- it feels like they're

headed toward a -- using a comparator in those trials that may not be registered in those countries and may be expensive.

So, it may be early for you to ask this -- to ask this -- early to ask this question, but I wonder where you all are thinking about with that particular question.

CHIBUZOR UCHEA: Yeah, that is one of the things that we've been thinking about. In terms of comparators and standards of care, they vary quite widely within -- within regions. And especially where we're looking at as well, a big issue is availability. And that's one of the key reasons why we're looking at the interoperability of our network with other regional networks. And that will allow us to leverage the expertise in those areas and also potentially bring in the ability to use comparators that are known and have -- that we have sufficient knowledge about in other regions to be able to implement that into -- into our potential region and have access to those compounds for use.

DAN RUBIN: All right. I do have a

follow-up to that issue. Just because of time, I maybe want to give you a follow-up and then go to Aaron, and then we'll go into the moderated discussion. But some of the later questions are big picture enough that I think they can address any issues here. But do you have a follow-up?

MANOS PERROS: Yeah, it is about the comparator but perhaps in the broader sense. I'm not a trialist. I'm trying to get my head around how this works. If you -- if you enroll a patient with pneumonia, at the point where you enroll him, you don't know if it's carbapenem (indiscernible) or carbon resistance, you don't know if it's enterobacteriaceae or pseudomonas. The standard of care in the comparator, it's different for each of those.

So, how does this work? A patient gets admitted, you don't start them on Colistin, but you would put them on Colistin the moment you diagnose CRE, for instance, in some regions but not in others. How do we do that in the clinical trial network context?

DAN RUBIN: I don't have the answer to that but, Aaron, did you have a comment? Or Dr. Uchea, did you want to follow up?

CHIBUZOR UCHEA: Yeah, that is one of the big problems that we know we'll be facing. It's something that we really have to look at. And we very much would value any additional input into that. But, yeah, that's one of the key tests that we're going to face.

AARON DANE: Yes. So, I just had a question for Scott, actually. So, Scott, I really like DOOR and what you can bring, but one of the things in terms of designing a trial with that as the primary endpoint is understanding the power and the risk of the study.

So, imagine a company's running a study and the company could sort of live or die by the outcome of that study -- and how you (indiscernible) and the assumptions you have to make. Because obviously with all the different categories, you've got to make a lot more assumptions than we normally would do, which is normally just what, response rate.

So, it's not using the approach, it's more how to assess the risk of the approach in the outset. I don't know if that's something you've looked at.

SCOTT EVANS: Yeah. Well, there's two ways to think about sizing the study depending on whether you want to size it based on a DOOR probability or a partial credit approach. And then, you know, as with any other hypothesis test, if you're going to do a hypothesis testing approach you've got to come up with a null and alternative. And the null's usually pretty obvious. But you've got to come up with what sort of effect size you want to see. And it's on a slightly different scale probability -- you're better off in one treatment rather than another, or -- and you've got to either take that at a high level with just thinking at that probability level or think about how there's going to be a shift in the DOOR outcome between two treatments and try to size to detect it.

Or partial credit is basically you're on a continuous outcome. You can think of it as a --

as a difference in -- a difference in means or a difference in proportions on a 100-point scale.

VANCE FOWLER: May I add to that? A second element to respond to Aaron's very reasonable question would ideally be to standardize the DOOR by indication. And, in fact, sort of one of the early -- sorry... So, the second strategy by which to avoid some of the challenges that Aaron raised would be to standardize the DOOR endpoint such -- by disease type.

And, in fact, one of the goals of ARLG 2.0, one of the early initiatives that Helen, in fact, is hopefully going to be leading, if I can persuade her to, is to -- is to develop just that for the four most common anti-infective initiatives or approval pathways. So, intraabdominal, complicated UTI, HAP/VAP, etc. That could then be publicly available to all sponsors to utilize as -- probably an exploratory endpoint. I doubt you're going to get to, you know, a primary efficacy endpoint with a registrational setting.

But to make that available so that it's a common standardized tool that could then be employed

with whatever anti-infective agent may be at question.

AARON DANE: Yeah, I think something like that would be really helpful. And the reason I raise it is because we often have a lot of angst, even on what the response rate is. Yeah, and that's a very simple measure. And we're not sure -- and that's where a key risk can be.

So, if we're starting to move to three or four categories, which clearly makes sense, is clearly more sensitive, but then that means there's much more consideration and much more risk with that. So, anything that could help understand that in the disease areas would be good, I think.

VANCE FOWLER: And I think it's very likely that the factors, the conditions that impact one patient's clinical response for an intraabdominal infection is likely to be profoundly different from that that he or she would encounter in the setting of HAP/VAP or skin and soft tissue infection, what have you.

So, the notion would be that you'd create it carefully, validate it, and make it -- and

then essentially unleash it and make it, you know, publicly available for all sponsors to employ in their own development programs.

JANE KNISELY: Do you have a related point? Okay, go ahead.

RIENK PYPSTRA: Yeah, just to continue on that, you're suggesting to develop these DOOR criteria for the purpose of clinicians. But as we discussed before, we have other people who are interested in this. In your talk, Scott, you mentioned patients. I think that is important. But we have in between category as well -- the payers, who are very important decision makers.

So, if we include a DOOR outcome or analysis in our studies, which ones should we then include? Of course, we want to serve the clinicians but can we include another for the payers? Only to be concordant?

VANCE FOWLER: Yeah, sorry, I actually meant the intent and the intended audience of these DOOR Tools would primarily not be the clinicians but, rather, would be industry. Because the goal is to

make available a tool much along the same lines as a guidance. Think about the guidance for HAP/VAP or ABSSSI or what have you. That this is a tool that -- you know, Compound A, Compound B from Company A and Company B, rather than create their own, have to go through their own validation, etc., that this is a standard external template against which to compare the performance of that compound.

RIENK PYPSTRA: But the way to build it --

VANCE FOWLER: The audience is -- the audience is indeed industry.

RIENK PYPSTRA: Yeah. But the way to build it we heard from Scott that it was 43 profiles. They submitted it to clinicians, they came to six categories. I suppose if you submit those 43 profiles to a payer, he may -- they may come up with a different categorization. And so that's what I want to bring in here, and it's actually a question to Scott. Is that something that you've considered and that is possible or what do you think about it?

SCOTT EVANS: Yeah. Well, the example

that I showed, we survey clinicians. But one of the questions we get is well, maybe you should survey the patients. Sometimes we forget about them in clinical trials. I think -- so, two comments: One is I think you could have multiple outcomes if -- in the sense that if clinicians, or industry, or patients, or payers are interested in different outcomes, you can construct those different outcomes and analyze them in studies.

