FDA CDER & NIH NCATS WORKSHOP

Regulatory Fitness in Rare Disease Clinical Trials

Virtual Workshop

Day 1

Monday, May 16, 2022

9:00 a.m. to 4:00 p.m.

Meeting Roster

Philip John (P.J.) Brooks, Ph.D.

Acting Director, Division of Rare Diseases Research Innovation
National Center for Advancing Translational Sciences (NCATS)
National Institutes of Health (NIH)

Katie Donohue, M.D., M.Sc.

Director, Division of Rare Diseases & Medical Genetics (DRDMG)
Office of Rare Diseases, Pediatrics, Urologic & Reproductive Medicine (ORPURM)
Office of New Drugs (OND)
Center for Drug Evaluation and Research (CDER)
U.S. Food and Drug Administration (FDA)

Sheila Farrell, M.D.

Medical Officer
DRDMG, ORPURM, OND, CDER, FDA

Raphaela T. Goldbach-Mansky, M.D., M.H.S.
Senior Investigator & Chief, Translational
Autoinflammatory Diseases Section
Laboratory of Clinical Immunology & Microbiology
(LCIM)
National Institute of Allergy & Infectious Diseases
(NIAID), NIH

Leslie B. Gordon, M.D., Ph.D.
Professor of Pediatrics Research
Warren Alpert Medical School of Brown University
and Hasbro Children's Hospital and
Medical Director and Co-Founder
The Progeria Research Foundation

Andrea L. Gropman, M.D.
Principal Investigator, Urea Cycle Disorders
Consortium
Rare Diseases Clinical Research Network (RDCRN) and
Professor and Division Chief
Neurodevelopmental Pediatrics and Neurogenetics
Children's National Hospital

Brendan H.L. Lee, M.D., Ph.D.
Principal Investigator, Brittle Bone Disorders Consortium, RDCRN and
Professor and Chair, Molecular and Human Genetics
Baylor College of Medicine

Kerry Jo Lee, M.D.
Associate Director for Rare Diseases
Rare Diseases Team, DRDMG, ORPURM, OND, CDER, FDA

Matthias Kretzler, M.D.
Principal Investigator, Nephrotic Syndrome Study Network (NEPTUNE), RDCRN and
Professor of Internal Medicine-Nephrology and Computational Medicine & Bioinformatics
University of Michigan Medical School

Janet Maynard, M.D., M.H.S.
Director, ORPURM
OND, CDER, FDA

Elizabeth A. Ottinger, Ph.D.
Deputy Director of Programs & Head of Project Management
Division of Preclinical Innovation, NCATS, NIH

Jennifer Rodriguez Pippins, M.D., M.P.H.
Clinical Advisor
Office of New Drug Policy, OND, CDER, FDA

Bita Shakoory, M.D.
Study Coordinator, Translational Autoinflammatory Diseases Section
LCIM, NIAID, NIH

Jeff Siegel, M.D.
Director, Office of Drug Evaluation Sciences
OND, CDER, FDA

Tiina K. Urv, Ph.D.
Program Director, Division of Rare Diseases Research Innovation
NCATS, NIH

Jie (Jack) Wang, Ph.D.

Clinical Pharmacology Team Leader

Division of Translational & Precision Medicine

Office of Clinical Pharmacology

Office of Translational Sciences (OTS)

CDER, FDA

Yan Wang, Ph.D.

Statistical Team Leader, Division of Biometrics IV

Office of Biostatistics, OTS, CDER, FDA

C O N T E N T S

C O N T E N T S (continued)

C O N T E N T S (continued)

C O N T E N T S (continued)

P R O C E E D I N G S

(9:00 a.m.)

Welcoming Remarks – Kerry Jo Lee

DR. K.J. LEE: Hello. My name is Dr. Kerry Jo Lee. I am the associate director for Rare Diseases in the Division of Rare Diseases and Medical Genetics, and the lead of the Rare Diseases Team at the Center for Drug Evaluation and Research, or CDER, here at the FDA. I am very excited to welcome you to our Regulatory Fitness in Rare Disease Clinical Trials Workshop, jointly presented by CDER and the National Center for Advancing Translational Sciences at the NIH.

CDER ensures that safe and effective drugs are available to improve the health of people in the United States and regulates over-the-counter and prescription drugs, including some biological therapeutics. We do not regulate gene therapies or vaccines. Those are in the Center for Biologics Evaluation and Research.

So why are we here today? There are over 7,000 rare diseases and conditions that

significantly impact patients and families. Despite an increase in novel rare disease approvals, there is still a tremendous unmet need for FDA-approved treatments for rare diseases and conditions. Rare disease drug development is complex; there can be limitations in our understanding of the natural history of a disease; challenges with endpoint selection; and the fact that small populations can also lead to challenges with trial design and interpretation.

All of us are here over the next day and a half to learn more about the fundamentals, best practices, and lessons learned when it comes to rare disease drug development that hopefully can help us in our work together to overcome these challenges.

This workshop focuses on academic investigators and those looking to learn how to bridge the gap between scientific discovery, academic investigation, and the regulatory aspects of drug development. Today's speakers from the FDA will explore topics such as adequate and well

controlled trials and core principles and fundamentals of trial design and interpretation, including analysis and dose ranging to maximize the effective use of small populations. You'll also hear from speakers in academia who will share their experiences.

As a reminder, this is not a forum to address specific questions about applications, but rather a forum to promote general understanding of the fundamental principles necessary to develop safe and effective therapies.

Some of you may have heard of CDER's new Accelerating Rare disease Cures program, or CDER's ARC program, whose mission is to drive scientific and regulatory innovation and engagement to accelerate the availability of treatments for patients with rare diseases. This event is an example of a type of engagement we really hope to support within the program, and we are so excited to be here to participate in what we hope will be just one of many future events.

And now I will turn it over to Dr. P.J.

Brooks, the acting director of the Division of Rare Diseases Research Innovation at the National Center for Advancing Translational Sciences to complete your welcome to the program today.

Dr. Brooks?

DR. BROOKS: Great. Thank you, Kerry Jo.

On behalf of NCATS and NIH, it's also my pleasure to welcome you to this meeting. As you know, at NCATS, our major focus is on translational science and improving the process of translation for all diseases, and a key aspect of that is understanding how to navigate the regulatory process.

So we were very pleased to have the opportunity to co-organize this meeting with our colleagues at the FDA, and very much look forward to the discussions, and clarification, and learning about the best ways to navigate the regulatory process.

So without further ado then, I would like to turn it over to Dr. Sheila Farrell from the Division of Rare Diseases and Medical Genetics in

the Office of New Drugs at FDA, who will be moderating the first session.

Sheila?

Session 1

Sheila Farrell - Moderator

DR. FARRELL: Thank you.

Good morning and welcome. I'm Dr. Sheila Farrell. I'm a medical officer in the Division of Rare Diseases and Medical Genetics at the Food and Drug Administration, and I'm the moderator for Session 1.

In this session, we have three speakers from the FDA Center for Drug Evaluation and Research who will be discussing different aspects of the approach to demonstrating substantial evidence of effectiveness for rare disease drug development. After all three speakers have given their presentations, we will have a question and answer period. Please submit your questions by clicking on the "Ask a Question" icon on the bottom right of the webcast player interface. We will try to get to as many of these questions as possible.

Now, without further ado, I'd like to introduce our first speaker. Dr. Janet Maynard is the director of the Office of Rare Diseases, Pediatrics, Urologic and Reproductive Medicine in the Office of New Drugs. The title of her presentation is the Approach to Demonstrating Substantial Evidence of Effectiveness for Rare Disease Drug Development: Overview Considerations.

Dr. Maynard?

Presentation – Janet Maynard

DR. MAYNARD: Thank you so much, Sheila.

Good morning. My name is Janet Maynard, and I'm the director of CDER's Office of Rare Diseases, Pediatrics, Urologic and Reproductive Medicine. In terms of my background, I'm a rheumatologist. Prior to joining FDA, I performed my fellowship, and then joined the faculty at Johns Hopkins Hospital, where I participated in research, patient care, and education. As a rheumatologist, I have helped care for patients with both common and rare diseases, which often have profound impacts on patients and families.

To tackle challenging public health issues, it is critical that we collaborate to advance public health for all patients. It is my pleasure to provide an overview of considerations related to demonstrating substantial evidence of effectiveness for rare disease drug development.

Next slide, please.

This is a standard disclaimer and disclosure slide. This presentation is not intended to convey official U.S. FDA policy, and all the materials presented are in the public domain.

Next slide, please.

Here is an outline for our discussion this morning. We will review FDA's regulatory framework; consider rare disease progress and challenges; discuss rare disease trial designs; and end with considerations related to innovation in drug development.

Next slide, please.

As background, the FDA's Center for Drug Evaluation and Research, or CDER, performs an essential public health task by making sure safe

and effective drugs are available to improve the health of people in the United States. An efficient predictable approval process is key to the development of innovative drugs.

Next slide, please.

It is important to consider the regulatory framework within which drugs are approved. To be approved for marketing, a drug must be safe and effective for its intended use. In terms of efficacy, there must be substantial evidence consisting of adequate and well-controlled investigations that the drug product will have the effect it purports or is represented to have under the proposed labeled conditions of use. A drug's effect must be clinically meaningful to patients.

In terms of safety, recognizing that all drugs have some ability to cause adverse effects, the safety of a drug is assessed by determining whether the benefits outweigh its risks. Safety is considered in relation to the condition treated, the efficacy purported, and the ability to mitigate the risk.

Next slide, please.

For product approval, data must support that the benefits of a product outweigh its risks. Benefits can be assessed by whether the product has a positive impact on how a patient feels, functions, or survives. Being able to describe clinical benefit is essential to making a decision about the favorability of the benefit-risk profile of a product.

Benefit-risk assessment considers the extensive evidence of safety and effectiveness submitted by a sponsor in an application, as well as other factors, including the nature and severity of the conditions the drug is intended to treat; the benefits and the risks of other therapies for the same condition; and any risk management tools that might be necessary.

Benefit-risk assessment in FDA's drug regulatory context is making an informed judgment as to whether the benefits, with their uncertainties of the drug, outweigh the risks with their uncertainties and approaches in managing the

risk under the conditions of use described in the approved product labeling.

Next slide, please.

Transitioning from a regulatory framework to rare disease considerations, we are seeing progress in rare disease drug development. Between 2015 and 2021, CDER approved 160 novel drugs for rare diseases, which was approximately 50 percent of all novel drugs that CDER approved. In addition, over 600 treatments for rare diseases have been FDA approved since the passage of the Orphan Drug Act. However, despite the significant progress, there is still significant work that needs to be done. Of the approximately 7,000 rare diseases, a vast majority lack an FDA-approved treatment.

Next slide, please.

This figure shows the progress in rare disease drug development over time. Specifically, this figure shows the number of novel drug approvals from 2010 to 2021. The columns are divided into the number of orphan novel approvals in green and the number of non-orphan novel

approvals in blue. The purple line indicates the percentage of orphan drug approval of all approvals in a specific year.

Since 2010, the number of orphan approvals has risen dramatically in the United States. In addition, the percentage of all approvals that are orphan approvals has also increased. In 2021, CDER continued to build on our previously successful years and approved 26 orphan novel drugs. That's 52 percent of all novel drug approvals by CDER in 2021.

In addition to novel approvals, every year CDER also approves additional uses for already FDA-approved drugs that help patients with rare diseases. These are called supplemental approvals. Our novel and supplemental approvals address a wide range of rare diseases that are often serious, and in some cases life-threatening.

Next slide, please.

Despite this progress, rare disease product development remains challenging. To help overcome these challenges, it is critical that we utilize

strategies and collaboration to facilitate optimal rare disease product development.

Next slide, please.

There are many challenges in rare disease product development. These challenges include small and sometimes very small patient populations. There can be genotypic and phenotypic heterogeneity within a disease. The natural history is, unfortunately, often poorly understood. These diseases are often serious and life-threatening and can be progressive with a childhood onset. There can be a reluctance at times to randomize to placebo.

In addition, sometimes we lack drug development tools, such as established efficacy endpoints. In addition, there may be limited, if any, regulatory precedent. It is important to incorporate regulatory flexibility while upholding our regulatory standards.

Next slide, please.

A key aspect of supporting approval is establishing substantial evidence of effectiveness.

This is defined as "evidence consisting of adequate and well-controlled investigations," including clinical investigations, by qualified experts by scientific training and experience to evaluate the effectiveness of the drug involved on the basis of which it could fairly and reasonably be concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or the proposed labeling.

Considerations related to substantial evidence of effectiveness will be covered in additional detail by Dr. Jennifer Pippins.

Next slide, please.

Substantial evidence of effectiveness is derived from adequate and well-controlled studies. These studies have the following characteristics. There is a clear statement of the objectives of the investigation and a summary of a proposed or actual method of analysis in the protocol for the study and in the report of its results.

The study uses a design that permits a valid

comparison with a control to provide a quantitative assessment of drug effect. There is adequate assurance that the subjects have the condition being studied. In addition, there are adequate measures that are taken to minimize bias on the part of the subject, observers, and analysts of the data, and assure comparability of treatment groups.

In addition, there are well-defined and reliable measures of assessing treatment response, and there's an analysis of results that is adequate to assess the effects of the drug.

Next slide, please.

The key aspect of today's workshop is to provide an overview of the fundamentals of drug development. Thus, we will first review frequently seen limitations or issues that we commonly encounter with rare disease trial design proposals, and then we'll consider strategies to address these.

Some common issues that we have seen include a non-randomized design when a randomized trial is feasible and ethical. In addition, we've seen

significant biases; for example, an external control or lack of blinding that cannot be adequately overcome in a specific drug development program.

Sometimes there's a limited understanding of the disease natural history to inform the trial design, including the study population, trial duration, and endpoints. Often, we see inadequate dose exploration, and sometimes a trial may be too short to detect a treatment effect, especially for slowly progressive diseases. If an endpoint is poorly chosen or a disease is very heterogeneous, sometimes we have to think creatively about endpoints to make sure that they are meaningfully assessing benefits.

Lastly, in some diseases that require dietary management, there can be limitations in the proposal if the diet is not optimized or standardized for those specific diseases.

Next slide, please.

These types of problems can lead to suboptimal inefficient trial design and biases. As

a result, the trial may fail to detect a treatment effect that exists or may show a treatment effect when there isn't one.

Next slide, please.

At this workshop, we will consider strategies to address some of these challenges. For example, it's important to understand the disease natural history as early and as comprehensively as possible. Also, it's important to utilize trial proposals that are designed to meet their stated objectives. We encourage frequent and early interaction with FDA and a specific review division that will be reviewing the protocol.

In addition, it's important to await FDA's review and comment before initiating a pivotal trial. Also, we should minimize uncertainties that we can control such as ensuring excellent trial conduct.

Next slide, please.

Rare disease stakeholders such as patients, families, and researchers can provide key elements

that can enable research and drug development for a rare disease. For example, stakeholders can help bring patients and families to engage with academic scientists. In addition, stakeholders can support the development of natural history studies and registries, which can provide both natural history data and facilitate the enrollment in potential future clinical trials.

This also facilitates engagement of other stakeholders such as industry and academia that may be interested in working in a specific disease area. In addition, stakeholders are very important in setting up patient-focused drug development or patient listening sessions, which can help develop greater clarity on what matters most to patients.

Next slide, please.

In terms of trial design, randomization and blinding are critical features for reducing bias. They should be the default approach when feasible and ethical. They are essential for detecting small but clinically meaningful effects. They are also very important for subjective or

effort-dependent endpoints.

It is important to note that there are trial design approaches that can minimize exposure to placebo; for example, utilizing dose response, delayed start, randomized withdrawal, or crossover designs. In addition, we have seen innovative proposals related to adaptive designs, master protocols, unequal randomization, and use of rescue criteria.

Next slide, please.

For proposals with non-randomized control, a major limitation is bias due to lack of randomization and blinding. Important questions include whether the treatment and control groups are comparable; if the endpoints are comparably assessed or impacted by lack of blinding; and is the control group comparable in terms of concomitant treatments, background standard of care, and endpoints available?

These should be considered when randomization is infeasible or unethical, also if the treatment effect is anticipated to be large,

and if the usual course of the disease is highly predictable.

Next slide, please.

FDA encourages innovative trial designs and creative thinking. Some examples include adaptive designs, master protocols, and novel approaches to endpoints. Regardless of the approach, prespecified analyses with type 1 error control are important to avoid data dredging and cherry picking.

Next slide, please.

The Food and Drug Administration is committed to facilitating the development of innovative, safe, and effective treatments and cures for patients who need them. I will discuss several select ways that FDA supports innovation in drug development, including patient-focused drug development; guidance documents; the Model-Informed Drug Development and Complex Innovative Trial Design Pilot programs; CDER's Rare Diseases Team; and CDER's Accelerating Rare disease Cures program.

It's important to remember that enhanced

flexibility and an efficient approval process have come while preserving our gold standard of safety and efficacy. At the end of the day, innovative therapies are only helpful to patients if they work and are demonstrated to be safe. So it is imperative that we ensure the right balance among patient access, sound science, and safe and effective products.

Next slide, please.

Establishing the therapeutic context is an important aspect of our benefit-risk assessments. Patients are uniquely positioned to inform our understanding of this context. PFDD, or patient-focused drug development, is a systematic approach to help ensure that patients' experiences, perspectives, needs and priorities are captured and meaningfully incorporated into drug development and evaluation.

PFDD efforts include FDA-led PFDD meetings; externally-led PFDD meetings; the PFDD Methodological Guidance Series; and the Clinical Outcomes Assessment or COA grant program. During

this workshop, you'll hear additional details regarding FDA's patient-focused drug development program.

Next slide, please.

Another mechanism to support innovation is through guidance documents that represent FDA's current thinking on a particular topic. These guidance documents are intended to provide guidance to different individuals depending on the content of the guidance. In the context of drug development, guidance is intended to assist drug developers in the development of drug products for the treatment of a specific disease or a type of disease, however, guidance documents are not roadmaps, as each development program has unique considerations.

Next slide, please.

FDA has issued several recent guidances that are relevant to the rare disease community. First, FDA issued a draft guidance for industry, entitled Real World Data: Assessing Registries to Support Regulatory Decision Making for Drugs and Biological

Products. This guidance was issued as part of the Real-World Evidence program and to satisfy, in part, the mandate under the federal Food, Drug, and Cosmetic Act to issue guidance about the use of real-world evidence, or RWE, in regulatory decision making.

This guidance provides sponsors and other stakeholders with considerations when either proposing to design a registry or using an existing registry to support regulatory decision making about a drug's effectiveness or safety.

In addition, FDA has taken steps aimed at advancing the development of individualized medicines to treat genetic diseases. Specifically, FDA has issued four draft guidances on topics related to individualized, investigational, antisense oligonucleotide or ASO drugs. These guidances cover topics related to clinical recommendations; chemistry, manufacturing, and control recommendations; administrative and procedural recommendations; and nonclinical testing.

Next slide, please.

In addition to guidance documents, FDA has other programs that are intended to facilitate drug development. For example, the Complex Innovative Design Pilot Meeting program is intended to support the goal of facilitating and advancing use of complex adaptive, Bayesian, and other novel clinical trial designs.

In addition, the Model-Informed Drug Development Pilot program is intended to facilitate the development and application of exposure-based biological and statistical models derived from preclinical and clinical data sources, referred to as MIDD approaches.

Next slide, please.

In addition to the innovative programs mentioned thus far, CDER has a Rare Diseases Team to help facilitate rare disease drug development. Established in PDUFA V, CDER's Rare Diseases Team facilitates, supports, and accelerates the development of drugs and therapeutic biologics for rare diseases.

The Rare Diseases Team is a multidisciplinary team located in the Division of Rare Diseases and Medical Genetics in the Office of Rare Diseases, Pediatrics, Urologic and Reproductive Medicine. Select activities include promoting advice to other review divisions on their rare disease programs; promoting rare disease consistency across CDER's Office of New Drugs, or OND; leading cross-cutting OND rare disease guidances, policies, strategic research, and workshops; developing rare disease training and education; and engaging with internal and external stakeholders.

Next slide, please.

As mentioned by Dr. Kerry Jo Lee at the beginning of this workshop, CDER recently announced the launch of the new Accelerating Rare disease Cures or ARC program. The vision of CDER's ARC program is speeding and increasing development of effective and safe treatment options, addressing the unmet needs of patients with rare diseases.

The mission of CDER's ARC program is to

drive scientific and regulatory innovation and engagement to accelerate the availability of treatments for patients with rare diseases. This is a CDER-wide effort with leadership represented from several offices throughout the center. The program is managed by CDER's Rare Diseases Team.

In its first year, CDER's ARC program will focus on strengthening internal and external partnerships with stakeholders and will engage with external experts to help identify solutions for the challenges in rare disease drug development.

Next slide, please.

In conclusion, the development of safe and effective drugs is central to FDA's mission. Rare disease development can be challenging, and it's essential to engage with FDA early and often during your drug development program. It's also important to learn as much as possible about your rare disease to optimize trial design. Also, you should ensure that your trials are adequate and well controlled.

Lastly, collaboration is key to facilitating

rare disease drug development. We are so appreciative for your participation in today's workshop and look forward to the discussion. Thank you very much.

DR. FARRELL: Thank you, Dr. Maynard, for that excellent overview.

Now, I would like to introduce our second speaker. Dr. Jennifer Rodriguez Pippins is a clinical advisor in the Office of New Drug Policy. The title of her presentation is Demonstrating Substantial Evidence of Effectiveness.

Dr. Pippins?

Presentation – Jennifer Rodriguez Pippins

DR. PIPPINS: Good morning, and thank you for that introduction. As mentioned, I'm a clinical advisor in the Office of New Drug policy, and my current work is focused on issues pertaining to evidence of effectiveness.

Prior to coming to FDA in 2009, I trained in internal medicine at Brigham and Women's Hospital in Boston, Massachusetts, as well as in pediatrics at Massachusetts General Hospital and Boston's

Children's Hospital, where I cared for a range of patients, including those with rare disease. I'm very excited to have this opportunity to be with you to talk about demonstrating substantial evidence of effectiveness.

Next slide.

Here is our standard disclaimer slide.

Next slide.

Stepping back for a moment, I want to provide some historical context. Between 1938 and 1962, drug manufacturers were only required by law to show that their drugs were safe. Over time, there was congressional concern about misleading and unsupported claims. Congress acted in 1962 with amendments to the federal Food, Drug, and Cosmetic Act, otherwise known as the Kefauver-Harris amendments, which included a provision requiring manufacturers to establish effectiveness with substantial evidence before approval.

Next slide.

The 1962 amendments to the federal Food,

Drug, and Cosmetic Act specified that one of the grounds for rejecting an NDA is a lack of substantial evidence that the drug will have the effect it purports to have. Additionally, FDA has also generally considered substantial evidence of effectiveness to be necessary to support licensure of BLA under the PHS Act.

Next slide.

The 1962 amendments also defined for the first time substantial evidence of effectiveness to be evidence consisting of adequate and well-controlled investigations, including clinical investigations by experts qualified by scientific training and experience to evaluate the effectiveness of the drug involved on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof.

Next slide.

Requiring evidence consisting of adequate and well-controlled investigations was significant because prior to 1962, it was not unusual for drug manufacturers to make claims about their products based on other types of data.

The requirement for generating evidence to adequate and well-controlled investigations was truly novel. Notably, the amendments specified investigations. The law's plural wording has generally been interpreted as indicating the need for at least two adequate and well-controlled trials, each convincing on its own, and is based on the scientific concept of providing independent substantiation of results.

Next slide.

Fast-forwarding to 1997, FDAMA amended the federal Food, Drug, and Cosmetic Act to allow for FDA to determine that a single positive adequate and well-controlled trial plus confirmatory evidence can establish substantial evidence of effectiveness.

I want to underscore that this mechanism to

establish substantial evidence of effectiveness may not always be appropriate. Since FDA needs to make a determination, based on relevant science, that a single trial and confirmatory evidence are sufficient, sponsors who are interested in establishing substantial evidence of effectiveness using this approach should seek feedback from FDA as early in development as is possible.

Next slide.

I previously touched on the scientific concept of providing independent substantiation in the setting of two adequate and well-controlled trials. In the one trial plus confirmatory evidence paradigm, it is the confirmatory evidence that provides substantiation of or support for the results of a single trial. It's also important to note that while FDAMA introduced the one trial plus confirmatory evidence approach to establishing substantial evidence of effectiveness, the act does not include a definition of confirmatory evidence.

Next slide.

The remainder of this presentation will

describe in greater detail these approaches to demonstrating substantial evidence of effectiveness. The content in the following slides, unless otherwise noted, is from an important document that I want to draw your attention to, the Draft 2019 Guidance titled, Demonstrating Substantial Evidence of Effectiveness for Human Drug and Biologic Products. I will refer to this publicly available document as the Draft 2019 Effectiveness Guidance.

Next slide.

This slide is the beginning of a figure that will serve as a visual summary of the Draft 2019 Effectiveness Guidance's approach to discussing substantial evidence in effectiveness. It depicts the two different approaches I've presented thus far, adequate and well-controlled clinical investigations, plural, seated on the left, and one adequate and well-controlled investigation plus confirmatory evidence, seated on the right.

Next slide.

First, I will focus on the left side of the

figure, the adequate and well- controlled clinical investigations approach.