One thing that we have done already with the DOOR outcomes in various studies is when we've implemented, for example, a partial credit analysis approach, although pre-specification and transparency is always the -- you know, so important in clinical trials, when it comes to putting a value on different outcomes, it seems to be avoided like the plague.

And it's a very interesting dichotomy. Transparency and pre-specification, but when it comes to writing down a value or trying to value different outcomes, nobody wants to touch it.

Now, acknowledging the fact that,

Number 1, that not everybody has the same value system, including payers versus clinicians versus patients, or even within those categories, the 20-year old woman is going to have a different value system than the 70-year old man, and to acknowledge that.

Now, one thing that we could do, as we get our DOOR outcome, we say we want to implement a partial credit approach. So, if I surveyed the clinicians and said, well, give me your grading key, and as -- to adhere to transparency and pre-specification, I could survey experts in this field and come up with sort of a population average of what people feel a grading system should be.

However, the actual analysis can also portray that if you want to deviate from that value system and would like to do some other value system, I can show you what the treatment effect is under that value system. And we have examples of this. In the CAZ-AVI Colistin study we did exactly this. And so, there are some mechanisms by which you can get information about that value system. One is to survey patients or clinicians or whatever the case may be.

The other is to get it in the course of the trial.

So, during the course of the trial, if you implemented, for example, a quality of life instrument, and I set up a DOOR outcome. And if I look at the patients in the most desirable category, those patients give me an idea using this quality of life instrument -- they give me an idea of what their quality of life is like.

Then I look at the quality of life in the patients in the second most desirable category and that perhaps in using this instrument, that the ratio of the quality of life of those patients compared to the patients in the most desirable category gives me an idea about how to score it with information coming directly from the patients themselves about how you impacted their life.

So, there’s ways I can get patient information, there’s ways I can survey, if you want to go to payers and other people, or go to clinicians who treat this disease, you can get that sort of information. Don’t be fooled that all of them are going to have the same answer. And so the idea is, is

to collect that information, design and analyze trials using that information, but acknowledge the fact that people value things differently. That's life. That's the way it is.

But I can show you how two treatments contrast each other according to any particular value system we like. And we have examples of this.

JANE KNISELY: Okay, so I have Rebecca over here has been waiting patiently, then Helen, then Nick, then Roger, then Pam, then John. And we're all on the DOOR question.

DAN RUBIN: And Sumathi told us to just keep this discussion rolling since the next questions are kind of related. We're getting at those anyway.

JANE KNISELY: Great. Okay.

REBECCA REINDEL: Just to follow on the DOOR tool, which I think is a really interesting way to look at benefit risk together. In the setting of a known safety risk for a product where you use the partial scoring system to assess the benefit and the risk of that specific safety concern -- with the development of tools that are standardized for a given

body site not allow you to be flexible in terms of assigning risk for specific safety concerns for a given product.

SCOTT EVANS: I'm not sure I completely understand your question. But one point that came up as I was trying to understand your question. When you're identifying endpoints and constructing a DOOR -- so, the big concern in the room is if I take -- if I try to construct some sort of ordinal outcome, could I manipulate in such a way that I make my drug look better than yours because of the way I construct it, perhaps favoring some components of outcomes more than others.

So, the outcome you're after is the outcomes that are sort of important and valuable to patients. They're not necessarily -- although, obviously, you're evaluating toxicity associated with particular interventions. But what you're measuring is patient outcomes and things that are important to the patient. So, I want to make sure that when I construct it, I'm not just including safety events that happened to occur with this intervention, which

would -- but not this other intervention. Because if another intervention, your control intervention has got their own sort of toxicity risk, if I either emphasize those or deemphasize those, I'm not playing a fair game.

So, what I want to measure is those things that are important for patients or that we see that perhaps patients aren't aware of, but do it in a fair way, not so specific to the intervention, obviously there's some toxicity concerns you want to do with that.

But if you're going to be transparent and open about how to construct it, of course, the concern of regulators and the concern of the audience and people is that you don't build an outcome that is artifactually going to make you look better than the next guy because of the way you do it. It's got to be about meaningful evaluation of what's happening to the patient.

HELEN BOUCHER: So, I just had a small comment about the ability of DOOR, by route of Scott's comment, to include aspects of the outcome that are

more safety related and for which there are values that payers care about. And so, as opposed to the traditional non-inferiority trial where we have kind of the standard safety display, DOOR could provide the ability to measure in clinical meaningful ways that payers care about things like renal failure and things for which there's not only mortality but also cost, that people agree on that could help add to the value of a product, even if it doesn't add to the indication, specifically.

NICK KARTSONIS: So, I've been thinking about this door thing for many years now and I actually find it very intriguing and actually think it's a fascinating way to kind of think about marrying up both the efficacy and safety aspects that kind of go into the development.

And I do agree, I mean, trying to get the patient perspective and the payer perspective is important. But I will tell you with every country now having its own HTA agency trying to get agreement from payers is almost impossible. But, I mean, I do think we're also in a very unique situation that we've done

all these clinical studies over the last decade and we should learn from them.

And so I guess I had two questions, and it's probably more, you know, directed to the FDA and their perspectives on this. One is you guys have had conversations with other regulatory agencies. Do you think they would be open -- I'm trying to think about this now, also trying to make Japan happy and Europe happy -- would they be open to that?

And second is, is there a possibility to use all these datasets that have been generated that you guys have access to and do these sort of analyses in a way that doesn't de-identify the comparators, if you know what I'm getting at -- but helps provide information that has value from an endpoint standpoint?

So, I know it's a bit of a loaded question but I'm just curious. Because I do think it's an interesting way to think about things moving forward.

SUMATHI NAMBIAR: So, for the first question we haven't had a lot of discussion about DOOR

endpoint. At least -- more recently we've been involved with one product. We've had a couple of discussions. And in one of those discussions the EMN was part of the meeting, so we met with the company together.

I don't think we heard any sort of concerns in broad terms but it's more the details. And that's exactly what we were struggling with -- you know, what were the different components of the proposed endpoint? And so I think there is certainly a willingness to work. I just cannot speak for them and say they will agree.

But what I did not hear at the meeting was we just cannot move forward with this. But I think for us, the bigger struggle was really defining which of the components were appropriate and how one could actually discern because it was so -- again, I don't have a lot of experience, but this one example that we worked through, I think, between the different categories there was very little separation. And the subjectivity that was involved in putting them in one category or the other was certainly concerning.

But there is room to improve and refine that and we're working to do that. So I think -- Dan, is that a fair assessment?

DAN RUBIN: I think that's fair. And as far as using the existing databases to try to come up with a better outcome, I mean, there's some precedent for that through the FNAH project to look at different datasets and try to define endpoint, but I don't think there's been any discussions about using datasets to try to inform more ordinal types of scales.

SUMATHI NAMBIAR: So, as a follow-up, Nick, so you meant not de-identifying the data actually -- and that's where it gets a little difficult, right?

NICK KARTSONIS: That's always the tough issue, right? I mean, I'm speaking obviously only for Merck but I imagine my counterparts might feel similarly around it.

But on the flipside of it is it's hypothesis generating, right? It's not hypothesis, you know, proving. So, in some ways you could make a

case for helping to move the field forward. And you've done that, right, with -- thinking about non-inferiority margins and how you guys came up with M1s and M2s and all that, which was incredibly helpful for the field moving forward. So, I just raise that as --

SUMATHI NAMBIAR: Good thought. Certainly something we can talk about.

ROGER LEWIS: So, I think the discussion about DOOR and the difficulties of, you know, everybody agreeing on, for example, the relative weights or even the ordering of different outcome states is interesting. I think it's an example where the more sophisticated approach to patient outcomes illustrates problems with existing outcomes that we've been able to not have to face.

So, for example, if you use 30-day all-cause mortality, some people die in three days in the ICU, some people die at day 29 in the ICU, some people wish they were dead but they're still alive in the ICU. And different people would value those outcomes in a different order. But somehow 30-day all-cause mortality just feels really solid. But it's the same

illusion.

So, I think the discussion about DOOR is really important and illustrative but I wouldn't want it to be seen as a criticism of the approach. I think the approach is absolutely solid. The right thing to do. And at some point you just have to -- for the purpose of defining the endpoint of a trial, defining the statistical characteristics and what defines an adequate and well-controlled trial, agree on what the order of categories are by disease state, just as Scott suggested.

But the fact that we don't have those same level of discussions for the rich outcomes that occur in the people who do or do not die in 30 days, that's just because we're choosing not to have those discussions, not because the same considerations don't exist.

JANE KNISELY: I think I have Pam next.

PAMELA TENAERTS: So, I think this is a fascinating way to look at a novel endpoint and I really think there might be a need for something to supplement the mortality, which is sort of

unidimensional in many ways.

But what I would say is there's opportunity for confusion if you do DOOR as per the physicians, DOOR as per the patients, DOOR as per the payers, and then you're kind of like, well, which ones weigh more in the decision?

So, my thing would be well, combine them all and come to a consensus where everybody -- it might be lower but everybody can agree to it. And I think that might be, you know, baby steps instead of sort of setting a new thing and potentially creating additional confusion.

VANCE FOWLER: Yeah, agree. And, in fact, one of the sort of work streams that's intended is obtaining the -- essentially, a patient-assessed quality of life assessment for those same four indications. We've already done it with staph aureus bacteremia and gram negative bacteremia where we establish quality of life in those patients.

And what was striking was two things. I guess the clinicians in the room would probably guess that the impact -- my hypothesis was that staph

aureus is going to -- it doesn't tend to kill you right away, it just sort of chews on you a long time. And that's exactly what you found -- that these patients who had ostensibly had been evaluated as cures, when you interviewed them extensively, you know, six weeks later or what have you, they were talking about, you know, terrible impacts on their lives. Having to wear Depends because they can't get to the bathroom in time. I mean, a profound impact on the quality of life. And these were ostensibly cures, what would've been cures.

So, and for gram negatives, it was really all about the first two weeks. If they're done in two weeks, by six weeks their quality of life is much higher than that with staph aureus.

So, yeah, incorporating that element, that's, again, part of the intent of creating this DOOR endpoint, is building patient assessment into that step forward.

JANE KNISELY: John Rex?

JOHN REX: So, the build I've got around DOOR here is that there -- I hear the plea

about people don't want to rate the number -- they don't want to put a number on it. There actually has been work done that I think maybe we just need to leverage it better. And I'm looking right now at the University of York and the University of Sheffield report that the U.K. is using to construct its value arguments around conducting a pilot program of buying new antibiotics. And there's a whole bunch of stuff in here about health-related quality of life and how you might think about measuring that in the setting of any infections.

And I've not decoded all of it and thought it through, but it actually makes me think further back to the ERG report from three or four years ago where I first learned about the value of a statistical life here and the way that the actuarial community estimates the value of staying alive for a period of time, and it feeds into the quality adjusted life here concept.

So, there are some tools that exist that would be familiar to the health technology assessment world that we may not know about. And so

my thought here about DOOR is, it's very interesting -- and that plot about the vancomycin showing that the lower exposures were actually net better, at least from my perspective, extraordinary.

So, this notion that the ID community needs to start folding in some of those measures -- and I'm just compliant that some of it exists and that's the theme here. We need to borrow that information and apply it here.

AARON DANE: I was just going to say in relation to that as well, the other area where there's a lot of work done with quantitative benefit-risk assessment, which I think we should look at as well for the same reasons. Because that's all about being very clear on what the trade-offs are, and you're very explicit about how you weight them and you can look at what happens if you change the weighting. So, I think, again, that's something which is similar which we can draw on.

SCOTT EVANS: Yeah, just one comment to that. Thank you for the information about the quality of life. DOOR has been used in other disease areas,

including stroke prevention trials, for example, where there has been a lot of quality of life valuation about what happens to patients -- strokes, MIs, and so forth, where primary endpoints trials are time to stroke, MI, or death, or maybe something else.

But deaths are worse than strokes,, and strokes with permanent consequences are worse than things that are transient, and you can factor in bleeds and all that sort of stuff. And there has been scoring systems set up in that area and/or analyses that have used that information.

AMY LEITMAN: So, I wanted to address a few points that were discussed just from the perspective of somebody who’s sort of getting entrenched in the world of PROs and PRO development because it’s difficult to get patient feedback and looking at the different kinds of infections. Survival is not enough for a patient. That’s really the goal, right, when you’re treating a patient? Generally speaking, you do want them to survive.

And I heard what was said before, that, you know, what a patient wants when they’re 20 and

what a patient wants when they're 70 or 80 or 90 may be different. And you may get to a point where a patient decides that they are not willing to continue on a journey. But, generally speaking, the goal is for the patient to live, to survive the treatment and survive the infection. But you're not treating an infection, you're treating a patient. So, how they survive matters.