Next slide.

The adequate and well-controlled clinical investigation scenario can consist of either two trials, as I've already described, or one large multicenter trial considered to be the scientific and functional equivalent of two trials, and I will describe these scenarios further on the next few slides.

Next slide.

In the scenario where there are at least two adequate and well-controlled trials, the second trial allows for independent substantiation of the results of the first. It's important to note that substantiation is not necessarily the same as replication; in fact, it's often more persuasive to have two trials that are not identical; for example, two trials, using somewhat different study populations within the same proposed indication or two trials for the same disease with different but related endpoints.

It's also worth noting that the designation of phase itself is not critical, and the distinction between phase 2 and phase 3 may not always be clear. Regardless of phase, however, the trials that contribute to our finding of substantial evidence of effectiveness must be adequate and well controlled, as further described in regulation.

Next slide.

In some cases, a single, large, multicenter trial can be considered sufficient on its own to establish substantial evidence of effectiveness. This is distinct from the scenario of a single trial plus confirmatory evidence, which I'll discuss momentarily.

The scenario of a single trial alone is not specifically described in statute. The Draft 2019 Effectiveness Guidance describes this scenario as a subset of the two adequate and well-controlled investigations approach, with the rationale that under certain circumstances there is no meaningful difference between the strength of evidence

provided by a single, large, multicenter trial and that provided by two smaller trials. Essentially, the large multicenter trials are considered both scientifically and legally to be multiple trials.

Next slide.

There are caveats to when such an approach ought to be acceptable, as outlined on this slide. The trial should demonstrate an effect that is clinically meaningful and statistically very persuasive on an endpoint such as mortality, a severe or irreversible morbidity, or prevention of disease with a potentially serious outcome.

A second trial might be impractical or unethical. Also, results are not driven by any single site; there are consistent effects across different endpoints and subgroups. Additionally, trial conduct must be thoroughly examined and found to be of high quality. It should be noted that negative findings from other trials could weaken the overall strength of the evidence and potentially might jeopardize such an approach.

Next slide.

Returning to our figure, I'll now focus on the right side --

Next slide.

-- and the one adequate and well controlled clinical investigation plus confirmatory evidence approach.

Next slide.

In some cases, FDA may determine that one adequate and well-controlled clinical investigation plus confirmatory evidence can demonstrate substantial evidence of effectiveness. As previously noted in this scenario, the confirmatory evidence, instead of a second adequate and well-controlled investigation, provides the substantiation of results from the single trial.

The Draft 2019 Effectiveness Guidance identifies factors FDA will consider when determining if such an approach is appropriate. These include such things as the persuasiveness of the single trial, the robustness of the confirmatory evidence, disease considerations, and whether it's ethical and/or practical to conduct a

second trial.

As I mentioned previously, sponsors considering such approach to demonstrate substantial evidence of effectiveness should discuss their intentions with FDA early on in development.

Next slide.

The Draft 2019 Effectiveness Guidance provides some examples of the types of data that may provide confirmatory evidence. These include clinical trial data for the drug in a closely related indication; mechanistic data; additional data from the natural history of disease; and scientific knowledge about the effectiveness of other drugs in the same class.

Next slide.

Having described approaches to demonstrate substantial evidence of effectiveness, I will end this presentation with a discussion of how FDA can exercise flexibility in this area. The Draft 2019 Guidance discusses this topic in some detail. Before presenting that content, however, I want to

turn back to statute and regulation for a moment.

First, statute. The statutory standard for substantial evidence of effectiveness includes an element of expert judgment. It says that experts, FDA, must make a conclusion about the data. FDA must make a determination that substantial evidence of effectiveness has been demonstrated.

Next slide.

The regulation that I'd like to direct you to is from the Code of Federal Regulations 314.105, which explains that the wide range of drug products and their indications requires FDA to exercise such judgment. It reads as follows:

"While the statutory standards apply to all drugs, the many kinds of drugs that are subject to the statutory standards and the wide range of uses for those drugs demand flexibility in applying the standards. Thus, FDA is required to exercise its scientific judgment to determine the kind and quantity of data and information an applicant is required to provide for a particular drug to meet statutory standards. FDA makes its views on drug

products and classes of drugs available through guidance documents, recommendations, and other statements of policy."

Next slide.

Turning back to the Draft 2019 Effectiveness Guidance, the final section of that document focuses on examples of situations when additional flexibility may be warranted. One way to exercise a judgment comes into place -- FDA's ability to fairly and responsibly rely on study designs that may produce less certainty in some circumstances, when appropriate. This reflects on the understanding that in some settings, less certainty about factors may be acceptable when balanced against the risk of rejecting or delaying marketing of an effective therapy.

FDA's decisions can take into account such circumstances as disease severity, disease rarity, extent of unmet need, and feasibility and ethical issues. However, while design and development program choices may result in greater or lesser degrees of certainty, in all cases, FDA must reach

the conclusion that there is substantial evidence of effectiveness. The statutory standard remains the same.

Next slide.

The Draft 2019 Effectiveness Guidance specifically addresses flexibility in the setting of life-threatening severely debilitating disease of unmet need and also in the setting of rare disease. The document discusses how flexibility can be incorporated in the approach to trial design, endpoints, number of trials, and statistical considerations.

Next slide.

In summary, today I've discussed that statute requires that substantial evidence of effectiveness be demonstrated. I've described different approaches to demonstrating substantial evidence of effectiveness: the adequate and well-controlled clinical investigations approach, which can consist of either two trials or one large, multicenter trial considered to be the scientific and functional equivalent of two trials,

as well as an approach, if determined to be appropriate, consisting of a single adequate and well-controlled clinical investigation plus confirmatory evidence.

I've also noted that statute and regulation both describe the role of flexibility, which is further described in the Draft 2019 Effectiveness Guidance. Flexibility may be particularly relevant in the setting of life-threatening severely debilitating disease of unmet need and rare disease. And with that, I'll end the presentation and turn it over.

DR. FARRELL: Thank you, Dr. Pippins, for that informative presentation.

Now, I would like to introduce our final speaker. Dr. Jeff Siegel is the director of the Office of Drug Evaluation Sciences in the Office of New Drugs, and the title of his presentation is the Role of Translational Science in Rare Disease Drug Development.

Dr. Siegel.

DR. SIEGEL: Good morning, everyone. Before

we start, I'd like to make sure everyone can see the slides and me, because when I first got on, I was unable to.

Please raise your hand if you cannot see the slides and me.

(No response.)

DR. SIEGEL: Okay. It looks like it was just me who was having that problem.

Presentation – Jeffrey Siegel

DR. SIEGEL: In any case, good morning, everyone. My name is Jeffrey Siegel. I'm the office director for the Office of Drug Evaluation and Sciences in the Office of New Drugs, in the Center for Drugs at FDA. I'm going to be speaking to you about the role of translational science in rare disease drug development.

Next slide.

Translational science really plays a key role in rare disease drug development -- I don't think that's a surprise to anyone -- and translational work, including biomarkers, unfortunately may not fulfill its potential in drug

development unless the discovery phase is followed by adequate analytical and clinical validation. Partnering with drug developers and consortia can allow translational science discoveries to fulfill their potential in drug development.

When I pause, you can advance the slides.

A resource, in case anyone is unaware of it, is the BEST resource. This is a site that contains explanations for the different types of biomarkers and how they're used in drug development.

Next slide.

Here's the list of different types of biomarkers. You've probably all seen these before. But I want to go through the implications this has for the work that you all do in promoting rare disease drug development.

Next slide.

Go back. Somehow the slides didn't work.

Okay. I just want to go through a couple of these and how important they are. Diagnostic biomarkers; in some situations, there may be a disease that has a common presentation, but there

are two fundamentally different genetic causes of it. In a case like that, having a diagnostic biomarker that distinguishes one type from another is really critical, and that would ordinarily be -- that should be part of the inclusion criteria for a clinical trial.

Next, prognostic biomarkers, these are obviously critically important. Imagine that you have a rare genetic disease that progresses slowly over time. It doesn't progress in six months; it doesn't progress in a year. It progresses in more like 3 years, 5 years, 10 Years. You can't necessarily rely on natural history studies to represent what's true now because there may be standard-of-care treatments that are actually effective but were never approved because there wasn't substantial evidence.

One of the reasons for this is that -- sorry. Someone sent me a text, and it's a little distracting.

Yes, one of the reasons for this is that in the old days, you would collect natural history

data based on patients who came to medical attention because of terrible, terrible consequences -- developmental delays and so on -- but now with genetic testing, we've learned that many of these diseases have a variable course. Some people may not present until they're adolescents. Some may progress when they're 2 years old.

So having prognostic biomarkers can allow you to match rare disease patients with the natural history controls, and that's something that really needs to be worked on more, but I think it's an important area for the future.

Monitoring biomarkers are measures of disease. They can be imaging biomarkers or panels of protein biomarkers. Lots of things can be considered for monitoring biomarkers, but they should measure something important about the disease and its progression.

Then you have pharmacodynamic and response biomarkers, including surrogate endpoints. These are pharmacodynamic biomarkers, so when you treat

with the drug, you can see an effect, and the effect reflects an impact on the target so that you can see what the drug is doing, hopefully rapidly, and then you can measure -- you can correlate the effect on the pharmacodynamic marker with long-term clinical outcomes, and that would represent a potential surrogate.

There are situations where things that you think would be good surrogates may not be because the substrate upstream of the missing enzyme may not necessarily have the effect of clearly being the metabolite that's responsible for the disease, so something to keep in mind.

Next slide.

When we think about using biomarkers in clinical development, we think about the type of biomarker it is, or prognostic, or enrichment, or whatever, and then how the biomarker impacts the clinical trial or drug development program. That's what's called the context of use. If it's to be used as a primary endpoint for approval of drugs for NPC, then that's how you would use it.

Next slide.

When we think about analyzing clinical trials using a biomarker, we think about the analytical validation and the clinical validation. The analytical validation has to do with whether the biomarker measures what it purports to measure, and whether it can be done with sensitivity and specificity, and is accurate and sensitive. Clinical validation, in contrast, has to do with the way the biomarker corresponds to a clinical outcome of interest.

Next slide.

Translational science could play a number of important roles in drug development programs. As Dr. Pippins has mentioned to you, one of the approaches for demonstrating substantial evidence of effectiveness, described in the Food, Drug, and Cosmetic Act, is with one adequate and well-controlled clinical investigation and confirmatory evidence.

When a drug's anticipated to be approved based on a single adequate and well-controlled

trial, there's a need for confirmatory evidence, and this confirmatory evidence can take many forms, some of which involve translational evidence.

I've shown in red the ones that involve translational evidence. There's clinical evidence from a related indication, which would not involve translational evidence. Mechanistic evidence could provide important support for a drug development program. Pharmacodynamic evidence in humans could provide important support. Evidence from a relevant animal model could provide important mechanistic evidence, assuming that the animal model is a phenocopy for the human disease.

Please advance my slides when I pause.

Biomarkers are integrated in drug development in a number of different ways. They can be incorporated as part of the drug approval process. Sometimes scientific community consensus is enough. Think of PTH levels for secondary hyperparathyroidism. Those trials never met Prentice criteria. That would be unnecessary because the mechanistic evidence was clear that

high PTH levels were the definition of the disease.

Then the other is through a program in my office, which is the Biomarker Qualification Program. With this program, once you're qualified, any drug development program can use the biomarker in their drug development program so long as it is the same context of use and the same type of biomarker in the validated assay.

Please advance my slides when I pause.

There are three interconnected paths to biomarker validation. One is through the Biomarker Qualification Program, like I just showed you; one is by scientific community consensus; and the other is, of course, through the drug approval process. Pharmaceutical companies, sponsors, can submit the biomarkers part of their program, and then it doesn't come to the Biomarker Qualification Program per se, but we get consulted, and then we would provide our input on the evidence for use of the drug in that particular drug development program.

Next slide. Thank you.

These are the different steps in the

Biomarker Qualification process.

Next slide.

I'd like to give you two examples of how biomarkers and translational science can be used in drug development programs. The first example is progeria. HGPS, as you all know, is extremely fatal, extremely rare, autosomal dominant segmental, and a premature aging disease.

Death is typically by heart failure at 15 years, but work from Francis Collins' lab and colleagues at other institutions identified lamin A as the responsible gene and demonstrated in animal models mutations in lamin A phenocopied HGPS, and the pathophysiologic pathway was determined to be persistent farnesylation of lamin A causing damage as cells age. Inhibitors of farnesylation ameliorate disease in animal models, lonafarnib, which is now approved for this dreadful disease.

Next slide.

I wanted to share how translational science contributes to developing effective therapy for HGPS. The first is genetic studies in humans

demonstrated the causal mechanism of HGPS, then the causal pathway was determined in animal studies to be this excessive farnesylation. The animal model recapitulated the human disease, making it really easy to test new drugs in the animal model to find out which ones were likely to work in humans. As you can see on the right in a study of mortality in progeria, this drug was shown to have a substantial effect on mortality in HGPS.

Next.

The next example I'd like to show you is AD-PKD. A consortium, put together by the C-Path Institute, the Critical Path Institute, related total kidney volume to progression of renal disease in autosomal dominant polycystic kidney disease. They developed a model shown here, where you could put in any set of baseline characteristics and show the rate of progression that was seen in patients with PKD. It's really quite remarkable because with any one set of parameters, you see very tight confidence intervals on what the progression rate is likely to be.

Next.

This model allowed us to determine, with quite a high level of precision, that total kidney volume was a prognostic biomarker for PKD. It was initially qualified as a prognostic biomarker based on modeling results, and subsequently it was applied in individual drug development programs. Data supported acceptance by the FDA review division as reasonably likely substantial evidence for accelerated approval.

Next.

I want to emphasize how important partnerships are. Partnerships can be a tremendous resource for bring together different stakeholders to qualify biomarkers in what would otherwise be a highly resource-intensive area. Academic groups may not have the funds or necessary data or samples to qualify biomarkers for regulatory decision making, but public-private partnerships like FNIH and the Critical Path Institute can play an important role in pulling these resources together and bringing together the different stakeholders to

be able to move the programs forward.

Public-private partnerships serve as intermediaries between patient groups, industry, academia, and regulators to develop novel drug development tools. There's a key role to collect trial data, share biosamples, integrate data sets, analyze and share data, and public workshops offer the opportunity for all stakeholders to share their views.

Biomarker developers may want to seek partnership with drug developers to assist in analytic validation, clinical validation, and incorporating the candidate biomarker in prospective clinical trials.

Next.

That's it. Thank you. I'm sorry there's not much time for questions.

Session 1 – Questions and Answers

DR. FARRELL: Thank you, Dr. Siegel, for that excellent presentation.

Now we will transition into the question and answer period for Session 1, as soon as all our

speakers are ready.

We've received a number of questions from the audience, which we appreciate. The first one, we received a number of questions regarding N of 1 trials and what our recommendations are in this space.

DR. MAYNARD: Great. This is Janet Maynard. I can start with that question.

We did receive several questions regarding N of 1, and I will say that there's been significant progress in the area of individualized medicine, where now with advances in technology, it really is possible to design drugs for an individual patient looking at their specific genetic defects. That of course raises very interesting regulatory considerations as we think about the normal drug approval pathway and how that might apply, where we're considering something that's being developed just for one patient.

This is a rapidly evolving space, and we are really committed at FDA to working with investigators. As I mentioned in my talk, we have

published four draft guidances from FDA that cover a range of topics, including clinical considerations, chemistry, manufacturing, and controls, nonclinical considerations, and even administrative considerations when you're working in this space.

So we hope that that is helpful. I think maybe one of the themes that we've had today during all of our different presentations is that it's so important to engage with FDA. So if you are working in this space and you have a specific question, please reach out to the relevant review division with your specific questions because we really want to work with you and address those questions as they arise during development.

I'll see if Jen or Jeff have anything they want to add to that.

DR. PIPPINS: Not to that. I think that covers it. I know we have many questions, so perhaps we'll move on to something else.

DR. FARRELL: Okay. Great.

The next question -- and we had a few of

these questions as well -- is when is a single-arm trial sufficient for the establishment of efficacy?

DR. PIPPINS: I'll take that. That's a great question, and it gets to the heart of so many different issues and raises a number of different considerations and topics, so you'll bear with me if I'm a little wordy.

The first topic the question raises is whether a single-arm trial can be considered an adequate and well-controlled investigation suitable for demonstrating substantial evidence of effectiveness, and as noted in the presentations, clinical investigations intended to demonstrate substantial evidence of effectiveness must be adequate and well controlled.

In the description of an adequate and well-controlled investigation, the CFR, the Code of Federal Regulations, states that the purpose of an adequate and well-controlled investigation is to distinguish a drug's effect from other influences. One of the key features of an adequate and well-controlled trial that allows it to accomplish

this goal is the use of controls, and the regulations describe a number of different controls.

By definition, a single-arm trial doesn't have a concurrent control group, and by concurrent, I mean a control within the same trial. But what's important to realize is that a single-arm trial can still be controlled, and that can happen in a variety of ways.

For example, as discussed already in these presentations, there can be an external control such as that drawn from a natural history study or from a placebo group from another trial. Alternatively, the control could be acknowledged external to the trial; for example, enough might be known about the disease that it could be concluded that the changes observed in the trial reflect the effect of the drug.

The classic example of this, everyone knows, is drawn from oncology, where tumors aren't expected to spontaneously shrink. So if tumors are observed to regress in a single-arm trial, there's

a basis on which to conclude that this represents an effect of the drug.

Now notably, as I already mentioned, there are many considerations to keep in mind when assessing whether or not an external control group, or a control group based on external knowledge outside of the trial, is appropriate, but certainly it's possible for a single-arm trial with an appropriate control to be adequate and well controlled, and therefore able to provide substantial evidence of effectiveness.

Now, I noted that the question raises a number of issues. Whether or not that single-arm trial -- that one single-arm trial -- is sufficient on its own to demonstrate substantial evidence in effectiveness, that's another issue, and that really speaks to everything that my presentation talked about in terms of the different approaches to establishing substantial evidence of effectiveness.

So I'll stop there and see if anyone else has anything to add to that.

DR. MAYNARD: Yes, that was really helpful, Jen.

I'll just add, to emphasize what Jen had discussed, this is really in a situation where we understand the natural history of the disease very well, and we have a good understanding of what would be expected in that disease.

We frequently have patient groups and other advocacy groups who come to us and say, "What can I do? I really want to help rare disease product development. What can I do?" And sometimes having information from a very robust natural history study can be helpful, not only in the setting of external controls, but also really to have a better understanding of the disease and the anticipated effect that it will have on patients, which really plays a critical role as we're thinking about the overall development program for a specific rare disease.

Some of the other questions we received in the meeting registration, in the context of external controls or single-arm trials was, how do

I get FDA's agreement on this when we're thinking about this in the development program? As I mentioned earlier, those are really conversations that should be happening with the review division, so as you're designing a potential study, really engage in those conversations.

FDA does have meetings around product development, where we meet with either folks from academic or sponsors to understand different questions that come up during development. So it's really important to have those conversations and think about the different considerations that Jen raised in the context of that specific development program for that specific disease.

DR. FARRELL: Okay. Thank you.

The next question is about biomarkers, and we got a few of these.

What are the most important questions the FDA is looking for when investigators are considering a novel biomarker as a primary endpoint to demonstrate efficacy through the accelerated approval pathway in these orphan drug indications?

Even before that, what should they be contemplating when they're thinking about novel biomarkers for a rare disease?

DR. SIEGEL: So let's imagine a couple contexts. We'll start with an easy one, and then turn next to a more difficult one.

An easy one is a genetic disease where there's a particular enzyme missing, and there's an upstream substrate that can be demonstrated to cause the toxicity. And if you don't have that increased level of the substrate, you don't have toxicity. That's the straightforward and easy way that you can incorporate biomarkers for regulatory decision making

In contrast, if you have a more complicated situation where in the animal model it, for instance, doesn't phenocopy the human disease, that makes it much more difficult, and if you have a biomarker where you can't be -- okay, let's imagine this situation.

You have a missing enzyme. You give a drug treatment in the animal model, and it turns off one

substrate but not necessarily another one, and you don't know for sure that the particular substrate that comes down is actually the one that's responsible for the disease. So making sure that they correspond is a really important aspect of what we do.

I think that's probably the main aspect of what I want to cover here. Thank you.

DR. FARRELL: Okay. Great. Thank you.

The next question is, what is the criterion to define a rare disease, and is this the same as an orphan disease?

DR. MAYNARD: In the United States, the Orphan Drug Act defines a rare disease as a disease or condition that affects less than 200,000 people in the United States. That's generally what we mean when we're referring to a rare disease, as defined in the Orphan Drug Act.

DR. FARRELL: Okay. Thank you.

The next question is about basket trials. Are basket trials an acceptable way of featuring clinical trials for rare diseases in non-oncology

indications with shared molecular ideologies?

DR. PIPPINS: I can take that, and just to make sure everyone's on the same page, it might be worth just reviewing a couple of definitions.

A master protocol is defined as one overarching protocol, and the key here is that it's designed to answer multiple questions. There are different kinds of master protocols, and this particular question is about the type known as a basket trial. Basket trials are designed to test a single investigational drug in the context of multiple diseases or disease subtypes, typically conditions that are related, such as the question mentioned with similar molecular ideologies.

The short answer is there are definitely ways in which basket trials could certainly play a role in drug development for rare disease. They're particularly attractive because master protocols, in general, in basket trials may offer certain types of efficiencies in terms of clinical drug development.

As will be discussed today and throughout

the entire workshop, there are various constraints and limitations that are created, or barriers are created in this setting of rare disease, given just the particular issues of having diseases with such low prevalence. So to have a tool like a basket trial that might provide certain efficiencies with testing different diseases or disease subtypes within a single protocol certainly is attractive, and there may very well be a role for it in drug development.

I want to point people to a couple resources. There's a really helpful, just general, opinion piece in the New England Journal back in 2017 by Dr. Woodcock and Dr. LaVange, which provides an FDA perspective on master protocols.

Then as alluded to in the question, most experiences that we've had with master protocols, or at least with basket trials, is in the setting of oncology. So while it doesn't directly speak to our topic today in terms of rare disease broadly, there are principles in a guidance put out and actually recently finalized by oncology about

master protocols that could certainly be useful. So I would point people to those resources.

DR. FARRELL: Thank you.

We've got a number of questions about real-world evidence. The first one -- and we might just ask for some comments from everybody if everybody has any -- is how can real-world evidence be used for confirmatory evidence for accelerated approval?

DR. PIPPINS: So I can start off with that one as well.

Fit-for-purpose, real-world data has the potential to generate real-world evidence that can be used to support a number of different regulatory type decisions. I'm actually going to answer this question for regulatory decisions more broadly and not just in the context of accelerated approval because it's relevant beyond just accelerated approval.

Again, just to make sure everyone has the same information, it's helpful to define a couple of terms. Real-world data is data relating to

patient health status and/or the delivery of healthcare routinely collected from a variety of sources. Real-world evidence is the clinical evidence regarding the usage and potential benefit to risks of a medical treatment that's derived from the analysis of real-world data. So you start with raw data, analyze it, and you can generate real-world evidence.

Real-world data can be used in different study designs to analyze it, so anything from randomized trials, including large single trials. It could be used as an external control arm in a single-arm trial. It could be used in observational studies. So there are different ways of using real-world data.

FDA has a robust real-world evidence program that includes guidance development, demonstration projects, and external engagement, all exploring the use of RWD and RWE in regulatory decision making. In addition, you will recall in my presentation, I referred to the Draft 2019 Effectiveness Guidance. That guidance does comment

and describes RWE as a possible source of confirmatory evidence.

DR. MAYNARD: Great. And maybe I'll just add a little bit on to what Jen is saying.

I think something that's really important when we're thinking about rare disease product development is really keeping the end in mind. Our goal is to have safe and effective drugs approved for patients and families living with rare diseases, and as we've sort of alluded to today, there are lots of different considerations in rare disease product development, and there are lots of different ways you can try and establish substantial evidence of effectiveness.

I think it's important, though, to keep the end in mind and consider how the different pieces of a development program will support that overall assessment of whether or not the drug is safe and effective for its intended use.

The questions are great because they've really alluded to a lot of the different creative thinking that we are seeing in rare disease product

development, and generally it's not a one size fits all. Each development program will have different considerations, whether that means related to real-world data and real-world evidence or the questions we were getting about N of 1 or basket trials. It's really important that we, of course, learn from other areas in rare disease product development but of course focus on the specific questions related to that development program.

Jeff, I don't know if you had anything else you'd like to add.

DR. SIEGEL: No. Thanks very much. I think you covered it very well.

DR. FARRELL: Okay. Great. Thank you.

We've received a question regarding global rare disease drug development and how we are working with our international counterparts.

Anybody would like to comment on that?

DR. MAYNARD: Sure. I can take that.

Rare diseases are, of course, inherently rare, and many of them affect patients globally, so

it's so incredibly important, especially for rare diseases, that we work with our international partners.

The Rare Diseases Team in the Division of Rare Diseases and Medical Genetics and the European Medicines Agency, or EMA, co-lead the international rare diseases cluster meeting, which is a confidential forum in which FDA, EMA, and other regulatory agencies convene to facilitate the exchange of information, including the scientific advice regarding rare disease drug development programs.