And I think what we see when we are dealing with patients, a lot of times what we see, and maybe this isn't seen as much with the ID physicians because the follow-up is done by another physician probably. And maybe it should be done more by an ID physician so they could assess how the patient is doing. And it might be an easier way to catch if there's going to be a reinfection, there might be earlier signs that a physician could catch.

But you have to look at how that patient is doing. And so often we see them and it's just -- they still feel like a train wreck. And it's months and months later -- and this is with all kinds of respiratory infections, and then you have to deal

with skin and soft tissue infections. If there's been a surgical debridement, don't assume that there's no permanent impact. They're worried about disfigurement and they have to live with that impact.

And, you know, we actually deal with that with NTM. We catch the patients who are disseminated NTM and, unfortunately, they've gone to those, you know, cosmetic surgery centers and gotten those infections and I mean, there were I think 17 deaths in New York City alone last year from that. And that's, I'm sure, not the only pathogen that's causing these infections.

So, you really -- you know, even if it's like an abscess on a limb, if it gets really bad, at some point you start thinking -- they start thinking am I going to lose my limb? Like, what's going to happen? So, there's a lot going on.

And then after that, let's say it's something with an abscess, you clear the abscess, it heals. Every little thing they become paranoid about. And it's something -- they get a bug bite and it's, oh, what's going to happen? Is this going to happen

again? What am I going to do? Are they going to be able to treat it?

So, it's really important to understand the patient journey, not just at the start of it, not just while they're going through the critical event, but afterwards. And really not just a month afterwards. For some of these patients it continues months or even years afterwards. Their body has undergone really a brutal assault. Some of these medications are really hard to take. Their immune systems have completely gone haywire trying to fight this infection. They're exhausted. They probably are having some kind of a nutritional deficiency, because when they're fighting an infection, they're just burning through calories.

You know, one of the things that our organization is starting to develop is a pamphlet on mental healthcare and a pamphlet on nutrition care for infection because those are things that patients have asked us for. And when I've talked to -- you know, I've talked to a sepsis survivor who was like, you know, the nutrition thing would probably be helpful

for people like us, who, when we're coming out of this, they've lost so much weight and they need guidance on not just how to put on weight but how to do it healthfully.

Like, you can't just pile in Ben & Jerry's as much as we'd like to. They can't just keep eating Ben & Jerry's all day. They have to figure out how to balance... You know, yes, you can have some of that but you have to balance the good nutritional intake and put those calories back on, and give your body what it needs to heal as well.

There was one other thing I wanted to talk about with respect to payers, and this is something that I'm seeing a lot in the patient advocacy community. So, the Institute for Clinical Effectiveness & Review, ICER, has a standard that they call Quality-Adjusted Life Years, QALY, or "Qually" as we call it. They start looking at various treatments and whether those treatment costs are justified by quality-adjusted life years.

And one of the things that we're dealing with now is that antibiotics are not

compensated well enough. And we're looking at developing these novel treatments, and it's really expensive. And these treatments are going to be expensive, and it's justifiable considering the cost and considering the value to society.

But we're talking about some of these infections like they are rare diseases. One of my concerns is that we'd better start bringing payers into the fold on this conversation as well as patients because I can guarantee you at some point down the road, ICER will turn its attention to some of these treatments and say, well, but, you know, two years later, they're still -- they're still really sick or they're getting sick again, they're getting pneumonia again. That is not -- I mean, believe me when I tell you that patient advocacy groups will kick up a huge fuss because any time this happens, it just seems like a particular group ends up getting targeted.

The patients suffer most. It's really up to the patient to decide what the quality of life is that they want afterwards, you know? Is the quality of life that they're going to get from the

treatment acceptable?

And I just wanted to -- you know, Roger's point about 30-day all-cause mortality really goes back to measuring out over time with PRO. Measuring out 30 days isn't enough. Measuring out 60 days might not be enough...90 days. So, it's -- we're at that point now where we have to start asking these patients what is your journey like? How long is it taking? Because when we start incorporating patient input into these clinical trials and we start measuring with PROs, one of the most important questions -- and this is the one that we're grappling with now, it's one of the many -- how long do we measure out?

And it's critical because it can tip either way when you're trying to decide do we approve this drug or not? Because you're looking at it -- well, is there clinical benefit? You may not see the full impact on the patient for 12 months or longer. It may take them that long to recover from any kind of infection state if it's that serious. So, we need to know for sure like, is there an extended period of

time that we measure out? So, we need to start finding these patients and talking to them, and listening to them.

JANE KNISELY: Thanks. I have Sara and maybe -- well, you're making your comment. Ah, thank you.

SARA COSGROVE: So, I wanted to circle back to Rebecca's comment about toxicity real quick with DOOR. Because the two big studies that have been done with DOOR related to toxicity have been done with Vanco and with Colistin, for which we have the luxury of 60 years of knowledge that both of those drugs are nephrotoxic.

And I think unless you have a signal already from a Phase I or II study, that if you're trying to do a DOOR analysis along with the registrational trial, you may be missing significant toxicity issues when you construct the DOOR. And I don't know if you've thought about those issues or how to handle it, or maybe that's a flaw that's tough to deal with and maybe DOOR is better done after, you know, the registrational trial.

VANCE FOWLER: Yeah, I totally agree. That's a great point. And it's one of the reasons why I feel like the ultimate contribution of DOOR, at least in the short term, will be in exploratory or secondary endpoints. I think the primary -- again, with, you know, the responsibilities that the FDA is charged with in terms of evaluating safety and efficacy -- I don't see a way around the safety component not being primary and not being evaluated independently for precisely the reasons you brought out.

So, I'm suggesting that DOOR Has a meaningful contribution for drugs that are already available and trials that are -- strategy trials, we can consider it as a primary. But I think for the purposes of evaluating new compounds, safety is going to have to be paramount. And, Helen, did you want to step on that?

HELEN BOUCHER: I would just add that by the time gets to Phase III, we usually know something about the target organs of toxicity from the preclinical and the early clinical. So, that's the

kind of toxicity we know we might be facing with this drug or class of drugs. Then there's that which we don't know, which Scott was talking about. And I think the DOOR, if you're going to apply it in a Phase III setting has to think about both. But the fact that you can do that is actually more than we can do in some standard trials. So, that's where there might be a benefit.

SCOTT EVANS: Maybe if I could just follow up. I think -- one thing -- a couple of points. One is if you use a DOOR endpoint you should analyze all the components to it. In and of themselves, they may or may not be interpretable. As I mentioned, you know, you measure duration of hospitalization, whether you can interpret that, you know, it depends on what's happening with other things.

I do think, though, that when you do get to late phase trials you will have an idea about toxicity. That's what the early phase trials are for. However, I do think that in the long run, the vision for a DOOR outcome, that the safety components that

weigh into it should not necessarily be intervention-specific.

So, if you're comparing two treatments, one may have a toxicity in one place and another has a toxicity in another place, the importance of a relative importance of those two things depends on their impact on patients. And so in some ways you want to be agnostic to sort of intervention-specificity. What you want is an outcome that evaluates the impact on patient lives.

One frustration I've always had in trials is we have all these rating systems for AEs -- severity, seriousness, is it related to treatment? Is it not related to treatment? Treatment-emergent, etc. What I would like to see is a rating that says how impactful is that adverse event on the patient? And that's actually what you would like to factor in.

JANE KNISELY: Sue, did you have a related point?

SUE CAMMARATA: No. Actually, I wanted to talk about clinical trial networks.

JANE KNISELY: Okay.

SUE CAMMARATA: But I didn't know what the time you'd want, because according to the agenda, it looks like this is wrapping up at some point here soon.

JANE KNISELY: Yeah. We are well into our moderated panel discussion, I think. So, let's just go ahead and note the discussion questions. So, we talked about a lot of things this morning. DOOR was a very interesting part of that but not the only part. And so I think we should open this up now to additional discussion. So, please go ahead.

SUE CAMMARATA: So, for me, I wanted to talk a little bit about the clinical trial network. Because I think on paper that always sounds fantastic in some ways. And I was actually part of the Pew meeting -- if it was out of 2016 or 2017, it just -- time flies. And I recall at that meeting, and I think it still is an issue, is what is the design and intent of that trial network?

If I recall, most of the pharma representatives were quite concerned about things like -- I'll say master protocols and that because of the

challenges of a particular drug, particular comparators, time, you know, how would you use this data?

I would say that I'm particularly happy to see that -- at least the Pew Trust is talking about pediatrics. I've said at that meeting and I still say it's not sexy, it's not exciting. We've talked about pediatrics. But if you want to be successful, I truly think that that is the one place3 you can start to get pharma potentially involved because it's less competitive, there' less risk. It's not really a money-making area for most companies. But it's one that every single company, whether you're big or small, has to...

I mean, I'm currently responsible for three pediatric programs, you know. So, it's something you have to spend money -- and it might be an area of the most success, but it absolutely is not sexy to everybody else except the pediatricians in the room. But it is one area that I would highly recommend thinking about because -- for any of the trial networks.

But I am concerned and I'm curious in this meeting, compared to that previous meeting, about the thoughts of everybody else about master protocols. I mean, for me, a clinical trial network that might be able to be up and running, cost effective, just to get a contract signed with an academic institution in the United States can take six to nine months. And that is always one of the other challenges of being able to do trials in the U.S. So, having that available with a site that's motivated is interesting. The problem is I'm not sure there's enough volume over time to do some of these trials.

So, for me, I was interested in this concept of sort of the idea of the strategic trial that might not be pharma-sponsored but might be sponsored in some other way to get information to keep some of these sites open. For example, in HAP/VAP, to answer interesting clinical questions that clinicians want to have answered but then also could be a site that participates in a clinical trial since -- for example, with HAP/VAP, the guidance is fairly clear in the general construct. But a company or a sponsor

might want to have something different because of the nuances of their product.

So, I'm just curious about h ow other pharma companies feel about it because I see lots of discussion but I'm not sure about the reality, for example, of a master protocol type system being set up versus something that's more flexible to handle both those strategic kind of questions and then the registration type of questions.

JANE KNISELY: Thanks for bringing that up, Sue, because I had the same question. I also was at that 2016 meeting and heard the same thing you did. So, we heard a little bit of it yesterday from David, so I'm curious to get some perspectives from companies about what are their perspectives on these clinical trials networks? How could they be useful? Sort of this third question here. So, go ahead, Nick.

NICK KARTSONIS: So, thanks, Sue, for raising that. In fact, the one area of it we've talked about internally where we'd love a network was pediatrics. I mean, because I do think you've raised all the right issues around that, that I think would

have value in terms of that.

You know, I sort of harken back to our recent experience with HAP/VAP. And each of our three studies was different. And we actually talked about this. Could we have done this in the network situation? But they each looked at different endpoints and they had different visits to some extent, and just some of the inclusion-exclusion criteria also varied a little bit.

You know, there's a host of things we worry about around the network. You mentioned master protocols and the factors that go with that, but then there's all the operational stuff of different safety tests for different drugs, how do you handle that? Blinding, comparators, how do you handle all of that stuff? And then there's the inevitable database issue, which, every company has its own database. And I don't know how other companies are but every time I feel like we've run a database, the lights dim a little bit at Merck, then they come back on. (Laughs)

And so it isn't without its shares of challenges of trying to share that data across the

databases. So, just mention that.

JANE KNISELY: Maybe we can hear from Rienk and David, and then go to Chibuzor.

RIENK PYPSTRA: Yes. I also agree that pediatrics is a very obvious one. I think HAP/VAP is also a possibility. It all depends on the level of integration. If you have a network of independent sites, you can indeed run a common protocol and you can organize yourself, and you may have some savings on the patient numbers and therefore on cost.

But you could go with the integration all the way to having a common database for the whole network. And the patient data that you're collecting in a continuous fashion are immediately available in that database. And then the step of adding an investigational or an interventional arm in that becomes much more simple.

It also overcomes some of the issues that we had about informed consent that were discussed this morning. If you were such a site, every patient admitted to your site would enter the hospital or enter the ICU. You can determine yourself where you

put that barrier. And at admission, get a questionnaire. We are a research site. Would you agree to participate in research?

If the patient says yes, from that moment on, the data of that patient go into the database. Whether you will or not need that patient doesn't matter yet. And then when you have a specific protocol, you re-consent the patient. You are now qualifying for this intervention. Do you still agree to participate in the research? That would be something that would facilitate it.

And such a network would have quite a lot of advantages. First of all, it would be of a very high-quality standard. The data that you're collecting in each site would be comparable because that's how the database has been set up. You don't have different physicians collecting different types of information in their patient loads.

So, I think that would be a very important homogenization of the data and making sure that we can compare the data across sites.

DAVID MELNICK: So, I agree with much

of what's been said. I think -- you know, perhaps we could uncouple the idea of a trial network from a master protocol. We don't have to do the whole shebang at once, you know. And there is the obvious saving -- you know, I love this term of a single use network.

We reinvent the wheel with every one of these trials. This has confused me from my start in industry, you know, 25 years ago. What do you do when you start a trial? You look at the FDA guidance and at the three last trials that had been done and, you know, you write your protocol and then take it to the agency and get it approved, and you're off assembling a trial network.