This is one example of communicating with our international colleagues because it's clearly important in drug development, in general, but especially for our rare disease drug development programs, where it's so important that we think about the considerations with our international partners.

DR. FARRELL: Okay. Great.

The next question is, what's the typical path for biomarker qualification using an IND, and

can this be shortened for a rare disease?

DR. SIEGEL: The short answer is yes, absolutely. The way this would be done is there's a pharmaceutical company sponsor who has an idea of a drug using perhaps an animal model, with evidence that the drug will effectively shut off the disease in the animal model. Let's just imagine that scenario.

The IND holder perhaps would be the pharmaceutical company sponsor, or if there was enough infrastructure to support this, it could be the clinical investigator themselves. And I think we need to work on developing that infrastructure because it's not available yet at many prominent institutions, and it would be easy to implement it.

The typical path would be that they would have the evidence demonstrated clearly that their drug will, in fact, turn off the disease process in the animal model, and then they submit their IND showing that it in fact does that and what their plan is for the first trial of safety, and then what their future plans are for testing the drug in

patients to demonstrate effectiveness.

As I mentioned before, this is not as easy as you might think because often the rate of progression with current standard care is completely different than it was in the past. You need to have prognostic biomarkers with very little interpatient variability.

DR. FARRELL: Okay. Thank you.

The next question is also on biomarkers. Could you address the options for developing clinical trials for rare diseases that progress slowly? And they're using an example of aberrant deposition of proteins that interfere with functioning but accumulate over seven years.

Could measurement of levels of the defective protein showing reduction act as a surrogate endpoint without the need to show prevention of disease?

DR. SIEGEL: Can you repeat the last part?

DR. FARRELL: They're asking if the measurement of the levels of the defective protein showing reduction could act as a surrogate endpoint

without the need to show prevention of disease manifestations.

DR. SIEGEL: Yes, absolutely. A lot of these diseases are slowly progressive as we talked about, so it may be difficult in the time frame of a clinical trial to see any clinical difference between treated patients and controls. It's a problem that we see often.

What you want to do instead is to provide evidence that the levels of the protein correspond in a prognostic way to clinical outcomes. And when you show that, then it can be seen as a surrogate endpoint, and then you can do a clinical trial, which potentially would be a single-arm study.

That's all something that would be negotiated between the pharmaceutical company sponsor and the review division. But if it's accepted as a surrogate endpoint, then that would be the basis for an approval for the drug with an adequate clinical trial.

DR. FARRELL: Okay. Thank you.

The next question is, could you please share

some insight on how a historical external control can make up for lack of randomization in the case of rare diseases?

DR. MAYNARD: Sorry, Sheila. I briefly lost audio. Would you mind repeating the question?

DR. FARRELL: Sure. Can you please share some insight on how a historical external control can make up for a lack of randomization in the case of rare diseases?

DR. MAYNARD: Yes. I think maybe, as Dr. Pippins mentioned earlier, when we're considering different trial designs, if we're considering using something like a historical control or some sort of external control, we need to think about the setting in which it's being used.

So generally, if we were using an external control, we would want to use it in a situation where the natural history of the disease is very well-defined. Also, the external control group would have to be very similar to the treatment group within the study, and then we'd have to make

sure that the treatments that were used in an external control are similar to what's being used in the study itself. In addition, often this is a situation where we would need to have very compelling evidence of an effect just so that we can make sure that it was not due to chance alone.

I'm not sure if I'm addressing it.

Jen, was there anything else you wanted to add or that I missed as I was trying to address the question, to make sure we got it?

DR. PIPPINS: No, just to say it's obvious, but it may be worth repeating, that the whole point of this is that we're trying to limit bias, so we're trying to really be able to discern that the effect that's observed is indeed an effect of the drug. So that's why you want these groups to be as comparable as possible.

DR. SIEGEL: Let me comment as well. Diseases like NPC progress very slowly, as we mentioned, so it may be very difficult to see an effect of the drug in the time frame of a clinical trial, but let's take a disease like methylmalonic

academia. There are investigators at the NIH who've been studying methylmalonic acidemia, and they have an amazing biomarker that seems to correlate with clinical outcomes in a very clear way, in a way that the substrate upstream of the missing enzyme does not, which is really remarkable, but that's their finding.

So in that case, the biomarker would be used as a surrogate endpoint, and it would be easy to show that this is what patients do currently and this is what patients do on this drug that effectively treats methylmalonic academia; very straightforward like that.

DR. FARRELL: Okay. Terrific. Thank you.

We've got a number of questions kind of asking a little bit more information on what specific examples of confirmatory evidence might be. Would anybody like to try to delve into that a little deeper?

DR. PIPPINS: Sure. I believe in one of my slides I talked about site examples, including four examples that are described in the 2019 Draft

Effectiveness Guidance. But among the various types of confirmatory evidence, there can be evidence from a clinical investigation conducted not for that specific disease but a closely related disease, where that information can be relevant and help to substantiate the results of a single trial.

Jeff touched on this somewhat, and in some ways the most examples we have today are confirmatory evidence drawn from information about the mechanism of the drug and/or pharmacodynamic effects of the drug that certainly can serve.

Additionally, we've discussed how RWE could potentially serve as confirmatory evidence, and then also information drawn about the natural history of disease. I want to note that, in that case, it's important -- the whole role or purpose of CE, or confirmatory evidence, is to provide substantiation of results, so if we're talking about natural history disease, information to serve as confirmatory evidence, we're not talking about information that's being used as, say for example, an external control for that single trial, but

rather we're talking about additional information that might provide additional confirmation of what's observed in, say, a control group for a trial; the concept being that if you're doing substantiation, you don't want something that's trying to substantiate itself. You want something external to the single trial in order to provide that substantiation.

So those are some examples, but I'll note that the 2019 guidance that talks about those four categories, those are examples. It's not intended to be an exhaustive list of the types of confirmatory evidence that are possible. So it's super important that sponsors engage the agency with regard to what they are thinking about.

DR. MAYNARD: Just to add a little bit on to what Jen is saying, the questions we received in the meeting registration, there was a lot of interesting examples, so I just wanted to make sure folks were aware of the resource we have.

Drugs at FDA is a website, which if you just Google Drugs at FDA, that's the easiest way to find

it, and it includes information, including reviews of approved drugs, and also includes the labeling information.

That can be a great resource if you want to look at different examples to see how FDA has articulated the review of specific applications, and that could be helpful as you're thinking about these questions about what is exactly substantial evidence of effectiveness or what are some examples of confirmatory evidence.

So I just wanted to make sure that folks were aware of that resource, and it can also be helpful looking for the most updated version of the labeling and things like that.

DR. FARRELL: Okay. We've got a couple questions on single trials. This question is asking, if we could provide some examples of rare disease drugs, non-oncology, that obtained approval on the basis of a single trial with confirmatory evidence, what is the process to communicate or get agreement with the FDA regarding use of one adequate and well-controlled trial with

confirmatory evidence?

Can the FDA provide a determination that one trial is adequate and well controlled during the IND stage, and if so, what kind of information would they need to provide to make this request?

It's a lot in that question.

DR. PIPPINS: This is a great question, and you're right, it packs a lot. It packs a lot in there. I can start off with some comments about process.

This is super important, but sponsors considering a development program consisting of one adequate and well-controlled trial plus confirmed evidence should engage as early as possible with FDA. There are a variety of venues for engagement with the agency during development, including milestone meetings such as even before the IND stage at the pre-IND meeting or end of phase 2 certainly. At these moments of engagement, a major central topic of discussion should be the anticipated approach to demonstrating substantial evidence of effectiveness.

Of course, whether or not the data generated by a development program, whether or not they're sufficient for approval, will ultimately depend on the results themselves. But certainly review divisions and sponsors can engage over the question of whether a single trial for CE approach appears to be reasonable.

The type of information -- which I think this is a great part of the question -- that should be provided to allow for such a discussion will include, at a minimum, the design of the single trial; what's anticipated about how persuasive its results might be; and information about the types and quantity of confirmatory evidence that are anticipated to be able to substantiate the single trial.

In terms of the nature of the discussion, the agency's ability to comment on the adequacy of the proposed approach is going to vary on the availability of the data from the program at the time of discussion. These may be somewhat iterative discussions in terms of from the pre-IND

stage to later on in development.

I don't know if others have additional things to add.

DR. MAYNARD: Yes, Jen. I completely agree with you. I think these discussions generally happen throughout development, but especially at the pre-NDA, or new drug application, or pre-BLA biologic license application meeting because at that meeting, really, when FDA and the sponsor can sit down and talk about the sponsor's anticipating submitting in their application. Generally, that would include consideration of the different trial, if it's one single trial, and what the confirmatory evidence would be.

Just to emphasize what Jen mentioned, at that meeting, FDA will have not reviewed the full details that are available from that because the sponsor would not have submitted the full details yet. But there can be an understanding and a discussion about what the anticipated scope of the development program is and how the sponsor is planning to support substantial evidence of

effectiveness.

That's the time, really, during development when those conversations are happening. And generally there is discussion and consideration of different proposals, and what are the potential strengths and weaknesses of the different proposals, and how those might be addressed when the application is submitted to FDA for review.

DR. FARRELL: Okay. Thank you.

In rare diseases, surrogate biomarkers can be predictive but not in all cases, especially in a heterogeneous disease population. Does regulatory flexibility apply here when not all patients see a benefit, despite showing a reduction in a surrogate biomarker?

DR. SIEGEL: I think that's actually an easy one. If you look at a particular disease, there can be different biologic subtypes that have different clinical courses, but within each one there would be, presumably, a similar course to the disease. And you would need to show that you've identified the key factors that determine when a

patient will progress or won't progress well on the treatment.

So in a situation like that, you would -- I think that's all I'm going to say. Thank you.

DR. FARRELL: Okay. Thank you.

Can you describe the difference between timing and development of a biomarker qualification in a surrogate endpoint for discussions with the FDA?

DR. SIEGEL: The way the Biomarker Qualification Program works is that there are three separate stages. First, the submitter submits a letter of intent, and we can have discussions in a pre-LOI, or pre-letter letter of intent, phase where we meet with the submitter and discuss what they would need to show to demonstrate that the product is an effective surrogate.

Once we're started on that, then the letter of intent would be accepted, and then we would go on to the next phase, which is the QP, the qualification plan stage, where the plan for analyzing the data and what data would be submitted

is submitted, and then we have opportunity to ask questions about it to make information requests to the submitter, who will provide explanations of their rationale for what they're doing.

Then based on that, once the qualification plan is accepted, we would proceed with the submitter putting the data together to support the drug being a surrogate endpoint. And at the end, they would submit a full qualification package where they would pull all the data together with the analyses that they said they were going to do in the qualification plan stage. Then we would look at the program, and if the data are supportive, we would accept the full qualification package and qualify the biomarker for the context of use, primary endpoint, as a surrogate or prognostic endpoint, whatever the appropriate context would be.

DR. FARRELL: Great. Thank you.

This is a question about getting FDA's input. This person is asking about the FDA feedback for rare and ultra-rare disease programs

if they've been working on fairly standard approaches but would like to reach out to individuals at the FDA to help navigate more novel approaches, and does anybody have any advice for that.

DR. MAYNARD: It's not fully clear to me from the question if it's in the context of a specific drug development program or if it's questions more in general. If it's a specific question about a drug development program like under an IND, then the best mechanism would most likely be to work with the review division. If it's a broader question, there are other forums which we can discuss more general topics, potentially something like a CPIM meeting, which we can discuss more general considerations related to facilitating drug development.

So it depends a little bit on the context of exactly the question, and that would be helpful to get an answer to. If it's specific, as I mentioned to a specific application, then I would interact with the review division, and more general, then

you could consider other mechanisms.

I don't know, Jeff and Jen, if there's anything else you wanted to add to that.

DR. SIEGEL: I'm good.

DR. FARRELL: Okay.

A number of questions on the difference between the different divisions; there's the rare disease group, and then there are divisions throughout the OND that deal with rare diseases but aren't actually the rare disease group. And they're just wondering about when they submit things to those divisions, are there other experts in those divisions, or what kind of expertise the divisions that aren't specifically rare diseases have at their disposal to help work through these programs.

DR. MAYNARD: Yes, that's a great question. The Rare Diseases Team, which I mentioned, is located within my office, the Office of Rare Diseases, Pediatrics, Urologic and Reproductive Medicine, and they help think about rare disease issues more broadly. But a specific application

that would potentially be for a rare disease would be within the review division with subject matter expertise. For example, if it was a rare rheumatic disease, that would be reviewed in the division that considers rheumatology considerations.

The Rare Diseases Team, though, is available to provide consultative service if there are any questions related to rare disease product development. We recognize with this significant increase that we've had in terms of rare disease product development, that rare disease considerations really now affect the Office of New Drugs and, really, CDER very broadly. So part of our efforts have been making sure we have resources so we can support the reviewers in all the different review divisions, who are specifically looking at those applications, by sharing knowledge and science about rare disease considerations.

So to answer the question, there is both a broad rare diseases team that helps answer cross-cutting rare disease issues, and then also specific input that would be provided from that

specific review division related to the application.

DR. FARRELL: Thank you.

Unfortunately --

DR. SIEGEL: I'd like to --

DR. FARRELL: I'm sorry. Go ahead.

DR. SIEGEL: I'd like to comment also.

This is a really interesting and important question. In the old days, it was very hard to find pharmaceutical company sponsors who are interested in developing drugs for rare diseases. That's completely not the case anymore. It's very viable financially for companies to develop drugs for patients who have a particular disease without any difficulty. These companies will partner with patient advocacy groups and get the support from that, and they know that if they have a successful drug, that it can be used to treat patients and demonstrate effectiveness.

So what I'm saying is that if you feel that you have an effective biomarker that is a surrogate, you should reach out to pharmaceutical

company sponsors and find companies who are interested, and discuss with the different ones, and find a company that you think will effectively promote development of a drug based on your defined biomarker pathway.

So just as I mentioned before, we recommend partnering with the Critical Path Institute and with the FNIH as public-private partnerships. Similarly, we recommend, when appropriate and at the right time, that biomarker developers should reach out to pharmaceutical company sponsors so that they can get the support for the analytical validation they might need, and in some cases clinical validation as well.

DR. FARRELL: Unfortunately, we have come to the conclusion of the time allotted for Session 1. We have so many great questions, including a lot of really great questions on trial design, which will be addressed in Session 3. So we're sorry we weren't able to get to all your questions, but we do encourage you to go to the other sessions, including Session 3, so maybe your questions will

get answered there.

I would like to thank all of our speakers for the excellent presentations and all the wonderful audience participation. We will now have a break, and we will reconvene at 10:45 for Session 2. Thank you.

(Whereupon, at 10:30 a.m., a recess was taken.)

Session 2

Elizabeth Ottinger - Moderator

DR. OTTINGER: My name is Elizabeth Ottinger, and welcome to Session 2. I am part of the therapeutics development branch at NCATS, where our program focuses on preclinical development for rare diseases and improving the translational processes to support the initiation of clinical trials.

In the first session, we heard from the FDA on substantial evidence of effectiveness needed to support drug approval for rare diseases, and in this session, we'll have three case studies from academic investigators who will share their

experience in rare disease clinical trials of diseases of very low prevalence. They'll discuss both their challenges along the way, but also successes to be able to show that a drug is safe and effective.

We have three talks followed by the question and answers, so please make sure you ask your questions in the right-hand corner button so that we can have that after the three talks.

I'd like to introduce our first speaker who is Dr. Leslie Gordon. She's the professor of pediatrics research for the Warren Alpert Medical School of Brown University. She's a professor at Department of Pediatrics for Hasbro Children's Hospital; a research associate, Department of Anesthesia at Boston Children's Hospital and Harvard Medical School; and she's director and co-founder of The Progeria Research Foundation, and she'll be sharing her story on the approval of lonafarnib for progeria.

Welcome, Dr. Gordon.

DR. GORDON: Thank you very, very much, and

thank you for asking me to speak today.

Are my slides going to be put up? I have just Dr. Ottinger's view.

(Pause.)

FEMALE VOICE: Hi, Dr. Gordon. Your slides are up. I can see them.

DR. GORDON: Oh, okay. That's interesting. I cannot see my slides.

(Pause.)

Presentation – Leslie Gordon

DR. GORDON: Well, thank you very much, again, for asking me to speak. This is an incredibly important meeting, and I'm really honored to be able to tell my story.

Next slide, please.

This is just disclosures.

Next slide, please.

These are some of the children with progeria, the children we are trying to save through our efforts in drug development. I've been asked to come sort of as a case study here to tell you what we went through in the story of lonafarnib

approval, now called Zokinvy, and it's a 20-year study in 15 minutes, so I'm going to try to streamline. But there's a lot I'll be skipping over, and a lot of efforts, and trials, and tribulations, and I'll be hitting the high points.

Next slide, please.

This is the picture of my family, and Sam you see here. Sam was born, and at the age of 2 was diagnosed with progeria. It's an ultra-ultra rare disease, and I'm sure you've heard this story so many times, rare diseases that are so rare that nobody knows anything about them, essentially, and that there's no place to go, and we didn't know if it even was a genetic disease.

So families do these things. They start foundations. We started The Progeria Research Foundation in 1999 to find cause, treatment, and cure for children with progeria all around the world.

Next slide, please.

Now, I'm just going to focus on just a couple of things here. Progeria has a prevalence

of 1 in 20 million, so there are about, today, maybe 400 kids with progeria throughout the world; very, very rare. The children all die of heart disease. The atherosclerosis that usually hits you and me in our 60s and 70s, hits them before the age of 10, and they die in their teens. This child on the right here, you see her born, but you see her on the right, and she's only 10 years old.

Next.

I'm showing you this because this is the set of foundational programs that we've built over time at The Progeria Research Foundation. One of the things that I'd like to point out that's most important here is that we have a registry program, and a medical and research database program, and a cell and tissue bank; all of the things that actually continue to be incredibly important in not only starting things off but continuing to succeed. We've talked a little bit here today about registries, and outcome measures, and natural history studies, and these things are incredibly important and have been in this story.

I see there's a little instruction here. I'm going to pause for a moment.

(Pause.)

DR. GORDON: These are the foundational programs, and I just wanted to point two of those out, and we'll be revisiting those later on as well.

Next slide, please.

Alright. We started in 1999. We supported some basic research, but we really wanted to discover the gene mutation for progeria and collaborated with some wonderful labs, including that of Francis Collins who discovered the gene mutation for progeria that was published in 2003, and really, we were catapulted into a new phase. That broke us open because now we could try to understand spring boarding from the biology of disease and identify treatments based on that biology of disease.

Next slide, please.

Progeria is caused by a gene mutation in lamin A, which produces a protein called lamin A,

and that protein is an internuclear membrane protein that has both structural and cell signaling effects. What happens in progeria is that there's a single-based mutation in 90 percent of the kids, and that mutation leads to the production of a shortened abnormal lamin A protein called progerin.

Next slide, please.

The key to element of progerin that I'm going to focus on today with lonafarnib is that that lamin A and, hence, progerin, goes through four post-translational processing steps, and you see that here on the left. On the right, you see progerin.

Now, with progerin, the omission of 50 amino acids creates a problem, and that problem is that the first step that lamin A goes through is a farnesylation step, where a farnesyl group is tacked on to the end, and it makes the molecule more attractive to lipophilic and more attractive to associating with nuclear membranes. Lots of proteins use this, and that's important that this mechanism is used by hundreds of proteins because

that's going to tell us something about why lonafarnib was developed for other indications.

But what we're looking at here on the right is an inability of progerin to be defarnesylated like lamin A is. So this toxic molecule is permanently associated with these membranes. Lonafarnib was the strategy that we first started to test, saying if we don't allow progerin to be farnesylated, will that help us to create a situation where it's not associated with membranes, and it can be metabolized more quickly, and it can be less toxic to cells.

Next slide, please.

Here you see just a couple of examples of the preclinical research. We got some FTIs, not always lonafarnib, but whatever we could get our hands on. What you're seeing on the top is a normal cell, a very abnormal nucleus in progeria cells, and then how treatment with farnesyltransferase inhibitor -- in this case lonafarnib -- helps those cells to normalize, and that was a critical in vitro experiment.

From there, now that we knew the mutation, labs could create mouse models of progeria, which we couldn't before, and some laboratories worked on giving those mouse models an FTI, and found some improvements. Here I'm showing you some weight improvement. There were other improvements shown as well, like strength, so we had some preclinical both murine and cellular evidence that this drug might work.

Next.

This is what we did. Lonafarnib was already being used in pediatric cancer trials. The RAS protein is farnesylated. There was a pediatric cancer trial going on at the Dana-Farber. They were already giving the drug to children with this cancer, so there was a maximum tolerated dose established in pediatrics.

We were really, really fortunate. We sought out a wonderful principal investigator, Mark Kieran at the Dana-Farber, a neuro-oncologist, who could serve as the PI for a clinical trial, and just repurposed this for children with progeria, and we

developed a team of clinicians who had never seen a child with progeria before but were willing to do this for these kids.

We then started an investigator-initiated trial with lonafarnib at Boston Children's Hospital, and this was investigator initiated, so we didn't need to agree at that time on a primary outcome measure for drug approval. We had a primary outcome measure, rate of weight gain, but we weren't asking for drug approval at that time. Then the drug company agreed to supply the drug for the trial, not as its pipeline, but for us to use, which was pretty amazing, and we've had that happen successively with the success of companies that made that drug. This was our big launch. That was in 2007, the first-ever clinical trial for progeria.

Next.

We brought the kids in from 13 different countries speaking nine different languages. They came in together. It was pretty intensive because we were not only looking at whether this drug was

going to work, and giving this drug, and seeing if it was tolerated well, but also, we needed to develop more on the natural history of progeria because we didn't know enough about it to have really solid outcome measures. We had run a beautiful natural history study at the clinical center at NIH, and that was our first natural history study of that kind, but we still needed to know a tremendous amount more.

Next slide, please.

We evaluated 28 children, and we saw some benefits. We saw a very modest rate of weight gain. It was statistically significant, but it was pretty small. But we discovered some really important things, and one of them that I'm going to focus on is an improvement in cardiovascular stiffness, basically.

We measured that in a couple of ways, something called carotid-femoral pulse wave velocity and something called echo density, and some other things. These were all secondary outcome measures, but we're learning along the way.

We're learning about progeria, and we're learning about what can change in progeria, and some things changed notably, and some things did not.

Next, please.

Now, I'm concentrating here just to teach you a little bit about pulse wave velocity because I'm going to come back to it later on. Carotid-femoral pulse wave velocity is essentially a measure of vascular distensibility, and children with progeria have very stiff vessels.

What you're seeing here is pulse wave velocity, the higher the number, the stiffer the vessel. This is caused by abnormalities in the vessel wall that have been shown on autopsy and in the mouse models. They have very high pulse wave velocity, and that was improved after two years of therapy, what you see here on the right, with lonafarnib.

This measure, the adult population, is a major predictor of adverse coronary events in the adult, and was back then. That's what we knew about it, and a decrease in the adult population of

1 meter per second correlated with lower incidence of heart attacks, so we were very encouraged by this and some of the other data we had as well that was more exploratory.

Next, please.

I'm going to show you the chain of clinical trials. This is what we did. It's highly unconventional, but I think it's really important to understand what we did and why. Here on the top, on the left, this is that first trial. I call it ProLon 1. All the kids were naïve, it was open-label, and there were 28 evaluated.

From there, we entertained another clinical trial that we slid right into. As the children from trial 1 were coming in for their final visit, we wanted to keep them on lonafarnib, and we held a trial, adding two drugs that we thought might be also beneficial over and above lonafarnib. We had this preliminary evidence. We were very excited.

None of the children went off of therapy; they just slid right into this new trial. But what happened then was extraordinary. After

that -- I'll call it the triple trial -- after that ended, was ending, we had more and more evidence that lonafarnib was beneficial, so we asked permission from the IRB and the FDA to not only keep children -- the children that had been on the triple therapy trial -- on lonafarnib, but switch them to just monotherapy while we continue to evaluate.

They also allowed us to bring in new naïve-to-therapy children and put them on the lonafarnib monotherapy without ever going on to the other two drugs, and that started in 2014, and actually through different trials is still ongoing now. I'm going to call that second group, if you look down the bottom of naïve to therapy, ProLon 2 because that's going to feed into this story I am going to tell you.

Next.

This is what we found. Now remember, survival was not an outcome measure in our clinical trials. We never imagined that we could tell in two years if the drug was going to extend survival,

but we embarked on survival studies using our international progeria registry, essentially.

This is incredibly important. I mean, this was a registry that we just created to communicate and keep track of everybody around the real world with progeria and make sure that the populations, that the families, that the children all knew what was going on over time, and it remains one of the most important programs that we have because it's a communication program about what's coming down the pike and educating people. Nobody is surprised when a clinical trial comes to fruition, and there's all sorts of communication both ways.

We did this study. Now, what I'm showing you here, the solid line is a control group of children who did not get lonafarnib. We had a historical, going all the way back, everybody we could find, to the initial publication on progeria, but we also had a concurrent control group, children that lived at the same time as the children who came into the trials.

Everybody that we knew of at the time we

started the trials was offered the trial. So that wasn't the problem; it was just that these were children we didn't know of at the time. So we put that all together, and we published it. The dashed line shows the children that were on therapy, but the therapy was either lonafarnib or triple therapy in this publication. It was sort of a long term, look what's happening here.