But those frontend activities -- negotiating the clinical trial agreements, you know, it's an incredibly laborious task and we pay an arm and a leg to CROs to do that work for us. I mean, it would seem to me that having a shared infrastructure that could include things like shared SOPs, you know, common CTAs, and then gradually move in the direction of, well, if there's a HAP/VAP network, we work on

endpoints. They're going to work and gradually move toward a master protocol that would be acceptable. So, it makes sense to me.

CHIBUZOR UCHEA: I totally agree with you, David. The two concepts are independent -- they're complementary but independent. We really need to be focusing on the initial development of the network, and that's what we've prioritized. The use of continuous master protocol is a brilliant idea. It's one that requires -- it's much more resource-intensive. We're a bit of a way from that at the moment. It requires further development and assessment of the investability.

Going back to clinical trial network in general, the one that we're setting up in Southeast Asia, we really want it to be able to facilitate parallel follow-on studies but the smaller type and optimization studies. And that's one of the reasons that I'm really happy that people are talking about pediatric indications, because that's something that we see as a key area that we can achieve a lot of buy-in. We want to get away from the whole waiting ten

years for a product to go through and start studies in pediatric indications. And we're working with Penta to try and leverage their expertise in this area and be able to build up our network to be able to run studies like that.

And I'll also go to a point that Vance made earlier, which is really important, about funding. We really need to be able to try to leverage government funding for these kinds of projects. They're heavily dependent on philanthropic organizations. But the resources needed, they're really deep.

Another way we're setting up our network is also for our investigators to be able to go out and get their own ground funding as well. So, our secretariat function will be able to support them in proposal development.

And I'd like to go back to a point that was made yesterday about bureaucracy, the potential for this to be like a heavily bureaucratic way of operating. One of the key reasons we're operating with a network secretariat is it will effectively work

as the external business face of the network. So, it will provide a single point of entry.

So, the whole idea is moving away from that bureaucratic approach where you're constantly trying to find various different trial sites. It's one point of entry, and then we build from there.

JANE KNISELY: Okay. So, a follow-up point on this topic? Sue? Okay, go ahead.

SUE CAMMARATA: Well, one of the questions I did have for you around the pediatrics -- and this is somewhat off topic but I'm interested to ask -- is around using those sites for drug development. Because, typically, you try to avoid going to places where you're not going to commercialize. It's just the ethics of it all. And -- but, unfortunately, with resistant bugs you have to go where you can go.

All the companies would love to be commercialized globally but it's not necessarily happening. So, is there some arm of that involved in this discussion that the Pew is taking about -- for products that are eventually studied? And then some

way to have access -- especially for the small companies that don't have partners in those areas.

CHIBUZOR UCHEA: In general, access is something that's really important to our organization and it is something that we're looking at. The key for us with selecting Southeast Asia as our anchor point is the access to public patient populations with hyper (indiscernible) disease and being able to look at the multidrug resistant and extensively drug resistant indications.

It was really refreshing to hear Steven's talk yesterday to see that this is being further explored, and that the data presented less concerns around generalized ability because that is one of the key reasons why not a lot of clinical research is done in these areas.

And what we're trying to do is we're trying to leverage that, we're trying to build the capacity, improve the trial site so that regulatory bodies are happy with the data. The big thing is about the value of ex-U.S. and ex-EU data. So, initially operating this trial site in this region

will help us to be able to leverage that and to show the potential advantages of being able to run clinical trials in these regions.

SUE CAMMARATA: This is actually a follow-up but it's directed more to the U.S. clinicians, because I keep hearing about the concern about not U.S. data. And I think for most of the pharma folks, and FDA, and other agencies, they're all comfortable with -- there are differences but this data is translatable.

And so I'm just curious because I know that there is this concern about the lack of U.S. patients. But I did a trial recently where we went out of our way and spent extra time and money to try to enroll in the U.S. and we got five patients versus 860 outside the United States. And that was a significant effort. We actually delayed our timelines to try to get those U.S. patients because everybody talks about it.

So, I'm just curious since many of the people here actually accept that. But I keep hearing the U.S. talk about it and I'm just questioning --

concerned or questions about how you feel about that. Because that is where I think, for the bugs, I mean, that's where we're going to have to go.

VANCE FOWLER: Yeah, I mean, I'd put that on the scale of concern about -- right up there with like a meteorite hitting the planet. It's just not -- it's just not a real concern. I'd much rather have, you know, data than not data. And, okay, there's going to be practice variation. I mean, we're setting up sites and got sites going on in China and all these other places. It's definitely different.

But guess what? You can randomize folks there. You can block randomize. You know, there's means by which you can address that. And at the end of the day, if you're -- you know -- so, no, it's not -- to me, it's not a meaningful concern.

JANE KNISELY: Additional responses?

SUE CAMMARATA: I was just going to say the same thing. I don't think -- I actually don't even hear this as a concern brought up amongst other infectious disease doctors. I think people would much rather have data in patients with resistant organisms.

HELEN BOUCHER: So, I will agree with my esteemed colleagues, but also just raise the point of view that we're in America, the most resource-rich country, and the question for us is, is it acceptable that our patients are not in these trials? And the questions are being raised by payers, and journal reviewers, and editors, and others, when there are zero or five out of 800 patients, there are -- some people think there are ethical issues.

So, I would just submit that it's at least worth of our consideration before we say, well, we're just going to give up on this enterprise in America.

VANCE FOWLER: I'm pretty read to give up. I mean, I think I'd much rather -- I'd much rather go ahead and get -- actually try to get some trials done and get some answers and some data somewhere. I know, I get banged around about the journalizability and all the ex-U.S., yadda-yadda, but for crying out loud, if we weren't in a crises, we wouldn't all be sitting here and investing two days of our time on this issue. And that's from all sides. I

think -- we're in a crisis, folks. The building's on fire, you know, and we need to do something. We need to get some data, we need to understand how to treat patients, do the right thing. Crawl, walk, run, fly. Thank you.

JANE KNISELY: John looks like he has a burning response. Pam's had her card up forever, so we'll go to her after John. And then I think we do want to give the audience an opportunity to ask some questions. So, if you have one, please come to the microphone and we'll continue to work here in the meantime. So, go ahead, John.

JOHN REX: So, Helen, are you concerned that we're in a position of looking like we're using the rest of the world as our guinea pig? I mean, is that sort of the idea you're getting at, that we need to be sensitive to that notion?

HELEN BOUCHER: I think that's part of the issue, and I think it's also the question of whether our patients deserve the chance to be included in these trials. And what we're seeing is that they're not.

So, I totally hear Vance and Sara and the need to get on with it, and we cannot let the perfect be the enemy of the good.

LINDSEY BADEN: But there are multiple problems to be solved here. I mean, we need high quality data, full stop. I agree with Helen. It's a -- it's a shame that that can't also occur in the Americas. But I don't think they're exclusive. On the other hand, we need high quality data more quickly.

JANE KNISELY: Okay, Pam. Sorry.

PAMELA TENAERTS: I'm actually not a patient person but... That's actually the point I wanted to make. So, with City, we started this effort in antibacterial work. A lot of the comments I got from people in the field was, well, we've been talking about this for ten years. What is going to be different about this time?

May I remind you that that was in 2012? And it feels like we're still having the same discussions over here. So, what I would like to say is, you know, we are underestimating the risk of what

we're doing right now and overestimating the risk of novel approaches. And I really would like you guys and girls -- because I've been told that you can't really say "you guys" because it's discriminatory -- that all of you sort of have an open mind. And when people have new ideas, to test those ideas. Make funds available. And I'm looking at the NIH because I don't know where else the money's going to come from. But to test those opportunities for new ideas. Because to just go off and do new things -- new isn't always better. Because if we thought everything was going to work then why are we even doing the clinical trials to begin with? Because everybody is convinced their drug is going to work when they go into development, right? Otherwise you wouldn't do it. But we've been proven wrong.

No reason to think that some of these things that we think are going to work to make this better may be wrong too and have unintended consequences. But I would like to say, like, just keep an open mind and start doing something. Maybe just start, and that'll change things.

I mean, you guys have come a long way, because when I talked about master protocols in maybe 2014, I heard, oh, that doesn't work in our field. So, I'm not saying -- but there's other things. Beg, borrow, and steal from other specialties. I mean, you know -- I don't know, that's just what I would like to say.

I think what we find the most difficult with the work we're doing is coming up with recommendations -- as people, you know, go back to the status quo, it's easier to sort of do what you're always doing than do something new. So, anyway...

JANE KNISELY: Roger, you've also had yours up for a while.

ROGER LEWIS: So, I wanted to make two comments about the issue of out of U.S. data and applied to U.S. patients. The first is just a request for statistical clarity when we're discussing the differences. There is a tendency, at least in informal conversation, to confuse differences in the overall average outcome of patients in different locations from the expectation of the treatment of

fact.

And the thing that is scary in terms of the use of out of U.S. data to inform U.S. regulatory decision making is if you think the treatment effect is heterogeneous as opposed to the background success rate of disease. And I just urge us when we discuss this, regardless of the position we're taking on it, that we're very clear to make a distinction between differences in prognosis and differences in efficacy.

The second point I would make is that -- and I think Helen was also referring to the Belmont Report principle of justice, that the population in whom the risks of research are borne should be the population that benefits from the results of that research. I didn't say it very well but it's in print.

That can -- ideally, there's a one-to-one correspondence between the population on which the experiment is conducted and those who benefit. But you can -- when that is infeasible, you can partially address the inequity by making sure that the benefit of the research is made available to the population in

which the research is conducted.

So, I don't want to -- I'm not in a position to comment on the feasibility of doing these studies in the U.S., and if doing so delays the availability of the agents globally, that's an important negative aspect for everybody. But we should address the question of whether these agents that are being developed are then available in the countries in which the research is conducted.

RYAN CIRZ: I'll be brief because that was essentially what my point was, was hearing... Well, I guess, first worrying about ex-U.S. predicting U.S. I mean, we do use mice to set our dose initially, if that makes you more comfortable. And, generally, if you are really rigorous, you get it right every time. We don't see a lot of failures except for sort of dosing errors that a lot of people think could've been predicted.

But that is a worry, and I think about this a lot. And regardless of whatever you thought about the plazomicin study, we did most of it in Greece and the drug has not been back to Greece in

years. And whether you believed it had that effect or not, those patients don't have access to it. And that is what's going to happen. And there is that little bit of risk of saying, let's go study it over there where we don't have a lot of it. Let's go study it over there and then we'll bring it back here to protect ourselves. And I think someday we've got to be prepared for how to handle that properly.

DAN RUBIN: All right, so we have another online question. Roger, this is a follow-up, I believe, from the earlier John Tamico question about borrowing. And remember, the earlier question was about CIAI, and the follow-up is that "The issue is not just a difference in treatment effects but in noise from source control and anticipated small numbers of patients at each body site. So, it is a body site issue, not an antibiotic site issue. Should this disqualify CIAI?"

ROGER LEWIS: Great, thank you. So, the hierarchical model takes into account the signal to noise ratio within each of the sites or the subgroups. The signal to noise ratio is influenced

both by, obviously, the number of observations you have within that group and also the variability you see.

So, if there's a source of nonbiased noise, just randomness in the adequacy of surgical site control or ancillary therapy, that can be handled pretty well within the model. If what's happening is the source of noise systematically makes a drug ineffective or, I guess, conversely, more effective -- but ineffective in a site, then that's an issue because then the site really isn't exchangeable. You know something's different about that site than the other sites.

And, statistically, it's completely analogous to a setting in which you know the drug doesn't get there. If you knew the drug doesn't penetrate the meninges even when they're inflamed, it shouldn't include meningitis as one of the infection types in a hierarchical model because you know it's different.

So, this is -- there was a brief comment in my presentation where I said that the

decision to include a site, or it could be a disease type, in a hierarchical model should be based on pre -- in learn phase or earlier data that shows that it's reasonable to think of this collection of diseases as potentially ones that would respond homogeneously to the treatment effect. If you know something's different and you're asking a separate question and you shouldn't be including them in the model.

DAN RUBIN: All right, I think that's going to close our moderated discussion. So, thank you, everyone, for making this a great event. And our next item on the schedule is Dr. John Rex is going to provide a summary.

JOHN REX: So, let me first say that this has been a really instructive day and a half. And Sumathi and I put our heads together and thought about the stuff we've learned the last day and a half and where it might go. And I'm nominated to talk about it, but this is really a group work product. So, all of you are here as well.

So, these are the big messages we're going to cover: AMR Enterprise in crisis. We can't

fix it all today but some things we can. What are the emerging ideas? What do we do next? And so, the fact that the AMR Enterprise is in crisis, that the Check Engine light is flashing was nicely demonstrated by several of the talks.

Late stage commercial failures have occurred and seem likely to continue. And even when successful, your stock gets shorted and your NPV goes down, and the return is less than that of even a moderately successful oncology product.