Next, please.

In 2015, we were actually in a discussion with the FDA about our next clinical trial of lonafarnib plus a drug called everolimus, which is still ongoing now, and we were talking in this trial about what would be acceptable outcomes for approval because we thought pulse wave velocity would be an excellent outcome for drug approval for this trial and future trials.

We went to discuss this with them, and it was a really, really interesting conversation, because at the time, they said well -- we submitted our packet. Our packet was pretty robust. It included the paper from Circulation, and they said

to us, "Well, right now, pulse wave velocity is not strong enough either in the adult population, and also you don't have something that says pulse wave velocity relates to cardiac outcomes or outcomes in progeria either, but we're really interested in your survival study. Maybe this kind of thing might be supportable if you take apart and only examine the monotherapy."

Now, I want you all to know that the triple therapy, the addition of those other two drugs, did not benefit kids any more than the monotherapy as far as we could tell and have published, but we didn't really know that at the time. And even if we had, they really wanted to see the monotherapy, so that's what we did next.

Next slide.

This is what we did, and this is the interesting part. We took ProLon 1. We had that by design, but it just so happened that just because we wanted to keep kids on drugs without interruption, because we cared so much about our population and essentially were running continuous

clinical trials to do so and were allowed to do that. We had another population of naïve kids that had never touched the triple therapy, and I'm going to call them ProLon 2.

Next, please.

This is what happened. Dr. Brooks showed you this as well. What you're seeing on the left is just ProLon1. Blue is the concurrent control kids. Now, remember those control kids don't come from the clinical trials, but they were from our registry. The red is children on ProLon1, that first clinical trial, and the number of deaths is obviously significantly decreased; and then on the right, you see ProLon1 plus ProLon2 in the dashed line that's above.

With additional analyses that we, and also Eiger, the drug company that I'm going to tell you about did, their label for this drug now says that it extends average &lifespan by -- I'm going to say at least, because that's all we can tell yet -- 2 and a half years. And even one more day, that's 26 percent of the kids' lifespan. Of course

it's never enough, but even one day more of a healthy, beautiful life is incredibly important.

Go to the next slide, please.

The next portion of this story is also interesting. It's a bit of happenstance and luck, but it is a big part of our story. There was a company called Eiger, and Eiger was interested in lonafarnib for an indication called hepatitis delta. I think that's still a rare disease, but it's much more frequent than progeria.

They approached Merck, and they got a license for that, and progeria was part of that arrangement and came along, in the sense, for the ride. But it was very attractive to Eiger because look at this data that we had on survival, and all of this other data that was certainly, we hoped, bringing the menu to them and saying, "Look, you really could get this approved. Let's partner." So we did so, they were interested, and that was wonderful. They are the IND holder and got approval in 2020, our very first drug approval for progeria.

So from trial to approval, we had 13 years of continuous lonafarnib administration. I don't know if a drug company would ever think that way or do anything like that, but we were just thinking about getting this drug safely into the children, and the FDA and the IRB were also thinking of the same thing, so it was pretty wonderful.

Next slide.

I just want to quickly tell you how I'm looking at this now because certainly you don't want your story to be over 20 years, 13 just in clinical trials, and we want to always learn and then compress to do better and go faster for our kids or anybody that we're trying to help.

For progeria, the things that we're looking at now are not likely to be repurposed drugs, so there is an added challenge of drug development. What you see here on the bottom is a mouse model of progeria and lonafarnib being about 25 percent effective for increasing lifespan. But there's a small molecule, there's RNA therapeutic, there's DNA base editing, all from scratch, all first in

human that we're working on, and they're incredibly important.

Survival is not going to work because now lonafarnib is standard of care, so everything has to be over and above what happens, and survival just isn't going to cut it for those, certainly in any reasonable amount of time, but also very difficult to tease out. So our responsibility is to tease out what are the outcome measures that are going to help us here.

Next slide, please.

Since survival isn't viable, we are concentrating on the things that I've mentioned and the things that you've heard about today, so that's pretty exciting. We're developing a progerin biomarker in plasma and have been working to do that for some time now, and I'll show you that in the next slide. But since this is a disease-causing protein, we concentrated on that, and that's going to be really, really important and also, again, still looking at pulse wave velocity and correlating that hopefully with survival in

progeria to show that this matters in our kids and that this will matter in clinical trials.

Next slide, please.

This is just to show you an unpublished, first look at what we found with our progerin in plasma. What you're seeing here is the decreased risk, percent decrease risk, for death as levels of progerin are decreased in the plasma of kids with progeria. So we're pretty hopeful that this will become a viable primary outcome measure if possible, although we know the bar is high, but we're pretty excited about it.

I just want to tell you that the story along the way is we had orphan drug status I think from 2011 on. That was incredibly helpful. The voucher system was helpful, very helpful. And I think what you're doing here is amazing and continuing to say we're entering new eras. We want to change; this can't be traditional anymore, and going with that, and creating new avenues for success for all of us in these rare -- but also the ultra-rare, which is even more difficult -- communities; incredibly,

incredibly helpful.

Next slide.

So thank you very much. Thank you for everything. Thank you for even thinking about all of this, and thank you for asking me to present.

DR. OTTINGER: Thank you, Dr. Gordon.

For our second speaker, we have Dr. Raphaela Goldbach-Mansky, and she is a senior investigator and chief in the translational autoinflammatory diseases section in the Laboratory of Clinical Immunology and Microbiology at the National Institute of Allergy and Infectious Diseases, NIAID, at NIH. She's going to share the journey towards a supplemental biologics license application for anakinra and rilonacept for a deficiency of interleukin-1 receptor antagonists, DIRA.

Presentation – Raphaela Goldbach-Mansky

DR. GOLDBACH-MANSKY: I'm presenting a successful submission of a supplemental biological license application for an ultra rare disease, deficiency of the IL-1 receptor antagonist or DIRA.

My colleague, Dr. Shakoory, will follow with an example of the submission that did not result in a successful approval, both highlighting challenges of ultra-rare disease drug approval. These are my disclosures.

My name is Raphaela Goldbach-Mansky, and I'm chief of the translational autoinflammatory diseases section at the National Institute of Allergy and Infectious Diseases at the National Institutes of Health. My group and program evaluates patients, pediatric patients, with rare inflammatory diseases that present with fever and rashes, and we aim to identify the genetic causes and characterize, the pathogenic pathways, and molecular targets for treatment, with the goal to develop proof-of-concept studies that provide better treatments for those patients.

Untreated disease results in organ damage, morbidity, and early mortality. There are 50 genetic causes of rare autoinflammatory diseases in the INFEVERS database. Of those, there are only five diseases that have approved treatments,

including the one that I'll be talking about today.

This points to a wider problem of rare diseases. The Orphan Drug Act defined rare diseases of less than 200,000 in the U.S. The monogenic diseases I showed you have prevalences of less than 1 in a million, with less than 300 patients in the United States.

The disease I'll present today has a worldwide prevalence of somewhere around 30 patients with that disease, and that actually illustrates the mounting challenges of a wider group of rare conditions, where 80 percent of patients with rare diseases suffer from around 300 diseases and 20 percent from over 6,500, illustrating a need to design studies for these ultra-ultra rare diseases that facilitate and accelerate a drug approval process.

What drove me to seek approval is the ability to secure access to long-term treatment, as patients with successful treatments who require chronic care often do not get assurance approval of prescriptions for drugs that are not approved for

their condition. Furthermore, if approved, the co-pays are often so high that patients can't comply, and they are not eligible for patient-assist programs because the drug they are asking for is not approved for their condition.

DIRA is a disease we discovered. A severe patient was initially treated empirically with the IL-1 receptor antagonist, anakinra, and had a tremendous recovery, and targeted gene searches resulted in the discovery of recessive loss of function mutations in a gene that encodes the IL-1 receptor, the endogenous IL-1 receptor antagonist.

The impressive treatments with recombinant IL-1 receptor antagonist, anakinra, forged a concept where mutations that regulate the proinflammatory cytokine IL-1 -- such as those resulting in gain of function of a sensor that is associated with increased production or with the absence of a negative regulator IL-1 receptor antagonist that causes DIRA -- result in amplified IL-1 signaling and therapeutic strategy to block results in clinical remission of the inflammation,

impressive results which really generated the proof of concept of a significant role of IL-1 in these conditions and was a compelling mechanism of action that supported the regulatory approval of DIRA.

We followed 9 patients at the NIH on a natural history study, where they received treatments, many through the NIH because they could not access drug at the outside. In 2013, Dr. Montealegre, who was a staff clinician in my group at that time, led a study, a pilot study, using a long-acting IL-1 inhibitor, rilonacept, and enrolled 6 patients DIRA, and started data analysis in 2014, showing that the drug, rilonacept, kept patients in remission.

First steps to a submission came from a discussion with the FDA in 2015, highlighting the challenges of providing treatment, which led to the FDA reminding me of the rare disease programs, or orphan disease designation programs, and reaching out to Regeneron, reminding them of the opportunity to file a supplemental biological license application.

In October 2016, after discussions, Regeneron agreed to file an sBLA for rilonacept in DIRA and a briefing book. The database formatting and a clinical study report, together with the analysis and publication of the data, occurred, and in January 2018, a Type B meeting with the FDA led to further discussion and to the FDA encouraging co-filing of a supplemental biological license application, including anakinra, the recombinant IL-1 receptor antagonist, which patients had received before they were switched to rilonacept.

Regeneron, the company, the maker of rilonacept, endorses the plan for a co-submission, and in March we held a conference call between Sobi, the maker of anakinra; Regeneron, the maker of rilonacept; and the NIH to discuss feasibility of the co-submission, which was pretty much endorsed and thought to be feasible. Regeneron completed, with a contract research organization, ICON, the regulatory documents, including the clinical study report and the formatting of the

data for FDA submission.

A short interruption came when Sobi management was unable to support a DIRA co-submission to drain sufficient resources. However, the NIH, or the NIAID leadership, provided me with funds to hire a CRO to help with the filing, and in that context, Sobi endorsed the co-filing and committed to drafting the regulatory modules and draft labels, which were required, and are required, to be submitted, including the preclinical data that support a supplemental biological license application.

The FDA had further requested that we define the study periods clearly. We had 9 patients, and all had pretreatment, IL-1 blocking treatment data, on anakinra, and six were switched to the rilonacept study, as I mentioned. After two years on the rilonacept study, five of those could not secure a drug through their insurances and switched back to anakinra; that at that time, we had received as a donation to support patients who were unable to obtain drug.

For the submission, anakinra and rilonacept had been approved for another IL-1-aided disease, cryopyrin-associated autoinflammatory disease with the subtypes of FCAS and Muckle-Wells, and anakinra for NOMID, and dosing and safety in these populations have been available.

Working with the CRO, we needed to extract the data, the anakinra data, that were collected on a natural history at the NIH, and from documents of hospital admissions, and outside physician records that were provided to us. The data were monitored by the CRO, the CRO assistant, with the development of a statistical analysis plan, clinical study report, and committed to helping with the summaries required for the regulatory submission and the draft labeling.

The statistical analysis plan had no formal sample size or power calculation, as this was retrospective data analysis. Remission rates were computed as rates with 95 percent confidence intervals for time windows that had retrospectively been established as meaningful: day 2 to 6 months;

6 months to 12 months; 12 months to 2 years; and greater than 2 years. Then at the final NIH visit, paired t-tests were used to compare baseline to the suggested time windows for the outcomes that I'll actually discuss in a minute, and hospitalization rates of pretreatment and on treatment where calculated.

Primary endpoint was remission, and that included absence of clinical signs and symptoms; that of DIRA were pustulosis aseptic osteomyelitis and elevated acute phase reactants, indicating systemic inflammation. CRP, an acute phase marker, had to be normal. Their absence of clinical disease, already graphic evidence of inflammation, and patients had to wean off glucocorticosteroids.

Secondary end points included reduction of glucocorticosteroids, and then normalization of markers of inflammation, including separate CRP white blood cell count and platelet count; normalization of hemoglobin; improvement and normalization of anthropometric and developmental

outcomes, including height, weight, and bone marrow density.

Hospitalization rates were requested by the FDA to be collected and were compared. We also had collected patient-reported outcomes, including a disability index, a disease burden module, as well as physician and patient global, as well as patient pain evaluations.

I'll summarize the efficacy conclusions briefly. In essence, all patients achieved inflammatory remission off glucocorticosteroids with anakinra treatment and the remission was maintained with rilonacept. Untreated, the mortality of the disease is estimated to be close to -- well, at least over 50 percent, and long-term survival of untreated patients are not known.

The growth parameters improved, and the hospitalization rate shrank from over 40 percent of the time alive to less than 0.6 percent through pretty much elective surgeries. Questionnaire data and patient-approved outcomes improved significantly. Safety of anakinra and rilonacept

were good, and drugs were well tolerated, and longer term safety data were available for the other diseases.

In addition to the stated documents, we submitted documents documenting the natural history of the disease, which mainly was a summary of the description of the patients that we followed at the NIH and a summary of the published literature. There are a total of 28 patients known; nine had died prior to making the diagnosis and nine were followed at the NIH. We also generated narratives on the 9 patients, summarizing pre- and post-treatment data.

In November 2019, a pre-sBLA meeting between FDA, NIH, and the two manufacturers, Sobi and Regeneron, took place, and in June, a parallel supplemental biological license submission of rilonacept and anakinra occurred with a successful approval in December of 2020.

Anakinra was approved for naïve patients at a starting dose of 1 to 2 milligram per kilo daily with a maximum of 8 milligrams, and rilonacept was

approved for maintaining remission in patients weighing more than 10 kilos.

With that, I want to thank all those who have been involved in this tremendous effort. I'd like to thank the FDA for the encouragement; Dr. Montelegre, Gema Souto-Adeva, Jenna Wade, and Lena Bichell for their work on extracting and generating the data on anakinra; the CRO, ICON, for their invaluable help in monitoring and generating the documents required; Regeneron and Sobi for their willingness to work together; and the tremendous compassion I've seen in many tools support; the submission for this rare disease and for their compassion towards patients; and the Autoinflammatory Alliance for their support.

I won't be able to answer questions in person, but I would be delighted to receive emails and support your efforts in any way I can, so please reach out. Thank you.

DR. OTTINGER: Thank you.

Our third speaker is Dr. Bita Shakoory, and she is also at NIAID in the translational

autoinflammatory diseases section. She's going to discuss baricitinib for autoinflammatory interferonopathies.

Presentation – Bita Shakoory

DR. SHAKOORY: Hello, everyone, and thank you very much for having me. I will go over our experience with the use of baricitinib in patients who have CANDLE, and CANDLE stands for chronic atypical neutrophilic dermatosis with lipodystrophy and elevated temperatures.

You have all heard, "If you hear hoofbeats, think of horses, not zebras." In rheumatology, we are trained to identify zebras among a huge herd of wild horses, based on hoofbeats, stripes, et cetera, et cetera. But in translational autoinflammatory disease section, we get to talk about dotted zebras, so we go beyond just identifying zebras.

So next slide, please.

By the way, that dotted zebra is actually identified in Kenya.

In this discussion, we are going to go

over a little bit of discussion about CANDLE and how baricitinib can be helpful in these patients. We're going to have an overview of baricitinib study in CANDLE, the challenges and obstacles that we have had, and lessons that we learned from communications and submission to the FDA, and how we have learned lessons in moving forward and improving the results.

Next, please.

The genetic discovery of the three monogenic interferonopathies between 2006 and 2014 provided us the pathomechanistic insights into type 1 interferon production in sterile immunodysregulatory conditions, and then led to clinical trials for blocking the interferon signaling pathway as a therapeutic strategy.

These three diseases, Aicardi Goutieres syndrome; the PRAAS/CANDLE, as we mentioned; and SAVI, which is STING -- now I'm blocking on the name of the disease. It's STING -- well, let's go to the next slide. I will tell you when I remember it.

After the disease was identified, it was very difficult to be able to treat these patients until we were able to treat these patients with JAK inhibitors.

In October 2011, we initiated a compassionate use and extended access study with a JAK inhibitor, baricitinib, and enrolled 10 patients with CANDLE, four with SAVI, and four with CANDLE life diseases, patients who didn't have a genetic confirmation but their disease phenotype did resemble CANDLE patients.

We enrolled these patients at NIH, and Dr. Vanderver at CHOP enrolled 36 patients with Aicardi Goutieres syndrome, and then later 5 patients with juvenile dermatomyositis were enrolled as well.

In the initial communication with FDA, the indication for baricitinib for interferonopathies could not be accepted, and we were asked to submit response data by disease only, so as a result, we focused on CANDLE. We had enrolled 10 patients and had seen most impressive clinical results in

CANDLE. The stars you see here are related to issues that we will get back to.

Next slide, please. If you can go back to the previous slide.

I have to also mention that the study that we initiated, at the time we started the study, there were no pediatric dosing, no PK or PD data in children, and there were no template or guidance for dose adjustment, and no endpoints or outcomes were defined.

So we had to basically start from scratch and do reductions, and do basically dose adjustments. We looked at all of the outcomes, endpoints, and we identified, basically, reductions in daily diary scores; corticosteroid requirements; quality of life; organ inflammation; and changes in biomarkers, namely interferon-induced biomarkers for defining the endpoints in this study.

Next, please.

This figure shows the impact of blocking the interferon receptor response by blocking the downstream mediator, JAK inhibitor, to a lesser

extent, to inhibitor on clinical features and biomarkers. Fifty percent of CANDLE patients actually achieved the clinical remission that included very strict parameters of no clinical symptoms. That include fever, rash, headaches, and musculoskeletal pain. They normalized their inflammatory markers completely, which includes CRP and ESR, and they basically were able to come off steroids completely.

In addition, all the patients who achieved remission normalized their interferon signature response gene and validated biomarker of interferon signaling. All the patients benefited. Even those who did not achieved remission, they still benefited from the drug, and they were able to have improvement in their symptoms, lower steroids, and improve quality of life.

This was the first time that patients with CANDLE actually faced a possibility for treatment, though optimal doses that were required for achieving this improvement were about almost 2 times the doses that were given to rheumatoid

arthritis patients that were 4 milligrams per day.

Now, we did observe reactivation of the BK virus and HZV, which we closely monitored. We did not see any of the safety signals that were observed in adult patients with rheumatoid arthritis.

Next slide, please.

These images basically show the face of patients with CANDLE, figuratively. In these images, you see how there's improvement in panniculitis in the face, mainly around the eyes. And in, basically, the middle image, you see the patient who is a 14-year-old girl. You can see the change from pretreatment stature to post-treatment stature in the 36 months after the start of treatment with baricitinib.

Next slide.

Here, you see the timeline for the baricitinib trial in 2011 to 2017. We undertook the compassionate use NIH protocol with Eli Lilly. In 2016, in parallel, you see what's happening with baricitinib. In 2016, baricitinib was approved in

Europe for rheumatoid arthritis, and in 2017, FDA rejected baricitinib for use in RA in the U.S. So what you see is the persistent remission in 50 percent of CANDLE patients in our study, and the narrow therapeutic window does not allow higher doses.

In 2018, while FDA approved the use of baricitinib in rheumatoid arthritis, we filed for sBLA for CANDLE with FDA, and in January 2020, at the time that we had an appointment to have a Type C meeting with FDA, FDA canceled the appointment because they felt there was not adequate data to make a risk-benefit assessment decision in this trial for this drug.

Next slide, please.

Basically, the main criticism was limited data and small numbers, but at the time, as I mentioned, there were just 10 patients that were identified. They suggested use of comparable external, historic controls, and then we decided in discussion with Lilly to undertake rigorous data collection and documentation of every single bit of

historic data. They suggested that we needed to use the historic controls that had comparable endpoints and show objective changes in core clinical outcomes, such as survival.

So we decided, okay, we were going to collect the data, but also, longitudinally, we were going to integrate the data from various physicians, hospitals, and define the flares based on withdrawal data whenever we had to withdraw any patients from the study. We documented the safety narratives and endpoints in order to address some of the FDA concerns.

They felt that there were limited data on safety, and because of the unblinded nature of the study, there was risk of bias. Also, they felt that the risk of the age of the patients and the disease on PK was not very clear, which we understood completely, but this had not been extensively studied prior to that. They felt that the outcomes were not very objective.

One of the points they brought up was caution against the use of proxies in their

reports. Keep in mind that some of our patients are very young, somewhere between 2 years-3 years old, and the daily diary is basically completed by caregivers, parents, and guardians. Actually, these diaries, this is basically considered observer-reported outcome and not proxy, which requires, basically, the proxy data entry would indicate that the person who is completing the form is actually entering their own experiences rather than the patient's experience, but our diaries clearly collect the data based on what is observed.

Next slide, please.

In order to overcome the challenges that were mentioned, we collected the historical data, and we did a complete literature review and combined information from every single patient that was done, and combined those with our cohort data.

Next slide, please.

We also included the dose-reduction data whenever we came across a patient that needed dose reduction, and we showed the increase in clinical flares and associated laboratory changes.

Next slide, please.

We submitted then an enhanced briefing package and tried to address the FDA feedback. We submitted all that in September 2020 to FDA, and then FDA granted the pre-NDA Type C guidance meeting. Keep in mind that, simultaneously, in 2018, there were safety concerns to arise about the use of JAK inhibitors in rheumatoid arthritis patients, and in 2019, based on these concerns, an extensive multicenter safety study, postmarketing safety study, in rheumatoid arthritis is started. So when we, basically, met with FDA in 2021, at that time the data from the safety study was pretty much emerging. At that time, in January 2021, the representatives from the rare disease office also were present in the meeting.

Next slide, please.

The feedback that we received, they felt like the data was inadequate to support risk-benefit assessment. To overcome this, they suggested a randomized withdrawal study. They emphasized that our endpoints were based on daily

diary score and that this was unacceptable.

The review of the published cases, which included all of the existing cases in the literature and any patient that was there with this disease, was inadequate. They also felt that our prospective endpoint data was inadequate and historic data was unclear, which included, basically, very detailed information, and they felt there was heterogeneity in the disease and in the treatment effect.

Then they felt the mission was not sustained; biological plausibility was not well explained, and there was risk associated with higher dosing, and concern about risk of thromboembolic events and serious infections, even though about 10 years of data did not show any of that in the pediatric population with interferonopathies.

So we looked at this basically objectively. There were modifiable aspects and non-modifiable aspects. We asked whether or not we had done the data justice by the way we presented it, and we

also felt like maybe publishing the data in peer-reviewed journals would be more helpful. We also looked at the FDA rare disease guidance document. Based on all of this, we felt like we had done everything we could in presenting the data in an ultra-ultra rare disease to FDA.

Our endpoint was not based on daily diary score alone; it was based on daily diary score as a small part of it, but in addition we had an extensive list of biomarkers and objective data collection by the physician. We also felt like maybe we could expand our patient cohort for the trials by collaborating with a couple of other centers worldwide. However, our patients were from various countries, and this would not add a very huge amount to our effort. We could also reference other small diseases and better defined treatment response parameters.

There were non-modifiable factors such as morbidity and mortality of CANDLE that we really could not do much about. There were patients in our cohort who were taken off the medicine because

of adverse events, who died as a result of not receiving any treatments, and there were concerns about the safety profile of JAK inhibitors that was out of our hands. But when we look at things from risk-benefit ratio, if these patients die or have significant morbidities when not treated, then it's kind of like these are relative in the sense of how bad is the disease, really, as Dr. Pippin was mentioning in the previous session.

We cannot do much about the number of patients or the length of historic data. The disease was discovered in 2010, and we started collecting data in 2011, so there wasn't much we could do about it. We couldn't do anything about negative publicity associated with JAK inhibitors and the barriers of multicenter studies, a coordination between U.S. and UK.

So all of this led to a decision, along with Lilly. We also felt like the patient-reported outcome component of endpoints, along with reduction of steroid dose and disease-specific improvements, were valid endpoints for the disease.

So based on all of the above, we did not feel that making any changes would make a difference in the response we would receive from FDA.

Next slide, please.

So after the meeting with FDA, after much discussion, we decided not to pursue withdrawal studies, especially because those patients who were stable on baricitinib would not be interested in it, and it was not ethical to try to remove them from the medicine.

After we did that in the summer of 2021, in September 2021, FDA issued a black-box warning based on postmarketing safety data in tofacitinib, baricitinib, and upadacitinib, so it kind of seemed like this was a bit of predicted response.

Let's go to the next slide, please.

We tried, but we failed. We failed all of these faces, all of these children that you see here; 11 years of hard work by NIH and Lilly, but most importantly we did not have approved drug for the patients and no approved treatment. Baricitinib is not covered by insurance companies.

Patients are not eligible for co-pays in this program. They can only receive this from NIH through on-site pick up, 11 years of trial participation, which is definitely not easy for these young kids.

We feel like we have failed these kids, and even though there is a drug that can really help them, we were not able to convince FDA that it would be worth approving it for them.

Next slide, please.

So there are lessons that we learned. We realize that there are things that an investigator can contribute such as detailed documentation and, basically, identifying the best outcomes for the disease; documenting the safety data; and flare and response criteria, which we were able to define for this illness.

We were able to learn and optimize our statistical analysis. We also were able to fine-tune enrollment of international patients in collaboration with other major centers. It's something that we're really exploring for our next

trials. We have learned the importance and the ways for IRB approval and patient consent. We have now sent in sample collection and sample storage for our future analysis as part of a network. We are building our infrastructure, and part of that is the platform trials and methodological innovations, as was discussed in the previous session.

Next.

Drug component, it's important to collect PK and PD data. PD modeling and dose-adjustment algorithms, we have learned they should be in place, then we need to, basically, have extensive data about biomarkers and metabolites as much as possible.

Next.

The protocol component, as mentioned, we have thought about crossover design, but this requires a placebo arm, and the placebo arm in a disease like this, where no standard treatment is available, becomes a problem and an ethical issue. The withdrawal study, as I mentioned, there are

ethical issues as well, and we're looking into novel trial designs.

I'm almost done.

Next slide, please.

I think the most important part for us is that we are hoping to start a dialogue with regulatory authorities about some flexibility for rare diseases and rare disease discoveries, innovative trial designs, and manageable regulatory requirements where it's not possible to undertake two trials, or it's not possible to define endpoints, and we have to kind of do this along the way.

We need to establish differences between adults and kids; that children are not small-size adults, and that all the adverse events that happen in adults necessarily do happen in kids and vice versa. The other aspect is that death is not the only poor outcome. As you saw in those children, even if a patient doesn't die, they may have complications that may be worse.

So we're hoping to be able to define

autoinflammatory outcomes that assure investigators of acceptance for existing and novel treatments that are yet to be discovered for rare diseases, and thank you.

Session 2 – Questions and Answers

DR. OTTINGER: Thank you to all our speakers. We will not have too much time; maybe for a couple of questions.

Are all the panelists on currently?

DR. SHAKOORY: Yes.

DR. OTTINGER: I don't know. There were some detailed questions, but I thought maybe to start with more of a larger question, if anyone wanted to take it. All of these were long stories of the winding road that you had to go on through the process, so I'm just wondering -- it's always when you look at the end and look back at the beginning -- is there anything really important or advice you'd like to give when someone starts this process of a possible drug to test that you've learned?

One thing was, Dr. Gordon, when you were

talking about the first trial that you did, the open label, was there anything, if you would go back, that you would do differently in the hopes of collecting more data?

DR. GORDON: It's a very, very, very good question. Everybody wants to know, how could you -- I want to look forward and say, how can we be better, and stronger, and faster? We just wanted to get into a clinical trial which we thought was something that might be helpful; every single child was going to die.

I can tell you about things that felt like they made a big difference in the long run that are kind of boring. We were in Excel spreadsheets, and you need to be in REDCap, or you need to be in something where, in the end, when you try to apply for your FDA approval, you don't have this mountain of, okay, how can we make this regulatory ready and audit ready?

Those are things that you can write those down. But not really, because we were in trial, and if we had waited until we had an acceptable

primary outcome for approval, we might not ever have started, because then that drug went away for cancer.

So I don't know that I regret any anything with that. I would say that learning from what we've done -- us, and everybody here, and others -- I hope it helps FDA to think about how they want to change things for folks with ultra-ultra rare diseases, and others to say how can we springboard off of this to be better, stronger, faster.

I think that's pretty general. I mean, we got in. We got in. We did what we needed to do. It may have ended without an approval, but we needed to see these kids on drug, and then once we realized we thought we had something, we needed to keep them on it. And everybody worked together to do that; an amazing amount of cooperation and collaboration.

DR. OTTINGER: Sorry. Did anyone else want to add to that?

DR. SHAKOORY: I think for us, not only has

early communication with FDA been important, but one of the things we are realizing is implementing factors that would allow -- basically expanding our infrastructure to allow a more efficient data collection and analysis, patient recruitment, et cetera, et cetera, so that we can make the best use of our time and the best use of the limited number of the patients that we have. That's one of the important lessons that we have learned. With so few patients, it's just more difficult if we don't make the best use of all the data that we can get.

DR. OTTINGER: I had one other general question, and it is the small number of patients. I was wondering how, you as both researchers and part of your disease communities, when there's multiple things that come along to test, how are communities dealing with that in terms of being able to run the clinical trials?

DR. GORDON: Bita, did you want to go, or I could go?

DR. GOLDBACH-MANSKY: I could maybe try.

Can somebody hear me?

I think this is a very good question, and I think we do need adaptive trial designs that allow patients with rare diseases from [inaudible – audio gap] -- to another. We can deal with multiple protocols. [Inaudible – audio gap] -- with a number of small patient cohorts. It's really unsustainable and it's quite stressful.

So I think we need to get assistance also by the regulatory authorities to use adaptive trial designs and to use, as baseline, the pretreatment data that basically can then be compared to varying drugs. I think there is no other way of dealing with such a challenge, and [inaudible] -- where we can be much faster in making these drugs available; otherwise, we'll always be running behind in our approval process.

DR. OTTINGER: I don't know if anyone else had anything else quickly to add, otherwise, there were a few specific questions. I don't know if you saw them and if anyone wanted to answer anything specific to what they saw of the questions coming

in.

I know, Dr. Gordon, there were a couple related to your project.

DR. GORDON: Well, I'm more than happy to do post-workshop postings, or emails, or anything like that, of course, and I'm sure we all are.

DR. OTTINGER: Great.

We're at 12:07, so I think we don't want to go too much longer. I think we'll probably end here, and everyone can answer individual questions and really try to address the questions that come in. We appreciate everybody's questions that did come in.

I just want to remind everyone that this is now a break for lunch, so we'll see everybody back here at 1:00 p.m. Thank you, again, for participating so far and really look forward to seeing you at 1:00. Thank you.

(Whereupon, at 12:08 p.m., a lunch recess was taken.)

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

(1:00 p.m.)

Session 3

Katie Donohue - Moderator

DR. DONOHUE: Good afternoon, everyone, and welcome to Session 3 on Core Principles for Clinical Trials.

My name is Katie Donohue, and I'm the director of the Division of Rare Diseases and Medical Genetics at the FDA, and I'm thrilled to be here with you today, with two panelists who are two of my closest collaborators, Dr. Jack Wang, who is a clinical pharmacologist, and Dr. Yan Wang, who is a statistician.

We're going to go through a couple of common challenges and a variety of potential solutions for those challenges when it comes to designing clinical trials for patients with rare diseases, and in particular making the most of small trial sample sizes.

With that, I want to introduce Dr. Jack Wang, who is a clinical pharmacologist. He's a

team lead in the Division of Translational and Precision Medicine, Office of Clinical Pharmacology at the FDA. I work with him closely. He knows more than almost anybody else about how to pick the right dose for patients with rare diseases, and he's going to start off today with a couple of slides, marching through some of those challenges and common strategies for how we address them.

So with that, I'll turn it over to you, Jack.

Presentation – Jie (Jack) Wang

DR. J. WANG: Thank you, Katie.

Good afternoon. My name is Jack Wang. It is my pleasure to participate at this workshop. I hope my presentation will be helpful for academic investigators and the pharmaceutical companies developing drugs for rare disease. The topic for my presentation is Dose Optimization for Rare Diseases.

Next slide, please.

This is my disclaimer.

Next slide.

Why are dose selection and optimization important? I would like to share the results of two surveys. The first survey, based on more than 300 new drug applications, showed that uncertainty in dose selection was the leading cause of failed new drug applications. The second survey, based on 40 approved new drug applications for rare genetic diseases, recently conducted by my acclaimed former colleagues at FDA, showed 82 percent of approved new drug applications had a dose-finding component.

Next.

With the importance of dose finding as background information, in the first part of my presentation, I will give an overview of clinical pharmacology principles in dose optimization. The second part will focus on the use of biomarkers in dose selection and as confirmatory evidence of effectiveness. We'll briefly introduce an adaptive trial design for dose selection and optimization. Takeaway messages will be provided, and a few case examples will be discussed in the presentation.

Next.

With a final goal of dose optimization and therapeutic individualization, for every new drug application, the clinical pharmacology reviewer will need to address two peer-reviewed questions. First, is the proposed dosing regimen appropriate for the general patient population? And second, is an alternative dose regimen needed for subpopulations?

To answer these two peer-reviewed questions, the reviewer will assess exposure-response relationships for efficacy and safety and to identify potential intrinsic and extrinsic factors that may influence the disease, exposure, and a response.

Next slide, please.

Specific clinical pharmacology studies are needed for the reviewer to assess intrinsic, extrinsic, and other factors affecting exposure and response of your drug product to ensure you have a complete clinical pharmacology program in your early interaction with FDA in the IND stage. For example, in the pre-IND meeting, you probably will

receive a long list of standard comments. It is important that you discuss with FDA your specific drug development program; which clinical pharmacology studies are needed; when do you need those studies; and what are the potential alternative approaches?

Next. Next slide, please. Sorry. Go back one slide.

Let's look at what an exposure-response relationship means. Exposure refers to different dose levels or drug concentrations. Without exposure information, it is not possible to evaluate exposure-response of your drug product, therefore clinical pharmacology reviewers will assess your IND protocols very carefully to make sure you are collecting PK data in the trial design. Response could include either desirable clinical response and undesirable clinical response. Exposure-response analysis is to relate the drug concentrations to clinical responses. This is often done by a modeling approach.

Next.

Exposure-response information plays a critical role in the regulatory decision making such as to guide the dose selection; to provide evidence of effectiveness; to recommend dosing regimen in specific patient populations like pediatric patients; and to assess substantial safety endpoints, for example, QT prolongation.

I have provided a few important FDA guidances in this slide and will provide a few case examples later in the presentation.

Next.

How about our current experience in dose-finding studies for rare disease programs? In the same survey we conducted for the 40 approved new drugs for rare genetic diseases, only 53 percent of the applications conducted dedicated dose-finding studies. Population PK and exposure-response analysis, however, were used in the majority of the applications.

The survey results indicated two things. First, there is a long way to go to convince other sponsors to conduct dose-finding studies in their

rare disease programs; and second, on the other hand, the good thing is sponsors are aware of the regulatory expectations on population PK and exposure-response analysis and have used it as alternative approach to dose selection.

Next.

Here are three case examples using population PK and exposure-response to support a dosing recommendation in an NDA or BLA. In the first example, lonafarnib, approved for progeria indications, the PK and exposure- response information supported expanding the indication from 2 years of age and older to patients 1 year and older.

In the second example, fosdenopterin, approved for MoCD type A, the PK stimulation supported a dose adjustment in patients less than 1 year of age. In the third, avaglucosidase alfa, approved for Pompe disease, PK stimulation was used to extrapolate the indication from 16 years of age and older to 1 year of age and older.

These three examples demonstrate a general

approach, using PK and exposure-response information to justify dosing in a subgroup of patients that were not started in clinical trials.

Next.

To summarize the first part of my presentation and to provide to you some important additional reminders, I would like to emphasize a few takeaway messages. First, conduct organ impairment studies and specify organ functions in inclusion/exclusion criteria of the study protocol.

Second, conduct at least in vitro DDI studies before the first-in-human trial and update is allowed on the prohibited co-medications in clinical trial protocols as DDI data evolves. Third, for oral drugs, investigate food effect early and specify food conditions in clinical protocol.

Fourth, include dose ranging as part of your drug development program, and number 5, as very important reminders, always validate your PK and PD assays, and use the to-be-marketed drug product, or formulation, in your efficacy and safety trials.

Next.

Challenges in the drug development program for rare diseases are often the challenges of clin pharm approaches to dose optimization. A very small number of patients in rare disease clinical trials is to a very low computational capacity in PK/PD analysis. Rare disease often has its heterogeneity in disease pathogenesis, which may confirm the exposure-response analysis.

Rare disease trials often do not have a well-defined clinical endpoint that directly reflects the mechanism of action of a drug. This together with confounding factors by disease heterogeneity make these partial response analyses less effective or informative.

Next.

There's also good news in dose optimization for rare diseases. As shown earlier, population PK and exposure- response analyses are well used in rare disease NDA/BLA submissions to facilitate dose optimization. The methodologies are ready to use. The results from rare disease clinical trials will

be more generalizable to the overall patient population because a high percentage of the patient population already enrolled in clinical trials.

To overcome the issue that clinical endpoints are not well defined to guide the dose, PD biomarkers can be used in dose finding and also as confirmatory evidence of effectiveness. It is important to involve dose finding at early stage. Success can be planned, and the dose optimization can be achieved by a successful clinical trial design.

In the next few slides, we will look at dose perspectives in detail.

Next.

The concept of confirmatory evidence has been introduced in Session 1 of the workshop. The regulatory framework allows the sponsor to demonstrate substantial evidence of effectiveness by conducting one adequate and well-controlled trial plus confirmatory evidence. Confirmatory evidence can be from different sources. From a clinical pharmacology perspective, very often these

will be the PD data from clinical trials.

Next.

Here is a list of a few things you should keep in mind when you use PD biomarker data as confirmative evidence. The selected biomarkers should be relevant to both the mechanism of action of the drug and the disease pathophysiology. However, the selected biomarker does not need a surrogate endpoint that has been validated to predict clinical efficacy outcomes, and the data is not necessary to be collected from the pivotal efficacy and the safety trial.

To show an exposure-response relationship of the PD biomarker data, support is used as confirmatory evidence. As a very important reminder, the bioanalytical assays for the PD biomarker should be validated.

Next.

In the survey we recently conducted among the 40 approved NDA and BLA for rare genetic disease, the majority of the dose-finding studies used the PD biomarkers or secondary endpoints.

Because PD biomarkers are usually more sensitive to treatment compared to clinical endpoints, the use of PD biomarkers in dose finding requires a smaller number of patients and a shorter treatment duration, which are desirable trial design features for the rare disease program.

Next.

Let's look at one example of using a biomarker as confirmatory evidence and to support a dosing recommendation. Fosdenopterin was approved by the DRDMG last year, indicated for patients with MoCD type A. Patients with MoCD type A have elevated levels of neurotoxic sulfite SSC. Urinary SSC decreased following treatment with fosdenopterin. As shown in the figures below, higher plasma drug concentrations were associated with lower urinary SSC or better PD response.

The exposure-response relationship supported the recommended dosing regimen and further supported the use of the biomarker data as confirmatory evidence.

Next.

There are three basic types of clinical trial designs to explore dose response or exposure response: crossover, parallel, the titration. The crossover trial design should use a reversible response endpoint. Parallel design is suitable for long-term treatment with chronic response and needs a relatively larger sample size. The titration approach is used in many rare disease programs because this approach could provide both a population and an individual exposure-response, and you need a relatively smaller sample size.

The big drawback of the titration approach, however, is the potential carryover PK or PD effect when the dose is titrated from one level. In this design, dose selection occurs at the phase 1 and phase 2 part of the trial. Different dose selections approaches could be considered such as parallel group dose ranging, individual dose titration, and in some cases using the maximum tolerated dose. The selected dose will then be evaluated for confirmation of efficacy in the phase 3 part of the adaptive trial.

Next slide.

Here are some takeaway messages for part 2 of my presentation. It is important you establish the comprehensive biomarker assessment plan in early phases of clinical development and have bioanalytical assays validated to use. Make sure you collect PK and PD samples or assessment plan in early phases of clinical development and have bioanalytical assays validated for use.

Make sure you collect PK and PD samples in all clinical trials for exposure-response analysis. When dedicated dose-ranging trials are not feasible for your program, consider using adaptive designs that incorporates both dose selection and confirmation of efficacy of the trial.

Next.

I want to thank my team members and colleagues in the Office of Clinical Pharmacology and all medical officers in DRDMG for their support. I also want to thank the planning and organizing committee of this workshop to give me the opportunity for this presentation. I think

knowledge sharing and collaboration are very important to bring new treatments to patients with rare diseases.

Thank you all for your time. I will be happy to take any questions in the Q&A session later.

Back to you, Katie.

DR. DONOHUE: Thank you, Jack. Your presentation sparked lots of good questions that we'll get to in a minute in the Q&A.

I do see that a few folks were having trouble with slides not advancing, so a couple pointers. One, try using Chrome as your browser -- we seem to have better luck with that one -- and then click "refresh." They do appear to be advancing, but those are two things you might try if it's not working for you.

With that, I'm going to move us into the next part of our talk. I'm wearing two hats in this session, a moderator and a panel member, so now I'm wearing my panel member hat.

If we can advance please.

Presentation – Katie Donohue

DR. DONOHUE: Okay. We're going to talk about endpoints. One of the things that I wanted to highlight is that when it comes to clinical trial design and rare diseases, obviously, endpoint selection is one of the toughest and most important decisions that we make. One of the things I wanted to touch on is this tension between what matters to patients and what scientists can measure well.

Often, there are aspects of the disease that contribute greatly to patient suffering and it matters greatly to patients. But for whatever reason, we don't have a good way to measure that scientifically. So a good endpoint is going to be in that middle part of the Venn diagram where it's important to patients and we can measure it well. Most diseases have at least a few symptoms or manifestations that are very important to patients but that we can't measure well, and those don't make good endpoints; they need to have both.

So when we think about what allows a scientist to measure something well, I often think

about back in the Middle Ages if you wanted to measure how high a horse was, you could use hands and put hands on top; so you can measure by hands. Well, that's not very precise. Now we might use a ruler and get a much more precise measurement.

So a good endpoint is something that we can measure precisely, and typically it's also something that changes fairly quickly or early in the course of the disease. How quickly that thing changes in the disease also really matters for a successful clinical trial, because if you pick something that changes slowly, you might get a clinical trial that's years and years long, or it might never work at all.

So finding that sweet spot of something that changes pretty quickly, that we can measure precisely, and that matters to patients, that's what's going to make a good endpoint. So that's one of first principles in clinical trial design, and I think one of those things that's really important to acknowledge.

Cognition, for example, is one of those

things that we know matters deeply to patients. It's very clinically relevant, but our tools for measuring it aren't very precise, and in most diseases, it doesn't change very quickly. So it often is not a good endpoint for trials, even when it's an important part of the disease.

Okay. Next slide, please.

One of the questions we got in the run-up to the conference was when can single-arm trials work? This is a source of great frustration, I think, for a lot of our stakeholders about when can they work, and when don't they, and why are we always saying we need randomized trials?

I think the hard part of this is that single-arm trials work when you are very lucky, so let's talk about this. There are three main factors, and one is, do you have an objective endpoint, something like an x-ray or a blood test, with lots of evidence, scientific evidence, to show us that a certain amount of change on the x-ray and blood test predicts a certain amount of change for the patient in the clinic? So we know that if we

see this amount on the blood test, we're going to see this much improvement in the clinic.

If you have an endpoint like that, a blood test or an x-ray, that everybody knows predicts how patients do clinically, well then we can start thinking about single-arm trials. But without an endpoint like that, in general, single-arm trials usually aren't going to work very well. At least that's one factor you can control, is which endpoint you're picking.

But there are two things that are crucial to a successful single-arm trial, and this is why I talk about it a lot, because we can't control them, and neither can you. The first is whether or not the natural history of the disease is stable over time. What do I mean by that?

I've included a reference down at the bottom of the slide. It's a really fascinating paper where some cardiologists took a look at three rare cardiac diseases. They looked at these natural history studies of these diseases, and they noticed that mortality was improving pretty dramatically in

some cases.

There was one natural history study where in the space of just two years, on average, patients were living 25 years longer. That's extraordinary. I mean, if we could bottle that, we wouldn't need any doctors anymore. But the problem is that there were no new treatments driving that difference in mortality. What changed was the availability of the diagnostic testing.

So within a very short period of time, it was much easier to get diagnostic testing done for this disease, so in a very short period of time all kinds of new patients were identified with this disease and data type as the others, but they had a much milder clinical phenotype, so those patients were living a lot longer.

Even though there was no new treatment, the natural history of the disease changed right underneath the feet of these investigators, and the truth is that that's happening for most of the genetic diseases that my division works with. None of us can control that. So you can start a trial,

and two years later, the natural history can be different because the genetic testing availability is different, so you've got to be able to guard against that in thinking about a single-arm trial.

The third factor is dramatically effective treatments. We know that with single-arm trials, there are potential sources of bias that are concerning, so you want to make sure that you're seeing a really robust treatment effect if you're going to rely on a single-arm trial to support a drug approval.

Again, this is not the kind of thing that you can count on up front; it's really about luck, and there are some exceptions to that. We know that often, for example, gene therapies do tend to have dramatic results, so that may be a scenario where you'd want to think about doing a single-arm trial because you're anticipating, and you have preclinical evidence for, the potential for a really dramatic effect.

But my portfolio has a growing number of drug development programs that have hit dead ends

because they've done a single-arm trial that looks promising, but it's not robustly persuasive; maybe because the natural history has aged a bit during the course of the trial; and maybe because the treatment effect looks a little bit modest, we just can't tell if it's the drug that's doing this or if the natural history has just changed. So often, there's no good path forward at that point. So I would say pursue a single-arm trial with caution.

Next slide.

Which brings me to the key point, which is, in general, in rare diseases, it's best to randomize the first patient, in part, because you can't control two of those key factors. A really good insurance policy for debriefing drug development in rare diseases is to randomize starting from the first patient.

Next slide, please.

The second core principle here is to be good stewards of the perception of equipoise. What do I mean by that? The reason we do clinical trials is to try and figure out whether or not a drug works

and to generate the scientific evidence to show that a drug is working. When we think a drug works, that's an hypothesis. That's a guess. We need to do our science. We need to do our experiments in the trial to prove that it's working.

So until we've collected enough evidence to prove that it's working, well, we don't know yet if it's working. That's what equipoise means; we don't know yet if it's working. So it's really important that all of our stakeholders be good stewards of this perception of equipoise, and that starts with patients. If you're in a clinical trial, it's important that patients not be on social media claiming benefit from treatment if they don't even know which treatment they're on. That's important.

Patients have a really important role to play in being good stewards of equipoise because if you want patients to enroll in a clinical trial, especially if it's got a placebo arm, then we all need to stay a little bit skeptical about whether

or not something is working.

This is also true for academics and really challenging for academics because, obviously, publishing positive results is what drives academic careers. Announcing good news is like the great privilege of being an academic, and I think the key here is to be really careful about how and when you describe those results.

So overstating the conclusions, concluding that a drug works based on an early-phase trial or a single-arm trial and publishing that before there's enough scientific evidence to really demonstrate it can create a huge problem because, suddenly, nobody wants to enroll in the control trial that needs to happen in order to get the drug approved.

That's one of my key messages, is that it's important for all of us to be good stewards of the public perception of equipoise in order to create the circumstances that we need to for clinical trials to succeed.

Okay. Next slide, please.

If you remember nothing else from my talk, I think this is probably the slide you want to think about. It touches back to a lot of what Jack covered in his presentation, and what it shows you is a strategy for doing some dose ranging in rare diseases.

We know that in most rare diseases, there are not enough patients to do a stand-alone, phase 2 dose-ranging trial and then two separate stand-alone, phase 3 confirmatory trials. We know that. This schema, this roadmap that I've got on this slide, is one option for how to do this in rare diseases, and it's often called a seamless design.

What it means is you start out by randomizing the patients. Maybe, let's say, you have 20 patients. You could randomize five each to these four arms: high dose, medium dose, low dose, or the control arm. Then you might follow them for a short period of time, a couple of weeks, a month or so, and look at what we call a pharmacodynamic endpoint. This is usually a blood test, a

biomarker, something that you can measure quickly. We think it probably correlates with the disease. We don't have to know that it predicts the disease. It's just something you can measure relatively quickly and easily that should give us some sense of how well the drug is working in the patients, and often we are surprised by the results of this.

Commercial sponsors tend to be the ones who do the best in the lowest dose ranging, and I'm often surprised by the dose that ends up being the one that gets carried forward. But the key here in this seamless design is that you can look at all of this evidence, you have an unblinded clinical pharmacologist who is specially kind of isolated and gets to look at this data, and they can say, "Oh wow. It turns out we really need the high dose for this program; we're not seeing much of anything with the medium and the low dose." But then those patients who were initially randomized to any one of the three treatment arms all get moved on to whatever that optimum dose is.

So maybe it's the high dose or maybe it's

the low dose. Whatever it is, everybody initially randomized to treatment moves on to the phase 3 part of the study on that optimum dose. Meanwhile, the patients who were initially randomized to control continue on control, and then you follow these patients out for a longer period of time for whatever your clinical endpoint is going to be.

This does several things. One, it lets you finish your overall drug development program sooner because you're essentially starting your phase 3 trial with your phase 2 trial, because the baseline you're going to use to measure your treatment effect is going to start at the beginning of that phase 2. Secondly, you get some dose ranging in, and as Jack noted, that is one of the biggest risk factors for a failed drug development program in rare diseases, is inadequate dose ranging.

So anything you can do anything in order to incorporate at least some dose ranging before you jump into your pivotal trial I think is a crucial factor for success.

Okay. Next slide.

Another way that we can do adaptive trials is to adapt the trial duration. I'd seen some questions about how do you design a trial when the natural history is sparsely described or really heterogeneous? Well, this is one of the strategies that we use.

We know how quickly patients progress can sometimes be very variable, and it's not at all uncommon for that to be a little bit different in a randomized trial than it was in whatever we were seeing in the natural history. So planning to adjust the length of your trial to what I call the Goldilocks trial, the just-right long enough trial, is a great strategy for de-risking rare disease drug development.

What this means is that you would plan to take prespecified interim analyses at designated intervals. You might say, okay, when two-thirds of our patients have hit the 6-month mark, we're going to take a preliminary look, and again, this is prespecified. You've got dedicated guardrails around who gets to look at the data and who

doesn't, and it's all written out in your protocol and statistical analysis plan. It's not one senior investigator unblinds himself every few months and looks at the data. That's not what we're talking about. But you've got your data safety monitoring board and you've got your plan with the appropriate guardrails to take an interim look at your data and see.