So, what are the elements that are within our grasp? And this was really -- I learned some things very helpful here. So, we know that push funding works, actually. CARB-X, BARDA, NIAD, Wellcome Trust, Novo REPAIR. They have lit a bonfire in the preclinical space, and there are some neat things coming. Can the science problem be solved? Can we find new antibiotics? I think the answer is yeah. It looks like we're going to.

We've also learned that we can reliably get products to approval -- many products to approval, with basic studies in well-understood infections. And

helpfully, these studies generalize reasonably well to the U.S. We just had a good conversation about that. Some more detail could be added here.

But, importantly, we can't generate the same quality of data for all uses. And other possible uses of a new drug are always going to exist and are important to clinicians. Five o'clock this afternoon you may have to treat a patient who's never going to have a labeled drug for their meningitis. And they're -- just pick any one of those things. There's nothing labeled for it so what do you do?

So, this flips around to what do we want? And broken down here, it's by important stakeholder group. So, physicians and ID physicians want access to all of the data, preferably interpreted for them in some way. Payers and P&T committees would like that same thing but they'd like a measure of the quality of the data. Patients would like us to hear their voice and think about how they feel net of the whole process.

And companies would like validated and acceptable mechanisms for promoting based on the data

on resistant pathogens and difficult infections. You know, we get it -- the CUTI's not where you want to wind up but it's hard to generate the rest of it. And FDA had to label per regulation. And remember the phrase, adequate and well-controlled.

So, two slides with ideas. The first one is -- the broad title here is Tell the Story. Theme A: Make clear to ourselves and our peers the limits on data generation. Everybody needs to understand what you can do, what you can't do. It's very easy to wish for me, especially when you've never -- you can't feel the complexity of producing the information. And if there's a better way to generate this data we would be doing it. We've searched long and far to find ways to generate the data and we are constrained.

We as a community also have to be very clear on the limits on the product label. And without rules, we would have the Wild West, arbitrary decision making. And the role of the FDA is to consistent -- one of their many roles is to consistently apply the rules so that we know what's coming. And the standard

of that is well-controlled, and that's the standard. Work with that.

But there is something else that I hadn't anticipated was going to come out of this, and this was this idea of sharing the other available data. Keeping in mind the limits on the label, we get the data published -- we should talk about the nuances of the trial. But there's something else that I think we could be looking for, and that is a sort of peer-to-peer communication. We're nominating IDSA here to the society for validating -- you can validate by publishing an informed critique of the available secondary data.

And I do mean this as peer-to-peer. It's 4 o'clock, I'm stuck with this patient, I might send my fellow to the library. He or she might find the right information. It's actually better if Helene spent some time six months ago reviewing it and she wrote down what she thought she would do in this unusual circumstance. It's the information that you would like to have.

And reviews like that can be used for a

number of things. They can be used by a company for discussion with payers. If written correctly, it's also the thing that when you want to use something off-label you can show to the insurer to say, shut up and just pay for it. That's what's in this document. Write it down in advance. And it might be usable for promotion. I'm not sure about that or not. We also need to include Europe in this conversation. So, this is all about telling the story.

The other part is Use the Data. Theme A: Be Clear on the Power of the Standard Indication. Modern non-inferiority sites are powerful tools. They do detect inferior agents, so it's an interesting presentation from Aaron about how few it takes to find the dog, you know? You're going to know pretty quickly if it's no good. They provide clear safety and efficacy comparisons, they facilitate initial approval, and they provide the basis, the launch pad to get you to the other stuff. And then better use of the data we already have.

A bunch of neat ideas here that are going to take some thinking. How do we borrow data

across indications? Could it be that we use different thresholds for different settings in terms of the data that we're willing to validate? And don't forget about the idea of (indiscernible) patient-oriented measures. And I'm going to add DOOR. It didn't get typed in here in time. But that's the idea that we should be working with.

And then, finally, generate the data more efficiently. I've heard some nice discussions about that. It's not a panacea. Platform trials show a potential to reduce cost and speed data generation. The idea that the sites would already have a contract in place. That right there saves a year in getting the trial actually running.

And the trial platform thing seems particularly true after initial approval is achieved. Studies in pediatrics, HAP/VAP, and rare infections would be especially suitable here. That was a theme that we heard when developing the concepts that you've heard about for the master protocols.

So, it's the last slide. Great conversation. And I want to give a shout out to

Sunita Shukla, who is sitting right there and was the ringmaster for making this meeting happen. Well done. Thank you very much for your help with that.

(Applause)

She had a bunch of people to chase down and she made it happen. Nothing is set in stone as to what's going to happen next but a subsequent debate seems needed on particularly the three topics listed here: How do we borrow data across indications? The idea of different thresholds for different settings -- what is adequate and well? And how do we include Europe in this conversation?

And I think if we start working -- those are things we can work on, in addition to the things the societies can work on, and I'm hopeful that we will have a next conversation. So, thank you to FDA, thank you to IDSA, thank you to Pew, thank you to the NIH for convening this session. Safe travels.

(Applause)

JOHN FARLEY: And I think I speak for all of us in the Department of Health and Human Services that have been part of this, that we have

gotten a lot of very useful input and ideas. And I also want to thank IDSA, and Pew, and NIH for co-sponsoring this event. We're committed to support some of the efforts that will come out of this meeting and we're also committed to reconvening this group at an appropriate time in the future because we think this was a really useful meeting. So, thank you.

CERTIFICATE OF NOTARY PUBLIC

I, MICHAEL FARKAS, the officer before whom the foregoing proceedings were taken, do hereby certify that any witness(es) in the foregoing proceedings, prior to testifying, were duly sworn; that the proceedings were recorded by me and thereafter reduced to typewriting by a qualified transcriptionist; that said digital audio recording of said proceedings are a true and accurate record to the best of my knowledge, skills, and ability; that I am neither counsel for, related to, nor employed by any of the parties to the action in which this was taken; and, further, that I am not a relative or employee of any counsel or attorney employed by the parties hereto, nor financially or otherwise interested in the outcome of this action.

MICHAEL FARKAS

Notary Public in and for the

STATE OF MARYLAND

CERTIFICATE OF TRANSCRIBER

I, SONYA LEDANSKI HYDE, do hereby certify that this transcript was prepared from the digital audio recording of the foregoing proceeding, that said transcript is a true and accurate record of the proceedings to the best of my knowledge, skills, and ability; that I am neither counsel for, related to, nor employed by any of the parties to the action in which this was taken; and, further, that I am not a relative or employee of any counsel or attorney employed by the parties hereto, nor financially or otherwise interested in the outcome of this action.

SONYA LEDANSKI HYDE

b

h