Then if there's a dramatic difference, and it turns out that the treatment is a whole lot more effective than anyone could have possibly hoped, well maybe you're done; you stop the trial essentially early. If it looks promising but it's not quite there yet, you continue the trial for several more months, take another look, and so forth.

So that's another strategy for revisiting drug development because it prevents you from the other major risk factor, which is too short of a trial. We see this all the time. When you have small sample sizes and a lot of uncertainty around how quickly these things progress, adapting the

duration of your trial gives you another insurance policy and protects you from stopping too soon for an otherwise promising therapy.

Okay. Next slide, please.

Estimands. For my clinical investigators out there, before your eyes glaze over, stay with me. This is a statistical concept, but it's actually really a clinical concept. You should never, ever, ever let your statistician off the hook until you've had at least one meeting where you talk about the definition of the estimand. What do I mean by this? Really, it's about intercurrent events. There's more to that definition, and I've included a footnote, and Yan, who's going to speak next, can talk more about this.

But the bottom line is that when we're talking about rare disease clinical trials, data are almost never missing at random. You know this if you're an investigator. You know your patients. Patients are committed to finishing these trials in rare diseases. They don't just like forget to show

up to their final trial visit because they got busy. These communities are devoted.

A statistical plan that just says, "Oh, yes. We assume that any data missing will be missing at random," it's not doing anybody any favors. Don't do that. It's wrong. Think about it. Think about it ahead of time. For most of these diseases, we can anticipate that some patients are going to have clinical events during the course of the trial that might interfere with our ability to measure and endpoint.

A classic example is the 6-minute walk test. Well, if you've got a disease where some patients develop hip dysplasia and might need a hip replacement over the course of a very long trial, you've got to think about that if your endpoint is a 6-minute walk distance. So a patient who drops out of the trial because they need a hip replacement, well, that data isn't missing at random.

So you want to think about that. What are the kinds of clinical events -- and maybe they're

infrequent in the disease but they happen, and they might affect my endpoint. Think about those things. Think about what things might happen to these patients clinically that would get in the way of your ability to measure the endpoint, and figure out how to incorporate that in your endpoint definition and into your analysis plan. You can actually increase your statistical power by planning for that, and planning for how you're going to account for that.

Similarly, with missing data, in rare disease trials you can also have data missing just by chance. Another example might be, again, a trial with a 6-minute walk test endpoint, a patient who has shown pretty dramatic improvement in the 6-minute walk distance over the course of the trial, we don't know if they're on placebo or treatment, but certainly they're doing a lot better, and then they have a car accident on the way to their final study visit.

What on earth do we do with that? In a small trial, that chance event in one patient can

really have a big effect on the results, because we want to think about and protect yourself from some of those chance events. So talk to your statistician about should we take an area under the curve approach. What can we do to protect ourselves from one or two chance events really derailing our estimate?

In a big trial with a thousand patients, you kind of don't have to worry about it. You can just say missing at random, and it'll work out, but small trials, we really can't count on that. So investigators definitely owe it to themselves to sit down with their statisticians and think through intercurrent events, chance events, and how you're going to want to plan for that in your analysis; so that's estimands.

Next slide.

Regulatory flexibility. My picture is not coming across. There's supposed to be a little balance underneath this. This is really about these broader principles at the FDA; how do we balance unmet need and scientific integrity when

we're applying regulatory flexibility?

I think, in general, we often tend to have a pretty broad agreement with our stakeholders about the degree of unmet need. We all agree that diseases that are more severe have more unmet need, and diseases that have no approved treatments or few treatments that are mildly effective, these diseases have unmet need.

Our differences of opinion with stakeholders' are pretty minor. Usually we all agree that this is a terrible disease and we need effective treatments. The question then becomes about when and how to apply regulatory flexibility, and one of the things that I want to share is that there are scientific factors driving the different kinds of regulatory flexibility that we can apply in a given situation.

There are times when there's a vast unmet need. There might be scientific reasons why we still need a randomized control trial and we can't use a single-arm trial. If, for example, the endpoint we're going to be using is a

patient-reported outcome measure, well, those almost always require randomized- controlled trials. As a general rule, you can't do a successful single-arm trial for those kinds of endpoints. You really need one of those biomarker endpoints like an x-ray or a blood test for a single-arm trial to work. So the kind of regulatory flexibility that you might apply has to be balanced with scientific integrity. Are the results of this trial going to make any sense? That's one factor.

Another one is around when can we use accelerated approval. And again, whether or not patients with a certain disease have a biomarker with lots of scientific evidence showing that a certain amount of change in the biomarker is going to predict a certain amount of change in the clinic, if you're lucky enough to have one of those biomarkers, one of those blood tests or x-rays with decades of scientific evidence showing that it's tied to the clinical outcomes, if you have one of those biomarkers, gosh, that's such a blessing, and

it makes it a lot easier to do trial designs in accelerated approval, but the state of that scientific evidence has nothing to do with how much unmet need there is.

So those things aren't always as tightly correlated as we might hope. Really, what regulatory flexibility comes down to is how much uncertainty is going to be acceptable for this disease. One of the main ways that we bring regulatory flexibility into the rare disease space is by requiring one well-controlled trial plus confirmatory evidence. I know that's been the subject of the entire panel discussions at this point, but that's a major source of flexibility that we often bring to rare disease drug development programs; so think smaller sample sizes.

There are a variety of ways that we can think about how to bring regulatory flexibility for a given drug development program, but it's driven in part by the unmet need and also by the scientific factors that are specific to that

disease. So we have to get a little creative about what can work here and what is feasible. That's what I wanted to touch on in terms of regulatory flexibility.

Next slide, please.

And that's it.

Next up is my colleague, Yan Wang. She's a statistician, Dr. Yan Wang, and she is one of the best statisticians in the building when it comes to rare disease trial design and thinking about how to maximize the chances of success, even with a very small sample size.

So without further ado, Yan.

Presentation – Yan Wang

DR. Y. WANG: Thank you. Thank you, Katie, for the kind words and introduction.

Good afternoon, everyone. In my talk today, I will focus on Statistical Considerations in Rare Disease Clinical Trials.

Next slide, and next one.

As a quick outline here, I will briefly discuss some key concepts related to trial design,

endpoint, and analysis. For sample size calculation, I will highlight three approaches that may be used to increase the chance of detecting a treatment effect. There is sample size through estimation, treatment duration adaptation, and global tests for multiple endpoints. I will conclude with a brief remark on the importance of having high-quality trial conduct and data collection.

Next slide.

Before I cover the statistical aspect of my presentation, I would like to first highlight the major challenges in drug development for rare disease, especially for inborn errors of metabolism, IEM. They include small and sometimes very small patient populations.

A rare disease is typically characterized as having fewer than 200,000 patients, but many IEMs have less than a few thousand patients. Their natural history is often poorly understood. It may affect multiple organs and tissues and have heterogeneous clinical manifestations. There is

often a lack of understanding and consensus on the efficacy endpoint. It is difficult to design trials for new drug after the first approval. Lastly, efficacy outcome measures usually have large variabilities, as shown in the next slide.

Next slide.

In this example, the efficacy outcome is the change from baseline at one year in the distance walk during a 6-minute walk test. The table shows the mean and standard deviation estimated using the data from two cohorts of patients with late-onset Pompe disease. In both cohorts, the magnitude of the standard deviation is more than double of the magnitude of the mean.

The figure on the left shows the individual data having a huge spread, going from a loss of 400 meters to a gain of 200 meters. The figure on the right shows no clear relationship between the baseline values and the outcomes at one-year.

Next slide, please.

The patients in these two cohorts came from two different trials, but they received the same

treatment. The question was, was the observed difference in the mean outcome due to chance alone or due to difference in baseline disease severity, standard of care, or procedures for the 6-minute walk test? Was the studied treatment effective? To answer these questions, we need a randomized placebo-controlled trial.

Next slide.

In our experience, randomized double-blind and placebo-controlled trial design is most commonly used because it is the most reliable design to determine the effectiveness of a drug for many rare diseases. In this design, randomization is used to ensure unbiased assignment of patients to treatment arms, and the assigned treatments are blinded to both the patients and the investigators.

Minimization and blinding are the most efficient strategies to minimize potential biases that may be caused by differences in baseline prognostic factors: placebo effect, observer effect, and differences in standard of care. Placebo control does not imply that the control

group is untreated. All patients should receive standard of care. This will limit ethical concerns.

Next slide.

Primary efficacy endpoints, these are the endpoints that provide key evidence of efficacy for drug approval. The most straightforward and readily interpreted primary endpoints are those that directly measure how a patient feels, functions, or survives. They can also be validated surrogate endpoints or validated clinical outcome assessments. A surrogate endpoint that is reasonably likely to predict clinical benefit can be used for accelerated approval.

In a rare disease trial, a composite endpoint is often used to capture the heterogeneity of the disease. It integrates or combines multiple measurements into a single variable. For example, for Fabry disease, a composite endpoint can be the time to the first occurrence of death, renal, cardiovascular, or cerebral vascular events.

Another example is the total Chorea score

for seven different parts of the body in patients with Huntington's disease. While a single primary endpoint is typically used, multiple primary endpoints may be selected to cover the range of treatment effect for some rare diseases. For example, the 6-minute walk test and FVC endpoint can be used as the primary endpoints in trials for patients with LOPD, MPS-I, and MPS-II.

Next slide.

Statistical analysis. The protocols describe clearly the principle features of the statistical analysis of the primary endpoints. The null and alternative hypothesis should define and indicate which parameters are used to quantify the treatment effect. For continuous outcomes, the treatment effect may be the difference in means or medians between the treatment groups. For binary outcomes, it may be risk difference, relative risk, or odds ratio. For time-to-event outcomes, it may be the difference in survival probabilities, restricted means or medians of survival time.

The protocols also include details on the

method for estimating and testing the treatment effect, the methods for controlling type 1 error rate, and the methods for handling missing data.

Next slide.

Sample size determination. One key major challenge question in trial design is how many patients should be enrolled? In principle, the sample size should be large enough to provide a reliable answer to the question. Does the test drug have a treatment effect? The protocol should provide detail on the four key elements impacting sample size calculation.

The first is the null hypothesis and the method for testing this hypothesis. The second is the significance level or alpha level, also known as the type 1 error rate. It is the probability of erroneously rejecting the null hypothesis if the drug has no effect. The lower the type 1 error rate, the more likely it is to avoid a false positive claim, and the more samples needed. While it is conventionally set at the 0.025 for a one-sided test or 0.05 for a two-sided test, a

larger type 1 error rate may be used for an ultra rare disease.

Next slide.

The third impacting sample size calculation is power, which is the probability of detecting a true treatment effect when a drug has an effect. The higher the power, the more likely it is to detect a treatment effect when it exists and the more samples needed. Conventionally, power is set at 80 percent or higher.

The last element is the effect size assumed under alternative hypothesis. It depends on the assumed treatment effect and the variability of the efficacy endpoint. For continuous endpoint, the effect size is the ratio of the treatment effect and the standard deviation of the efficacy endpoint as shown in this equation here. The larger the effect size, the easier, it is to detect an effect and require fewer samples.

Next slide.

How to estimate the effect size in sample size calculation? In principle, effect size should

be estimated based on the minimum effect, which has clinical relevancy, or published data, or the result of an earlier trial in similar settings. However, for rare disease without approved therapies, there are often limited or no data available to estimate the effect size. In our experience, rare disease trials are typically sized based on the assumed large effect size, however, most drugs have a moderate effect size if they have an impact.

Next slide.

This slide shows the effect size estimated using the data from three randomized placebo-controlled trials. The trial is for patients with MPS-1, the second trial for MPS-2, and the third trial for LOPD. For the 6-minute walk test endpoint, the effect size ranged from 0.48 to 0.6. For the FVC endpoint, the effect size ranged from 0.27 to 0.65.

Next slide.

Here are some examples of sample size and power calculations for placebo-controlled trial

with 1 to 1 randomization ratio. To attain a power of 80 percent, a sample size of 33 per arm is needed for effect size 0.7. For effect size 0.6, a sample size of 45 per arm is needed. For effect size 0.5, a sample size of 65 per arm is needed.

In our experience, most trials for IEM have a sample size less than 30 per arm, and thus these trials are underpowered with a power of less than 50 percent to detect a statistically significant treatment effect if the test drug has a moderate effect size 0.5 or less. So a question is, how can we increase the power to detect a treatment effect in rare disease trials?

Next slide.

In the next few slides, I will briefly discuss three approaches that may be used to increase the power to detect a treatment effect. They are sample size re-estimation, treatment duration adaptation, and global tests for multiple endpoints.

Next slide.

Sample size re-estimations. This method is

used to address the uncertainty on the assumed effect size in sample size calculations. Based on interim data, this method investigates the validity of the assumed effect size and increase the sample size if the conditional power, the interim data, is promising.

The conditional power is calculated based on the assumption that the future effect size will be the same as the one estimated from the interim data. If the conditional power is promising, for example, over 50 percent, the sample size can be increased to attain a higher power; for example, 80 percent. If the conditional power is favorable, for example above 80 percent, the sample size will not be increased.

Next slide.

Here is a hypothetical example of trial designed with a sample size re-estimation. The trial starts with a planned sample size of 33 per arm based on an assumed large effect size 0.7 for the 6-minute walk endpoint to obtain a power of 80 percent. This trial planned to increase the

sample size up to 50 per arm if the predefined interim analysis is promising.

The interim analysis is run after the first 20 patients per arm, and the estimate effect size is .55, which is 20 percent smaller than the originally assumed effect size. Because the treatment difference is smaller, reduced from 35 meter to 30 meter, at the same time, the standard deviation increased from 15 meter to 55 meters.

Based on the internet data, the conditional power is 65 percent and is promising. The sample size is increased to 45 per arm, which is a 36 percent increase from the original planned sample size to attain a conditional power of 80 percent.

If this trial is designed with a fixed sample size strategy based on effects size of 0.55, a sample size of 54 per arm is needed to obtain a power of 80 percent. This will represent a 20 percent increase in sample size compared to the adaptive design with sample size re-estimation.

Next slide.

Treatment duration adaptation, Dr. Donohue mentioned earlier. This approach is used to address the uncertainty on the treatment duration needed to demonstrate efficacy. Adaptation is based on the analysis of an efficacy endpoint assessed at a predefined interim time point for all patients.

If the analysis shows convincing efficacy, the randomized treatment can be stopped early, prior to the predefined maximum duration, Tmax. If the analysis does not show convincing efficacy, all patients remain on their randomized treatment, and the final analysis is based on the endpoint assessed at Tmax.

In other words, this design consists of two or more efficacy endpoints, one assessed at the interim time point and one at the maximum time point, Tmax. This trial can stop early prior to Tmax if the endpoint at the interim time point meets the predefined success criteria for efficacy.

Next slide.

In our experience, many trials fail to provide conclusive evidence of efficacy likely due to inadequate treatment duration. As illustrated in this hypothetical example, a placebo-controlled trial has a fixed randomized treatment duration of 6 months. At 6 months, all patients have the option to receive the test drug in open-label. The efficacy results at 6 months numerically favor the test drug with a p-value of 0.4 for treatment comparison.

The outcome of the patients in the test drug will continue to improve after 6 months, but without a concurrent placebo control after 6 months, this trial fails to provide conclusive evidence of efficacy.

Next slide.

If this trial is designed with a treatment duration adaptation, patients will continue with their randomized treatments for another 6 months because the first 6-month results are not convincing. The trial will have a greater chance of showing significant results at 12 months if the

longer treatment duration produced a larger treatment effect.

Next slide.

The third approach to increase power is using global tests for multiple endpoints. When a test drug is anticipated to have effect on multiple endpoints in a small trial, it is desirable to perform a global test on the multiple endpoints so that one can make a single probability statement about the drug effect.

In this table, we use a hypothetical trial to illustrate the concept of global tests. This trial has two primary endpoints, FVC and 6-minute walk test. When tested separately, both endpoints failed to show a treatment effect at the significance level of 0.05. On the other hand, the two global tests, O'Brien Rank-Sum and Test-Statistics-Sum, produced a p-value less than 0.05 indicating that the drug is efficacious.

Next slide.

Here are some details about these two global tests. The O'Brien Rank-Sum is based on the sum of

the ranks of the data from the multiple endpoints for each patient. Each combines data at the patient level and is typically used for continuous or ordinal endpoints. The Test-Statistics-Sum is based on the test statistic for treatment comparison for each endpoint. It combines test statistics at the endpoint level and is used for all types of endpoints, including binary endpoints and time-to-event endpoints.

Next slide.

As illustrated in our simulation, when a drug has an effect on multiple endpoints, the global tests are more powerful compared to the conventional testing approaches. In this figure, the blue line is the power curve based on the Test-Statistics-Sum, the purple line is based on the O'Brien Rank-Sum, and the black line is the Hochberg method, which is a conventional method commonly used for testing multiple endpoints, and the green line is testing a single endpoint.

As shown in this figure, the power of the global tests are consistently higher than the power

based on the conventional testing approach. For example, for a sample size of 30 per arm, the power of the Test-Statistics-Sum is 15 percent higher compared to the Hochberg method. Compared to the method of testing a single endpoint, the power of the Test-Statistics-Sum is 25 percent higher.

Next slide.

High quality of trial conduct and data collection are essential to the success of a rare disease trial. To obtain quality trial data, the trial sponsor should follow the ICH E6 guidance that covers the principles of good clinical practice.

According to this guidance, trial sponsors should implement and maintain quality assurance and quality control systems to ensure that the trials are conducted and data are collected in compliance with the protocol, good clinical practice, and the applicable regulatory requirements.

Quality control should be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly.

Methods and procedures for outcome assessments should be standardized to reduce noise. This will help to increase statistical power.

For example, in a placebo-controlled trial with a sample size of 35 per arm, we expect a treatment difference of 35 meters in the 6-minute walk test endpoint. If the variability of the outcome is decreased from 60 meters to 54 meters, a decrease of 10 percent, the statistical power can increase from 67 percent to 76 percent, an increase of 13 percent.

To conclude this slide and my presentation overall, I would like to emphasize that trial execution is as important as trial planning. Thank you for your attention.

Session 3 – Questions and Answers

DR. DONOHUE: Thank you again, and thank you to our audience participants for your wonderful questions. We have gotten dozens of them, and I'm going to try to address as many of them as we can in the 15 minutes or so that we have left.

Jack, a couple of really good questions for

you on the dose ranging piece. First up, how does the FDA determine if dedicated dose-finding studies are required before initiating a pivotal clinical trial in a rare disease?

DR. J. WANG: Yes, that's a good question. Thank you, Katie.

When we are in a dedicated dose-ranging study, as you have heard from my presentation, dose-finding and dose-ranging trials are very important for a rare disease program. From a regulatory perspective, though, if the sponsor is asking whether it is required, it is not required by regulation but it's something really needed for your program.

How we determine when a dedicated dose-ranging study is needed, it can depend on many factors. For example, what kind of nonclinical model and efficacy you have and whether you have any healthy subjects' biomarker studies, and whether you have any experience from other relevant disease populations because there are often many drugs developed for many indications. We often see

some sponsors do a rare disease program for an approved drug, so the dose-ranging information from other programs can be helpful.

It also can be dependent on the target patient population. For example, if the sponsor wants to do a rare disease for a pediatric indication, we often want to see some proof of concept and/or dose ranging to make sure there's a direct prospective benefit.

Also, you have heard from the presentation when it's not feasible to do dedicated dose ranging, then you can do an adaptive trial dose-finding study to put dose finding on the confirmatory efficacy trial.

I hope those considerations are helpful for the question.

DR. DONOHUE: Thank you, Jack.

Another question is, how does FDA determine which subpopulation studies are required to support registration in the treatment of a rare disease? I might even broaden that and say, can you comment on when during development do we tend to require the

different clinical pharmacology studies, and why?

DR. J. WANG: Yes. That's also a good question. Actually, most of our IND sponsors often have these kinds of questions in their IND meeting package. For specific drug development programs, the sponsor needs to discuss their IND specifically from studies, what study is needed, and what other approaches, as I mentioned in the presentation.

To give very brief advice, very often dose separation studies label the issue, and it can be conducted as postmarketing studies if the sponsor has their pivotal efficacy and the safety trial already done, and the data is promising, and they are eager to submit their NDA/BLA. Yes, in those situations, organ impairment studies can be done as postmarketing commitment or requirement.

In some situations, we require the sponsor to conduct, for example, an organ impairment study before the pivotal trial. For example, if the sponsor has an indication that it's a liver disease, we certainly want to see how liver impairment, hepatic impairment, affects the PK

before they conduct the pivotal trial; otherwise, we are not able to determine a good dose for their efficacy and safety trial.

Yes, thanks for the question. I hope it was helpful.

DR. DONOHUE: Thank you, Jack.

We had several questions about can you do a seamless design with a gene therapy? Essentially, what do you do with treatments that might have carryover effects? These are good points. The seamless design isn't going to work in all situations. There are going to be some treatments like gene therapies that are sort of one and done, where that's not helpful.

Can you comment on that aspect of when does a seamless design work, when doesn't it, and what might the alternatives be?

DR. J. WANG: Yes, that's also a good question, Katie. As you know, we do not regulate gene therapy in CDER. I think we can look at some other applications in CBER to see their general practice. But in CDER, we do have some similar

therapies like antisense and siRNA.

For those treatments, very often, we need to look at experiences from other drugs of the same class to see other successful stories that we can use a similar approach. Yes, most of the cases will rely heavily on the nonclinical data, and also you need to make sure the trial has a very good monitor for both the efficacy, biomarker, and safety.

I don't think we have a straight answer for those unique cases. I think that it will be very specific for the drug and for the patient population.

DR. DONOHUE: Thank you, Jack.

Now I'm going to send a couple questions to myself. We got some very good questions about flexibility, regulatory precedent, and second generation drug development and what constitutes available therapy. These things are all kind of tied together.

Starting with what constitutes available therapy, does it have to be FDA approved? The

short answer is no. I tend to take a very pragmatic approach to this. If a therapy is still widely available that almost all of the patients are taking it, then it's available therapy, so you've got to deal with that in designing your clinical trial.

It does present challenges. If it's unproven and any potential effect is modest to fair, you might be able to persuade patients not to take it and just stay on a placebo instead, particularly for a shorter trial duration, and that gets into the ethics. If everyone is taking the drug, and if everyone believes strongly that the drug is working, even if it's not FDA approved, you're probably not going to be able to randomize patients to placebo, so you're going to have to think about developing a new therapy as an add-on therapy to that.

So you've got to deal with the reality of the facts on the ground as you're designing your trial in terms of what is going to be ethical and what is going to be acceptable to patients. Those

are key factors.

Some great questions about if you have regulatory flexibility with the first generation drug development program, what does that mean for the second generation drug development program? I'm so glad that this question was posed because I think it's really critical, and it goes right back to when should we accept single-arm trials?

What are the hidden costs? If the FDA approves the first drug for a disease based on a single-arm trial, it makes follow-on drug development really challenging. If you look at drug pipelines for other diseases, most drugs are mildly or modestly effective. Most patients end up needing to take several different medications to manage their disease.

The way those medications get developed often is with what we call noninferiority designs, where you randomize patients to the first gen therapy, and then your new drug that you're developing, and you're trying to show that this new drug is basically as good as the old one; at least

it's no worse.

Now, conventionally these often require four times more patients than that first generation trial showing superiority to placebo, and it also means that you had to have a randomized trial with a placebo arm for that first generation therapy. So in order to do this standard follow-on drug development paradigm, the first gen trial has to be randomized so that you can develop what's called a noninferiority margin in order to show that follow-on drugs are at least as good as the first gen therapy.

So if that first gen therapy gets approved based on a single-arm trial, if there's no randomization, there's no noninferiority margin to inform follow-on drug development. So it can really paint patients into a corner where, yes, they have an approved therapy, but we've now made it incredibly difficult to develop second and third generation therapies for those patients.

So that's one consideration, and it's an important one in thinking about a therapeutic

pipeline for a given patient population.

What can we do about that in terms of the noninferiority designs when that sample size isn't going to be feasible? For a good example, actually I think you could look at the Nexviazyme program. That was a second-generation drug development program that relied on a noninferiority margin, and crucially the first gen trial was randomized, so that might be a good example. But these are really thorny challenges, and they're some of the more interesting scientific questions I deal with. We're all going to have to put our heads together to think of some solutions. Those are some preliminary thoughts on some of those questions that have come in.

I do want to pivot to several questions that came in from a statistical standpoint to ask Yan about.

Yan, if you would turn your camera on, please. When selecting component endpoints in site global testing, how do we make certain that we don't re-measure a small nonclinically important

improvement twice, making the power appear larger? Is there any strategy to ensure that global testing covers a broad spectrum of physiological and clinical changes over the course of the study?

As a theoretical example, measuring walking distance and leg cycling ability to likely assess similar things, but maybe a combination of walking distance and seated arm peddling can capture some seated fitness improvements as well.

I think, essentially, this question is getting at, how do you pick the components of your endpoint? There are other questions about how do you make sure that you're still controlling for type 1 error when you have one of these global endpoints? Then what are the implications for that in terms of labeling?

Those are the three main questions that are coming in about your multicomponent or global hypothesis test.

DR. Y. WANG: Thank you, Katie, for the question. Regarding the first question, I think the question asks which components should be

included in the global test or which endpoints, including the multiple endpoints?

I think this is more a clinical question because it depends on the drug mechanism, mechanism of the drug and the disease indication. We know for LOPD, often you can use both endpoints FVC and 6-minute walk test as the primary endpoint because we believe that the drug likely will work on both endpoints.

It also depends on the property of your drug. For other rare diseases, if we don't know the drug well enough, we are not sure which component will be helpful to include in a global test so we will have more power. I statistically cannot address that question.

The second question, can you repeat again the second question? I know the third one is how you're labeling if the drug has approval. That's the third question. The practice is more to follow the composite endpoints. Say for a composite endpoint, you have the time to event like death, randomized as composite endpoint. If the trial

makes it, you summarize the results, what's the probability of the clinical event by treatment group and the treatment difference? Yet, at the same time, you also look at each individual component.

For the global test, I think we follow the same principle. In the table in one of the slides I showed, you will present the summary statistics for each component endpoint. In terms of the global test, once the drug is approved, we don't need to provide details about the p-value of the global test in the labeling. That's not necessary. Once we make the decision that the drug works, then we just focus on describing the effect size for all the endpoints in the labeling.

I think the second question is about controlling type 1 error rate. That's the same question, applying to composite endpoint. A trial can make it based on composite endpoint and based on global test, but it's not guaranteed which of the component endpoints will show a statistical difference, but that's okay, as long as they don't

show harm on one of the component endpoints.

There's no type 1 error issue here because the global test, it tests a single hypothesis, the null hypothesis that the drug doesn't work for any endpoint. The alternative hypothesis, the drug at least works for one endpoint, so there's no multiplicity issue here when we use the global test.

DR. DONOHUE: Thank you again.

One last question here about, can you use the global hypothesis test for these multicomponent endpoints to address heterogeneity and power optimization?

DR. Y. WANG: Yes. The answer is yes. Actually, I think the global test can be very flexible. The example we use is often like, say, the trial has two primary endpoints, which means every patient has two primary endpoints. The global test can be applied in this situation to account for the heterogeneity of the disease.

A trial can include two types of patients. One patient, say, they walk well, so 6 minutes is

not a good endpoint for this subset of patients, and they only have problems, say FVC. You can have a subgroup of patients that only have one endpoint, the FVC endpoint as the primary endpoint. You can have other patients, and their lung function is ok and works normally, but there 6-minute walk test is not so good.

So you can have two different subpatient populations enter into the same study, but with different endpoints, and the global test can combine the evidence for these two patient populations with two different endpoints together to make a single statistical statement.

DR. DONOHUE: Thank you, Yan, and thanks also to Jack. Thanks to all of my panelists, and also all of the participants for asking such great questions.

I think the key takeaways here are there are a handful of tools in the box that we use for dealing with rare disease drug development over and over and over again. One of the first is seamless design to make sure that we've got dose ranging so

you can use all the same patients in your phase 2, and then move them right into phase 3 and not have to have separate pools of patients.

So those seamless design strategies are really important because as Jack noted, dose ranging is really important. Inadequate dose ranging is often one of the major contributors to failure in rare disease drug development, so anything that makes that more feasible is going to help.

A second strategy is the adaptive duration of the trial by extending the length of the trial as needed. This helps us deal with a lot of the uncertainty around the natural history and how quickly patients are going to progress.

Then as Yan noted, these multicomponent endpoints with a global hypothesis test across all the pieces is another core strategy for improving power, for addressing heterogeneity, and frankly, for also increasing sample size. If you can broaden your enrollment criteria because you can measure benefit across a range of endpoints and

enroll all of the available patients at all available ages, you can increase your power that way, too.

Those are three of our best strategies for dealing with some of the common challenges in rare disease drug development. I thank everyone for your questions, and thank you for having us. Take care.

(Whereupon, at 2:25 p.m., a recess was taken.)

Session 4

Tiina Urv - Moderator

DR. URV: Hi. Welcome back. My name is Tiina Irv, and I'm a program director from the Division of Rare Disease Research Innovation, formerly known as Office of Rare Disease Research, at the National Center for Advancing Translational Sciences at the NIH.

This session that will be next will illustrate the challenges of designing and conducting rare disease clinical trials that are fit for purpose from a regulatory perspective. The

participants in this session are all PIs from the Rare Disease Clinical Research Network or the RDCRN. Our first speaker will be Andrea Gropman. She is a division chief of Neurodevelopment, Pediatrics and Neurogenetics at Children's National Hospital, and she's also one of the principal investigators of the Urea Cycle Disorders Consortium.

Andrea?

Presentation – Andrea Gropman

DR. GROPMAN: Thank you, Tiina, and thank you, everyone, for giving me the opportunity to present. I'm going to be wearing two hats and talk about two distinct challenges in bringing and advancing science from the bedside or the bench to clinical trials for rare disorders.

Next slide, please.

These are my disclosures in terms of my funding and my work as medical and scientific advisory board member.

Next slide.

I'll be talking about drug development in

two classes of disease. One is the urea cycle disorders, shown here on the left, and I'm one of the co-PIs of the Urea Cycle Disease Consortium, and the other is for two rare mitochondrial disorders, LHON, Leber's Hereditary Optic Neuropathy-Plus, and MELAS, which is a mitochondrial encephalopathy, lactic acidosis, and stroke-like episode.

Next slide.

I'll talk about the history of drug development and the Urea Cycle Disorders Consortium, or UCDC, which I'll use as the abbreviation; clinical trial readiness from UCDC in terms of biomarker discovery projects; preclinical studies to inform trial design and how the UCDC expertise helped in development of new therapies for these rare diseases; and how we facilitated a phase 4 study for approval treatment for an even rarer urea cycle disorder.

Next slide.

Urea cycle disorder is shown here, and the role of the urea cycle is the disposal of waste

nitrogen via the conversion of ammonia to urea through a series of enzymatic reactions. A deficiency of an enzyme or a transporter in this pathway, which is responsible for converting ammonia to urea, can result in the accumulation of toxic levels of ammonia, first in the blood, and then, unfortunately, ultimately in the brain, and the resulting encephalopathy from this hyperammonemia can cause death on the one extreme, or more often neurologic impairment.

The long-term management of urea cycle disorders is not very satisfying. It requires a low protein diet with supplementation of essential amino acids and other nutrients that are lacking from that diet; ammonia lowering agents; and an emergency protocol for use because despite the diet and the other medications, these patients are still at risk, or many of them are still at risk, for hyperammonemic episodes.

Next slide.

What are the current treatment options and what is the treatment landscape for urea cycle

disorders beyond the diet? We have at our disposal oral sodium benzoate, which conjugates with glycine and causes excretion of a non-toxic hippuric acid in the urine; sodium phenylbutyrate, sodium phenylacetate, which conjugates with glutamine and allows for excretion of a non-toxic phenyl, acetyl glutamine in the urine; and more recently, glycerol phenylbutyrate, which is a pre-pro drug and allows for conjugation with glutamine and excretion as a non-toxic phenylacetylglutamine in the urine, has a slower release and uptake than sodium phenylbutyrate, sodium phenlyacetate, and we have arginine for infusion.

Next slide.

In addition, there's a very rare urea cycle disorder, NAGs, or N-acetylglutamate synthetase deficiency, which is responsive to a medication called N-carbamyl-glutamate.

Next slide.

Over the course of the last 16 funded years in the RDCRN, we've conducted a number of studies and protocols. The most expansive is our

longitudinal study of urea cycle disorders, from which we were able to leverage data for subsequent clinical trials. For example, we've had randomized clinical trials of low versus high dose arginine in arginosuccinate lyase deficiency, and a number of biomarker studies involving the brain, and ultimately the liver to poise us for participating in clinical trials, as shown here. We've also worked with several pharmaceutical companies for either clinical trials, randomized clinical trials, or a post-surveillance protocol.

Next slide.

These are three of the trials that we've been involved with. One was with Orphan Europe at the time, now Recordati, and this was for a compound, Carbaglu, or N-carbamoylglutamate, for that NAGs deficiency.

The product was a synthetic form of the N-acetylglutamate. Basically, the product was approved in 2010, and we've been involved in conducting the postmarketing surveillance under an RDCRN protocol. We were able to show that the

Carbaglu was effective in a subset of patients, with one of the proximal disorders, carbamoyl phosphate synthetase 1 deficiency, but not ornithine transcarbamylase deficiency.

Then this work was extended. We were able to leverage this and to study this through an R01. That was Dr. Mendel Tuchman, who was able to perform a multisite team of investigators to look at this further, and also to perform post-surveillance marketing. So the involvement of Orphan Europe was supplying the drug and placebo, but the trial was supported by both NIH as well as philanthropic funds.

The next major clinical trial that the UCDC was involved with was the FDA approval of Ravicti, which is glycerol phenylbutyrate. This is the nitrogen binding agent, and we were able to provide de-identified aggregate data from the longitudinal study to inform the clinical trials and basically introduce the UCDC investigators, who would serve as consultants and site PIs.

Then more recently, we've been involved in

an enzyme replacement therapy for arginine deficiency, again providing de-identified data on arginase deficiency patients who were enrolled in the longitudinal study to inform the clinical trial design, and the company now has an active phase 1/2 clinical trial for this arginase enzyme replacement therapy.

Next slide.

With regard to the study for the glycerol phenylbutyrate for urea cycle disorders, this was the study design. We had a phase 2 and a phase 3, originally starting with adults, then bringing the age subsequently down. Because there are ethical issues in treatment of patients with rare disorders, especially if they have a drug that works, really having it as an add-on initially is the way to go. They also do this with epilepsy trials as well, as you can't just take someone off a medication that's been tried and true -- and maybe not totally effective but at least providing some efficacy -- and put them on an unknown.

We looked at both the short- and long-term

effects of ammonia regulations. Initially, we had the patients first on their stable dose, and then add on to the new agent, switching to equivalent dose. This was over a 12-month period, a long-term treatment period. We had 100 individuals, 51 adults and 49 pediatrics, across the multiple sites of our urea cycle consortium. They had monthly visits looking at ammonia and plasma amino acids.

Next slide, please.

We evaluated the 24-hour ammonia regulation as well as long term, and this was published in 2013.

Next slide.

Plasma ammonia has been a standard and acceptable surrogate endpoint for these clinical trials, and a lot of this knowledge came from clinical observations, so looking at what type of biochemical abnormalities presented in patients in the throes of a hyperammonemic crisis; so again, taking information from the bedside to clinical trials using the data from enrolled subjects -- .

Next slide.

-- and also using the longitudinal data to power clinical trials in the UCDC. Many of these slides are from Sandesh Nagamani, who has graciously allowed me to present them today, and this is actually a study with Brendan Lee, who's our next speaker and used to be in our consortium. So really, evaluating sample size for primary neurocognitive outcome endpoints in this condition were powered using data from neuropsychological assessments in the longitudinal study.

Next slide.

Our involvement in phase 4 studies in this very rare disorder, NAGs deficiency, performing the Carbaglu surveillance as part of a UCDC or RDCRN protocol, this was the only surveillance protocol for this particular drug that was approved in 2010, and this effort was led by Nick Ah Mew, who's one of our site PIs.

Next slide.

To date, many of our studies have focused on biomarker identification, so long standing with

neuroimaging, and now more recently with liver; comparative efficacy studies that we've conducted looking at standard of care versus liver transplant; randomized-controlled studies of ammonia lowering agents; and evaluation of novel therapies.

Next slide.

I wanted to contrast that with some more recent experience that I'm embarking on with colleagues at GW. We had the benefit of the urea cycle drug development studies to work with pharmaceutical companies, but now we're back to the academic center.

Two disorders in particular we're interested in are this Leber's-Plus and MELAS, which are both disorders of oxidative phosphorylation in the mitochondria at complex 1. Both of them cause devastating disease for which there is not very effective therapies out there.

Next slide.

MELAS and Leber's-Plus are progressive neurodegenerative disorders. They do share some

similar features but also have very different clinical manifestations. On top of that, even within the disease and within the same family, there may be a broad clinical spectrum of presentation in terms of what the symptoms are and the ages of onset and the severity.

Now, they're both maternally inherited, and pathogenic variants in these two genes affect oxidative phosphorylation. In MELAS, the variants tend to be heteroplasmic, whereas LHON, they may be near homoplasmic levels.

Next slide.

I've had the opportunity and quite gracious to work with this very talented group of researchers who have developed what they call the, Mito-EpiGen Program. They've been doing preclinical work initially with MELAS in fibroblasts to gain insights into the biomedical and pathogenic signature.

Dr. Chiaramello's lab has designed a strategy for using multi-omics in this particular disorder, for which there isn't really an effective

animal model, to look at preclinical effects of drugs.

Next slide.

Using the preclinical work in fibroblasts, we can look at what we already know about the biochemistry of these patients, is that they have dysregulation of complex 1. They have alterations in many bioenergetic pathways such as glycolysis, oxidative phosphorylation, TCA, and fatty acid oxidation as well.

This could possibly be a model for precision medicine and testing various compounds in patients. Also, we know that there's a downregulation of the arginine biosynthesis pathway, which may be important in that there was uncontrolled, basically a clinical observation that arginine may be helpful in patients with MELAS in particular, and this has not really gone through a clinical trial as yet.

Next slide.

But the challenges of clinical trials in academia are many, so funding; responding to multiple review cycles, namely IRB; establishing

clinical trials and material transfer agreements with sponsors and medical centers; finding the resources within your institution; patient recruitment, protected time, and the large amount of associated paperwork.

Next slide.

About a year ago, NCATS came out with an RFA describing the opportunity for a basket clinical trial to evaluate drugs targeting shared molecular etiologies in multiple rare disorders. It's a two-part grant with UG3 and a UH3, comprised of an exploratory and a developmental phase award, which is a cooperative agreement like all U awards are.

Next slide.

The rationale was that, currently, companies and investigators are looking at drugs targeting shared molecular ideologies, but the standard approach in clinical trials has been to focus on one disease at a time, and usually the disease that's picked, even within rare disorders, is one that is less rare than the others.

But as Dr. Donohue said, you really need to

balance the rareness against the scientific rationale. So this approach of picking the more common of the rare results in clinical trials in which only the most common rare diseases exclude patients with the least common diseases, even though the scientific rationale may be stronger in that disease that is of lower prevalence.

Next slide.

Taken from the wording of this RFA, this was proposed as a potential solution to adopt a basket trial approach that's been developed for tissue agnostic oncology drugs for clinical trials of drugs that target molecular defects common to anatomically different cancers, and to apply this to rare disease.

Next slide.

There are variations on this theme. The basket trial tests one or more drugs on one or more diseases. There's also an umbrella trial, which is slightly different and tests one drug on different mutations but in the same disease. Then of course, you've all heard about N of 1 trials, where you

basically have a drug developed for one particular patient who has a particular DNA variant. These can involve multi-omics, data mining, and ultimately may provide information about clinical decision making.

Next slide.

The UG3 phase basically is the translational, and then if that is successful, there's transition to the UH3 phase. The UG3 phase will depend upon the maturity of the project at entry, and then those projects that have met specific milestones can then go on and be eligible for transition to the UH3 phase, which will support a small clinical trial involving at least two different diseases. This is a cooperative agreement, so along the way, NIH program staff are involved in the planning and execution of the projects.

Next slide.

Conducting clinical trials in academia, especially now with a basket trial approach for rare disease, which has never been tried, is

certainly going to be complex in the design and patient access. How do we access rare disease patients? Well, luckily there's RDCRN for mitochondrial disorders and patient advocacy groups.

Other things that need to be considered are what would be the cost of the budget to conduct this; what are the roles of the staff and responsibilities; and how do we establish governance and oversight?

Next slide.

For those of us who have not done this outside of academia, navigating the FDA website can be difficult, especially since a lot of our hospitals use encrypted email, and just looking around at the site can be arduous, so I'm looking forward to the talk tomorrow about how to do that.

Next slide.

We're going to focus on two ultra-rare diseases, MELAS and LHON-Plus. These are studied by the RDCRN, the NAMDC, which is North American Mitochondrial Disease Consortium; again, the

challenge to recruit these patients, however, understanding that these patients don't have access to effective treatments; repurposing a drug that's been previously used in solid organ tumors and being able to reactivate studies into new patient populations for new indications. These are the challenges and the goals of this project, and basically, the patients share a common etiology with complex 1 deficiency and have a chronic energy or ATP deficit.

Next slide.

Some of the issues that may come up when one tries to embark on a clinical trial are what's our preclinical data? Well, we don't have an animal model, but we have to think of new ways around us because not every rare disease has an adequate animal model. But we have a fibroblast, so will studies establishing the preclinical efficacy of different pharmacologic compounds be enough for this proof of concepts in these two new populations?

Next slide.

But there has been published literature using the compound that we're interested in, in embryonic cortical neurons, hippocampal neurons, and other neuronal cell lines.

Next slide.

So here we go, embarking on uncharted territory; so really need the advice and guidance of the FDA going forward and need to think about new ways to approach the study design, and the retention of patients, and also measuring the efficacy of these drugs, as have been previously discussed.

I wanted to acknowledge all the clinical and research partners. Dr. Nagamani is one of the co-PIs of the UCDC; along with Cindy LeMons, who's the executive director of the National Urea Cycle Disease Foundation, which is the patient advocacy group; all the UCDC PIs, patients, and the families; and Dr. Chiaramello and her lab over at GW, and I thank you for your attention.

DR. URV: Thank you very much, Dr. Gropman. That was really wonderful.

I want to invite the audience to please send in any questions they have that we'll take at the end of the presentations. You're able to do so from your screen.

Next up, we have Dr. Brendan Lee. Brendan is the chair of Molecular and Human Genetics at Baylor College of Medicine. He's also the principal investigator of the Brittle Bone Disease Consortium.

Brendan, take it away.

Presentation – Brendan Lee

DR. B. LEE: Thank you, Tiina, for this invitation. It's been a great meeting, and I hope to share with you some of the work we've been doing in the Brittle Bone Disorders Consortium. I think it illustrates very nicely many of the points that have been touched on this morning and this afternoon. Our benefits, we also suffer from being a common rare disease, so to speak.

Next slide, please.

These are my disclosures.

Next slide.

The Brittle Bone Disorders Consortium covers a host of diseases, which originally were termed "osteogenesis imperfecta." This is one of the three heritable disorders of tissue that Victor McKusick described in the '50s in his treatise. As such, I think it is characterized by the variable expressivity that we see in many of the genetic disorders affecting connective tissue.

As some of you may know, the main features have been low bone mass and brittleness of bones, something we focus on clinically and in trials, and their associated deformities and, hence, fractures. But it is important to keep in mind -- and this is relevant in considering composite endpoints -- that this is a connective tissue disorder with extraskeletal manifestations, including in dentition; in hearing; in lungs, and ligaments, and tendons, for example.

As you can see in the x-rays, though, there is a real variation in terms of severity, and heterogeneity of clinical presentation is the hallmark with features of the condition, which is

incompatible with life, all the way to some minor risk of fracture that one may not even know they have this condition.

Next slide, please.

I'm going to sort of start with the end in terms of what are the lessons that we've learned in terms of translation of rare bone diseases, especially the Brittle Bone Disorders Consortium, have taught us.

The first is that, actually, the structural functions of the mouse and human skeleton has been remarkably conserved through evolution, and this has supported strong clinical translation, not only in rare disease, but in common diseases, as you'll see. And this has impacted in terms of how our natural histories have really progressed.

Now, the clinical endpoints, however, in these rare disorders have suffered from enormous clinical heterogeneity, and this is first initially reflected in locus and allelic heterogeneity, so now many genes that contribute to the phenotype, as well as many mutations in genes that contribute to

heterogeneity; but also with now what is functional standard of care, where drug treatments have actually impacted the natural history, and this was also alluded to in how it impacts the development of actual approved drugs.

There's no question that a theme throughout has been the early partnership and collaboration between NIH, industry, patient advocacy groups, and academic researchers are key to identifying unmet and sometimes unknown needs; accelerating research; performing the natural history studies which we hope to power the endpoints that are coming for FDA-approved studies; and accelerating early-phase trials, as you can see from brittle bones consortia; also leveraging the human experience, both in terms of dosing, dose response, and toxicity for potentially applications, or newer applications, to drugs that have been studied in the context of repurposing, even if it's repurposing non-previously approved drugs.

Next slide, please.

The statement that there's been great

conservation -- and the mouse has been a superb translational model for structural targets of treatment -- I think it's evidenced by this; that there have been really superb and many successful drugs that have been approved for the treatment of a common disease, osteoporosis, in terms of how it impacts bone formation by the osteoblasts, shown on the left, and bone resorption, by the osteoclasts, shown on the right, and really changing this balance to improve and increase bone content.

I think the best example of these have been the bisphosphonates, shown on the right, drugs that inhibit the function of osteoclasts, moving forward to drugs that, in fact, target signaling; drugs that block rank-ligand signaling to the osteoclasts, for example, denosumab, an antibody that is very effective on the anti-resorptive front; and similarly on the anabolic front, forms of parathyroid hormones, which in pulsatile fashion stimulates bone formation; and most recently powered by rare disease genetics, mutations of sclerostin or the development of antibodies that

block sclerostin to increase bone mass by blocking Wnt signaling.

Now, this slide is important because the experience of pamidronate and the safety margin of this drug led its use to be developed in the late '90s by Francesco, Herrera, and others, and this has now become a de facto standard of care, especially in pediatric OI, and has impacted the natural history of this disease, and in fact, how we even consider performing controlled clinical trials for approval.

Next slide, please.

This slide demonstrates one of the challenges I pointed to. There are now many, many types of, quote, "OI," which contribute to the spectrum of the Brittle Bone Disorders Consortium, and while the majority of the genes include genes that involve structure and post-translational modification of collagen, there is enormous heterogeneity with its underlying mechanistic heterogeneity and, hence, really are beginning to lead us to focus on genotype specific groups when

we think about targeting mechanistic-based therapies.

Next slide, please.

Bisphosphonate is, in fact, an accepted de facto standard of care, but it is not FDA approved, as is often the case in rare diseases. Its use has been studied in multiple trials, but this is an excellent review by Bob Steiner and others in terms of bisphosphonate therapy in OI.

As you can see, it is a standard of care, especially in children with severe OI. There have been multiple trials that have been performed, and I take quotes from the conclusions. "It is unclear whether oral or intravenous bisphosphonate treatment consistently decreases fractures, though multiple studies report this independently, and no studies report an increased fracture rate with treatment." So it doesn't certainly harm patients in terms of fracture rate, but clearly it's been variable whether a clinically important endpoint, i.e., fracture reduction has been met, and there are many reasons for this.

At the end, "The studies included do not show bisphosphonates conclusively improve clinical status in people with OI." That's a pretty daunting statement when you think about the fact that this is de facto standard of care; even though I think clinicians and patients would report the anecdotally enormous benefit.

I think this is, again, reflective of the enormous heterogeneity in this population, where you can study a patient with OI, and they may suffer hundreds of fractures, but at the same time, another patient, depending upon where they are in their life -- so it's not only the genotype, but also the impact of environment, the life course, and their age where they may have had only one or two fractures in the past recent years. You can imagine how the distribution of such events clinically can totally confound powering a study when you're looking at fracture endpoints.

Next slide, please.

It's because of this that the Brittle Bone Disorders Consortium was formed, and it is at now

over 14 clinical sites across North America to try to begin to document the natural history of this, and now, really, the natural history of this in the age of bisphosphonate use and how that can inform many of the things that we've been talking about today.

Next slide, please.

What have we achieved to date and as a take-home message? We have the largest cohort of patients with osteogenesis imperfecta, following now for the past eight years. There are close to a thousand such individuals. In the studies that we've performed, we've actually identified clinical signals not previously appreciated or studied; for example, the risk of postpartum hemorrhage, impact of pain, anxiety, and other neuropsychological endpoints.

Importantly, and not surprisingly, we were able to quantify the effect sizes of different subtypes of OI, and this really helps to begin to address the variable expressivity as it confounds sample sizes in considering powering trials. This

includes multiple measures such as growth, which is a major aspect of OI, especially the severe type; pulmonary function, a really confounding measure; mobility, including measures which have been accepted by the FDA like the 6-minute walk test; hearing loss; and increasingly important patient-reported outcomes that impact quality of life.

What is clear from these studies is that these are truly, as Victor McKusick himself described years ago, broad connective tissues that target elements beyond bone, in which I think inform us to begin to think about composite endpoints to increase the power of potential studies.

The consortium and the data generated, as actually very nicely demonstrated by Andrea in the previous talk, is a basis for academic, industry, and advocacy partners to come together to power and design clinical trials. There have been some good examples of this. Actually, a study performed and done by investigators within the BBDC on an

anti-TGF beta strategy has now been moved forward for further development by Sanofi.

Then again, as a model for engaging academic investigators and multiple centers, industry sponsored studies focused on the agonist, sestrusumab, being referred by another company partnership, Mereo and Ultragenix.

Next slide, please.

This is a study which I think illustrates both the power of the preclinical model in terms of translating not just efficacy potentially, but also dose finding in the preclinical model to the clinical scenario. This I think spans a spectrum in rare disease, and while they're completely absent preclinical models as in mitochondrial disease that Andrea touched on, and then on the opposite end of the spectrum, we are blessed with a really powerful preclinical model in terms of structural components of the skeleton.

Here, we had shown several years ago that an increase in TGF-beta signaling in bone was, in fact, a common mechanism in multiple forms of OI

preclinically that impacted either the structure or the post-translational modification of collagen, as shown in the top; and that by blocking TGF-beta, one could effectively restore bone mass and bone strength, as shown in the micro CT image on the top, on the right.

Now, what is important is that this mechanism is reflective of the broad connective tissue disease because, in fact, the pulmonary disease that we see as an altered alveolarization of the lung, shown on the left -- wild-type in the middle, and model recessive OI, and then a partial rescue with ID11 -- really extended beyond the skeleton.

Next slide, please.

Within the context of the BBDC, another very important, I think, lesson is can we then validate preclinical findings, such as what I showed you, in human tissues? This was an example where leveraging large consortia, we're able to obtain tissues, bone tissues, from OI patients, as well as control subjects, and show -- using a multi-omic

analysis, that whether you look at histological features, as shown in the top middle and top left where you see osteocyte density features of OI, or RNA sequencing analysis on the top right, where it showed the increase in TGF-beta signaling that we saw in the preclinical models, and ultimately on the protein level, whether by Western blot analysis or reverse-phase protein array on the bottom -- that in fact, again, in the human scenario, there was increased TGF-beta signaling, again, correlating human and mouse pathologies.

Next slide, please.

This then drove us, in fact, to perform a single-dose study, looking at the safety of fresolimumab, a pan-anti-TGF beta antibody, that had been studied by, first, Genzyme, and subsequently Sanofi, in the context of other diseases such as cancer and sclerotic diseases.

We took advantage of that human experience to, in fact, repurpose this study drug to osteogenesis imperfecta. And in fact using again the previous industry experience in terms of dose

context but modified for the pharmacodynamics that we would expect for bone remodeling, we actually studied the drug over a prolonged period of time after a single dose, a dose for 1 and 4 milligrams, and saw biomarker changes, shown below, in terms of osteocalcin and C-telopeptide, the pro-collagen one and pro-peptide, which are markers of bone turnover for resorption and formation, respectively.

In fact, we saw a very strong dose response, which was consistent with the mechanistic data because, in fact, the features of mouse, as well as human OI bone, is a high turnover disease where formation and resorption are uncoupled. In fact, this suggests that that turnover, the sort of ineffective high bone turnover, was potentially corrected in this cohort.

Next side, please.

Now what is interesting, though, is in these even few subjects, we began to see what the preclinical models also predicted. If you look at the top slide in the range from mild, to moderate, to severe OI, you see a listing of both

types -- IV, VII and III -- as well as mouse models that were studied.

In fact, the mouse model, by us and as well by other groups, had shown that there was increased TGF-beta signaling in all these models. But at the doses that we used to correct the bone mass, we only saw a robust correction at the moderate model under the spectrum, and at the most severe end of the spectrum, including this case, the JRQ model, which is a severe connective tissue disease model, there was insufficient TGF-beta at the doses we used in the other models to actually lead to correction of the phenotype.

In fact, that's sort of what we saw in terms of phenocopy and what we see in the human patients. We see a robust increase in bone mass, which is quite significant, given the context of how we know osteoporosis drugs work in general, that at 3 and 6 months, in these models, the model form of OI, kind of IV, but as we moved to some of the more severe forms, we saw really no significant effect, and maybe even a decrease, albeit, again, relevant

to some of the points brought forth earlier in this small sample size, that this may have been confounded by clinical events like fracture and immobility, given the more severe phenotype.

But irrespective, I think this underscores a couple of key points, that robust preclinical models may predict not only potential efficacy but also dose response when we start thinking about the translation in the human context. Based on these studies now, in fact Sanofi's moving forward with this trial, thinking about, in fact, exactly the type of patients and the genotypes that we'll be studying in subsequent phases.

Next slide, please.

That mechanism in terms of the translation actually can also inform clinical trial data that were previously unexplained. This is one of the largest clinical trials that we had performed, looking at an anabolic that was already FDA approved at the time for osteoporosis, teriparatide, in adults with OI.

We saw this differential effect in mild OI,

on the left, versus more severe OI, on the right. Interestingly, going in the reverse scenario in terms of modeling the human scenario with the mouse data to try to explain the clinical effect, you can see in the next slide what we found was that, in fact, the reason we think that there was a lack of efficacy in the more severe models of PTH was due to the increase in TGF-beta, because it had been shown in cell studies by others that TGF-beta can stimulate PTH receptor insensitivity.

In this animal modeling of that context, you can see inhibition of TGF-beta. Using both subtherapeutic doses of PTH and 1D11, we had a synergistic effect causing an actual complete rescue of the bone mass phenotype, again underscoring this strong bidirectional translation in the mouse versus the human data.

Next slide, please.

Another important element that I think leads us to begin to think about composite endpoints has been, in fact, our ability, using this large cohort, to stratify clinical features like

mobility. In this study by Karen Kruger and her colleagues from our consortium, they were able to begin to quantify the 6-minute walk test based on the clinical classification of OI, types I, III, IV, for example, as well as an additional type V, which can be common in certain populations. You can see how, in fact, especially in the more severe type III, that it may be an effective use in terms of as a potential endpoint.

Next slide, please.

Another area we're really beginning to focus on has been quality of life, and in this case, a pediatric measure of mobility, both upper extremity, physical function, and transfer and basic mobility. And by again incorporating this into a large natural history study, we're able to begin to obtain data to really define the endpoints, in such patient-reported outcomes and observer-reported outcomes, on how to begin to power studies, whether they are two-group comparisons versus a crossover type design, that was talked about previously.

You can see the kinds of numbers that would be required, again, underscoring that many of the trials that have been done to date in the context of bisphosphonates, which, again, I pointed to in the Cochrane review, were significantly underpowered when you think about endpoints like this type of quality-of-life measure.

Next slide, please.

So really, we can begin to do this not only in terms of measures that are specific to areas of the instrument, but also, again, with the different clinical severities; so type I, type III, and type IV, again, using in this case in adults, with an adult tool, the SF-12, a brief version of the SF-12, we're able to, again, calculate the potential sample sizes for crossover versus parallel design. You can see, again, the potential dramatic numbers that might be needed, depending upon the clinical types that are being focused on.

Next slide, please.

Another point I would like to touch on is that biomarkers will potentially be very important.

In fact, biomarkers have been shown to be effective in the generic, quote, "physiological states," and one excellent example of this is a type X collagen biomarker from the growth plate, and was published previously to be an outstanding marker for linear growth, in children especially.

Again, taking advantage of our consortium, we performed and asked whether we could use this as a biomarker for growth. What we found was, in fact, the effects were quite opposite; that in especially the shortest patients, shown on the right, type III and IV, that this biomarker can actually be distributed widely and even could be increased, given that these were the shortest patients. Almost in reverse correlation, that could be seen in OI patients, again, underscoring that growth plate dysfunction can affect biomarkers that previously have been studied to be effective surrogates.

Next slide, please.

To end, I think that we have begun to leverage the BBDC infrastructure and the expertise

in the community. I think the industry partnerships to accelerate downstream studies is an example. A good example of that has been the collaboration with Sanofi, but also industry engagement of investigators broadly, as Ultragenix and Mereo with anti-sclerostin in OI.

In all cases, natural history and longitudinal data are really beginning to inform clinical trial design and sample sizes, and then ultimately, expanding patient advocacy networks to increase capacity will be the key. I've not had time to touch on this, but PCORI work at our consortium, as well as work by our tag partner, the Osteogenesis Imperfecta Foundation with the Rare Bone Disease Alliance, is increasing and expanding these lessons throughout.

Next slide, please.

I will end there with the acknowledgements of the many team members that have contributed to this. Thank you.

DR. URV: Thank you, Dr. Lee. That was truly wonderful.

Next, we will move onto Matthias Kretzler. Dr. Kretzler is a professor of internal medicine, and he's also a research professor of computational medicine and biology. He is also the principal investigator of the Nephrotic Syndrome Study Network or NEPTUNE.

Take it away, Matthias.

Presentation – Matthias Kretzler

DR. KRETZLER: Tiina, thanks a lot for the introduction, and thanks a lot for a fascinating symposium, where I think we are really getting at the heart of some of the key impediments that slow-poke us down in the rare disease community. One of the key features, certainly, we experience in our disease domains, and what you also heard from Brendan and Andrea already, is the heterogeneity of what presents syndromic diseases to us clinicians.

Next slide.

You can see my disclosures all available on this, my employment with the University of Michigan.

Next slide.

I would like to use specific cases in our RDCRN Nephrotic Syndrome Network of Rare Glomerular Disease, to delineate a strategy, which hopefully will be applicable to diseases of interest to you as well, and how we can move from syndromic classes to mechanistic disease categories, really, using the incredible advances in translational sciences we are witnessing right now.

In our diseases, in the nephrotic syndrome field, is a syndromic disease classification that really brings people together who suffer from glomerular filtration barrier failure, heavy proteinuria, general [indiscernible] stage, and loss of kidney function. But as you have heard by the speakers beforehand, this is a highly heterogeneous disease. We know by now that there are more than 65 different monogenetic lesions and different genes that can cause a disease, and the series of environmental exposures can also lead to loss of kidney function. They are highly variable along the same lines as you heard and familiar.

It's the same lesions, and we see differences in manifestation from clinically silent proteinuria to rapid loss of kidney function in childhood.

So how can we get a handle on that heterogeneity? Here, we have the opportunity as nephrologists, that we do actually obtain, as part of the diagnostic workup of our patients, fine-needle percutaneous kidney biopsies for histological diagnosis, and that gives us, obviously, a window to define the structural damage patterns present at the time in the patient's history at a biopsy visit.

We also can use the emerging molecular strategies to define the molecular stage in a cell and tissue context-specific manner of a given patient at the given time. In addition, in kidney diseases, we have the special advantage that we can get liquid biopsies. We can get urine samples that carries cells, molecules, metabolites, proteins from the affected nephrons into the urine, and are readily available then for biopsies.

And over the last six years, we were very

fortunate that cell biologists have developed important stem cell derived kidney organoids as excellent patient and individual specific model systems of the alterations of the glomerular filtration barrier.

Next slide.

With this approach, we now can generate deep clinical phenotypes, and in our cohort we capture over 1100 of those patients with the structural patterns of the disease, and then to continue genetic and genomic disease pathophysiology to define cross-cutting disease mechanisms if we have multiscalar data integration platform in place to do that around our prospective cohort study --

Next slide.

-- so that we can actually identify the different outcomes in prospectively ascertained patient cohorts. We can link these outcomes to the determinants at baseline and see which of these are good and poor, and then obviously mine those patients with poor outcomes, what are the underlying molecular events, and bring them to

targeted therapies.

Most excitingly over the last six years, we were able to leverage particularly biofluids of urine-based assays. We actually developed patient-level activity assessment of the molecular mechanisms putting their nephrons at risk, and thereby on an individual patient level can assign a disease activity and the given time, and then bring these patients to the respective trials.

Next slide.

This really is a philosophy which we envisioned in the NEPTUNE study funded by the NIH now for 13 years. From the get-go, we take these observational cohort studies to functionally define our diseases for improved mechanistic disease stratifications so that we can have an expert panel categorize patients, and bring those patients to the targeted therapies; so we break the conundrum that we had multitudes of clinical trials in our space failing, despite the fact that we know that some of these compounds were active, but only in a small subsegment of the patients.

Next slide.

With this philosophy in place, we have established similarity like the other rare disease networks you saw today, a comprehensive network across North America, which bring these people to studies as early as possible in their disease course.

Next slide.

With this, we have established now enriched partnerships from patients, natural kidney donors, who were actually instrumental in getting the network initiated in the first place. Ancillary projects and data sharing tools are available for studies inside the U.S. and with our global research partners around the globe, and very critically, for all translational and clinical projects, you have heard today, very robust public/private partnerships governed by the framework from the National Institutes of Health for our federally funded cohort studies.

Next slide.

This approach, we now have established from

over 700 patients active in the study with a framework of knowledge around the diseases, so that we can get those syndromic diseases and use information from cross-sectional demographics to whole genome sequencing and urine single cell based RNA profiling approaches to define different disease strata in patient populations.

Next slide.

We are bringing that information together into what we refer to as the NEPTUNE Knowledge Network, where clinical morphological and molecular information is brought together. It's searchable because it is the tranSMART data platform for access from our ancillary study investigators from public and private entities, and then really follows three main questions our patient participants ask us from the get-go, where is my disease coming from; where is it going to, and what therapeutic options we have available?

Next slide.

With this approach, we have over 180 ancillary studies by the international

glomerular disease community available, leveraging different aspects from our cohort studies, and conversely bringing them the insight from our studies on clinical samples, data generations, back to our data sharing instruments to drive our discovery instruments forward.

Next slide.

I would like to give you one example relevant for the disease heterogeneity, where we use the multiscalar data integration approach to define mechanistic subgroups and bring them now to targeted therapies.

Next slide.

This study started off using the gene expression signatures, which we have generated from microdissected nephron, segments out of the kidney biopsies. There's a NEPTUNE cohort. Here you see the subcohort, which is syndromically classified for FSGS and minimal change disease. And yes, you can see out of these gene expression profiling by RNA-Seq, we get three main concerns as cluster groups defined T3, T2, and T1.

We then leverage --

Next slide.

-- and we have a sister cohort in place in Europe, the ERCB, using the same procurement strategies and generated identical data, and it's the same analytical platform. We identified three subgroups there as well.

The next slide.

Our sister network, the H3CKD Africa network from sub-Saharan Africa, we're indeed generating similar subclasses --

Next slide.

-- and by carefully evaluating our data sets, we could show that, indeed, the signatures between North America and Europe and North America and African sub-Saharan data sets were tightly correlated, showing that, indeed, what we are capturing is a robust signal.

As you can see on the left lower panel, our conventional FSGS and minimal change diseases were actually contributing to each of these three clusters, confirming our initial hunch that, yes,

these were syndromic and not mechanistically defined studies.

The beauty of the expression-based classification of patients is that you can look on this --

Next slide.

-- and you can actually ask what is different between cluster 3 and cluster 1 and 2, for example. And in this specific instance, using different bioinformatic data mining strategies with network analysis and upstream regulators, we identified that in this specific setting, the cluster 3 patients were significantly different from cluster 1 and 2, mainly due to TNF-driven differential regulation off the kidney tissue in the expression profiling studies.

Next slide.

That got us very excited because our study teams on the experimental trial side already had tested the TNF inhibitor on adalimumab, the Nephron 2 trial and the NEPTUNE framework, and had to stop the study due to futility because only

20 percent of the patients responded with the treatments without an ability to increase stratified patients for targeted therapies at that time.

Next slide.

We therefore developed, in the bioinformatic core facility, out of our expression data sets the TNF activation score. You saw these regulatory hierarchies, so you can ask which transcripts are known to be TNF dependent in their activation state, and then we took these expression levels of these TNF-dependent transcripts to identify on the patient level the activity of the pathway in the kidney tissue.

In these waterfall plots across North American, European, and the African cohorts, you indeed can see a high heterogeneity of the TNF activation score across the study participants with the cluster 3 patients showing the highest activity scores present. Well, that's a good starting point, so we could at this time now enter a study to obtain tissue biopsies, profile, and then bring

patients to targeted therapies.

Next slide.

However, the group asked can we do more? Can we identify where these TNF signals are coming from and develop non-invasive surrogates of those? Here, we take advantage of the fact that we now can assess transcripts in the cell-type specific manner in a single nuc RNA sequencing data sets of our hierarchical --

Next slide.

-- NEPTUNE biopsies. We were able to identify several of the downstream transcription targets of the TNF pathways. And as you can see in these bubble plots, among the panels of cells from podocytes to proximal tubular cells, the TNF activation low in blue and TNF activation high in red, the activation is actually taking place across many different similar compartments, so an intrinsic activation state of the kidney and not just of infiltrating immune cells.

Next slide.

With this, we now were able to ask, A, do we

have an adequate model of this ubiquitous activation of kidney under stress with TNF precedent here? We took advantage of our participation in the NCATS kidney on a chip and Trial on a Chip effort to test if we can use our kidney organoids as a model system for TNF activation.

Next slide.

And indeed in the organoid system, we can show that it's the same TNF activation score transcriptionally based, which works in human biopsies, and showed beautiful dose and time responses to TNF stimulation of the kidney organoids in a dish.

On the right side, you can see that, in addition, we not only saw robust activation of the transcriptional readouts, but supported and coded by these transcripts were also determined in the organoid supernatant. I can get indeed some of these parameters might be capturable in a non-invasive manner.

Next slide.

With this, we evaluated, similar to the in vivo state of the kidney biopsies, a similar contribution. And similar to the kidney tissue in the patients, in the kidneys on a dish we saw also very robust activation of the downstream transcriptional activation surrogates of the TNF pathways, interstitial tubular cells, and glomerular filtration cells and podocytes.

Next slide.

With this, everything enhanced, a biomarker core facility of Neptune 2, to the right, dove into the existing proteomic data sets we had on file from our participants, and now correlated the blood and urine proteome signatures for the downstream TNF activation surrogates with the intrarenal transcripts.

This you can see among a panel of known TNF-dependent transcripts, CCL2, uMCP-1, and TIMP1, and showed tight correlation between tissue and urine normalized for urine creatinine and allowed, actually now in a non-invasive manner, to assess the intrarenal tissue activation score.

Next slide.

With this, it is now possible, on an individual patient level, dynamically to measure the TNF activation inside the kidney in a given patient at a given time point, and then compare that patient with the existing NEPTUNE population, and map the activity state of the patient among a spectrum of glomerular diseases already on that cohort.

Next slide.

With this approach, we now return back in the experimental therapeutics working group in the RDCRN. NEPTUNE at right initiated a phase 2 proof-of-concept study, where now we use the TEB, the target engagement biomarker, assays to bring the right patients to the TNF inhibitions, and then follow them throughout the TNF exposure to see if, A, the biomarker, and B, the outcome proteinuria is responsive to the intervention.

Next slide.

This was an example of how one can use, in our specific instance, tissue level but

potentially, although non-invasive, surrogates to map a specific pathway activity.

Next slide.

We have seen in our field excitingly, finally, the influx of the reality of potential molecular mechanisms targeted by the network. And one of the key questions now is, as we see multiple agents being called to these heterogeneous diseases, can we develop a strategy to bring the right patients to the right trials, at the right time? That's a philosophy --

Next slide.

-- which we are pursuing with the NEPTUNE Match approach, where we take our knowledge network, we define non-invasive surrogate -- as I have shown you for the TNF inhibition -- for the clinical trials that are being called to our patients with a rare disease.

We profile these patients on the clinical side for the activation state of devised molecules, potential surrogates for target activation in the trials, and then bring these patients to the

various trials of the independently executed clinical trials by our NEPTUNE Match private partners to undergo the clinical trial exposure.

At the end of the trial, patients return their outcomes back to our predicted target activation. We can see if this stratification approach indeed enriches for outcomes and gives the expected power and frequency.

Next slide.

This is a novel concept, at least for our rare disease space. Obviously, in oncology there are precedents of how to execute that. We have developed a rigorous training protocol for our network to transmit that information robustly to map, measure, and report our findings to study participants and clinician investigators, and then to have robust statistical models in place with the retrospective assessments of kidney health outcomes.

Next slide.

With this I would like to wrap up. I hope I have given you an overview of how integration of

multiscalar data sets in heterogeneous diseases can help you to identify a subgroup of patients of molecular pathways, many of which cut across our conventional disease categories to bring the right people to the right trial, at the right time, and we see the Clinical Trials.gov number of -- several of the trials who are active in that framework as we speak in the NEPTUNE framework.

Next slide.

This has all --

Next slide.

-- not been possible without the long-term support from the NIH, from the patient interest groups, and NEPHURE Kidney International.

Next slide.

We have a lively rare disease community cutting across many different knowledge domains, interest groups, and continents --

And final slide.

-- to a very dedicated team here in Michigan who makes all this work fun, even in times of significant challenges to all of us. Thank you for

your attention.

Session 4 – Questions and Answers

DR. URV: Thank you so much. That was wonderful, Dr. Kretzler.

Now we have time for a few questions. Feel free to submit any questions you might have at this time. I have a couple here for you all. The first one is for Dr. Gropman, and the question is, why would basket trials allow drugs to be approved more quickly?

Dr. Gropman, what do you think about how basket trials could speed up the whole pace of trials in drug discovery treatment?

DR. GROPMAN: Sure. I think some of the reasons that come to mind would be you're looking at more than one disorder at the same time, so cutting down on the cost and the time.

If you have multiple arms representing the multiple disorders that have both shared and divergent endpoints, using that aggregate data with fewer subjects and less time in the interim analysis could potentially lead to a quicker

approval of these types of study designs using the basket trial, the statistical power with less subjects, and also the fact that the traditional way to do clinical studies was to look at one compound and one disorder, do that trial, then go back and look at another disorder with that same compound; so time essence by enrolling multiple arms, I believe.

DR. URV: Terrific. Thank you so much.

We have a second question for Dr. Lee.

Could tissue engineering be an option in the treatment of OI?

DR. B. LEE: That's an excellent question and I think could be approached from two contexts. One is in the context of translation, clinical translation, and preclinical translation, and then the second from a clinical efficacy perspective.

I'll take the first one. Broadly thinking, I think tissue engineering approaches, an example of the preclinical space would be what actually Matthias touched on and what NCATS has supported in terms of tissue on a chip.

I think one potential, which has not been exploited in the connective tissue space, is to actually model on a chip abnormal matrix by putting, for example, OI cells onto that matrix. That would be actually very powerful in terms of screening both biologics and small molecules on impacts on matrix directly.

That's one area that we as a field have not tackled. We focused on modifying the cellular components, as I touched on in our work, but it's been hard to tackle the qualitative issue of that normal matrix.

I think in the clinical space of tissue engineering, in terms of thinking about whether we can engineer tissues with cell therapy, for example, either artificial matrix, or matrices, there's no question that's in play in the targeted tissue repair domain.

For example, in these more generalized connective tissue diseases, you can impact, for example, fractures that occur and/or joint disease, and there is an absolute application in a more

targeted tissue engineering application, and that of course is still limited by a host of other different regulatory rules around that.

But I would say that's going to be an important component of all genetic diseases and rare diseases, where there's a degenerative component where you lose a tissue and it's not something you can replace easily in the context of connective tissue cartilage, for example. Once you lose it, it's gone. So I think that that aspect of tissue engineering for it there will be critical.

I think systemic treatment is our very high bar, partly because of just targeting and getting the tissue in the cells that make that tissue throughout the whole body. So I think more systemic treatments will be probably the highest bar and perhaps lowest likelihood at this point.

DR. URV: Okay. Dr. Lee, we have one more question for you.

With multiple candidates in the pipeline for OI, how will future companies be able to recruit patients for the disease?

DR. B. LEE: That's an excellent question, and this I think was hopefully -- at least my belief -- alluded to in the talk that Matthias gave. I think the approach previously has been recruit as many people as possible to try and cover for the heterogeneity. I think that, actually, recruiting fewer patients, but more homogeneous patients, whether it is by molecularly stratifying them, clinically stratifying them, both will be important.

I think we touched on that a little bit in our consortium. I think if you look at even the bisphosphonate experiences, the few trials which did reach an endpoint in terms of fracture were, not surprisingly, the ones which had the more homogeneous clinical populations.

So I do think, hopefully, companies, as well as investigators, in general, will begin to really stratify this in terms of potentially heterogeneity, or getting towards more homogeneity, and perhaps also stratifying response, as they are more mechanistically targeted therapies.

As I pointed to, the most severe patients didn't seem to respond as well to the doses of TGF-beta. Well, one could approach that by saying, well, there's more in TGF-beta, and we need to up the therapy, and that's certainly one possibility. But another is that there could be another mechanism that's dominating that group and, hence, targeting a therapy for that group, specifically in a true genotype-specific fashion, would be the answer. So I think there's still a lot of room to play in the future.

DR. URV: Okay.

Dr. Kretzler, could you expand on that from the NEPTUNE perspective as well?

DR. KRETZLER: Yes, Brendan, I think this is absolutely on target. This is why the networks and the cohort studies can become so powerful, because on one hand, that prospectively can define what subsegments in your populations are present and have reached disease subtype present in play; what is the expected trajectory of these disease subtypes, the outcomes, and their response to

current exposures.

Then use that information, the genetically associations and potentially invasive or non-invasive surrogates to stratify your patient populations going forward, and that then starts to scale. If you have multiple agents coming into the domain, you can identify which segment of your population is most beneficial.

And that might not be a scalable solution if you are one molecule or one trial strategy, but if you bring a community together where you now have multiple efficacies together, then there's a strong scientific and I think also a strong economic role in collaborating along those platforms in an intelligent basket trial design framework.

DR. URV: Thank you, Dr. Kretzler.

I have one more question that I'd like each of you to answer, and that is, you come from consortia that are well established and that have been around for many years. My question to you is, if you're a new academic researcher in a newly established or a very young area of research for

rare disease, what are the most important things to have in place? I guess we could start in the order that you presented.

Andrea?

DR. GROPMAN: Yes. So I think definitely an infrastructure that supports clinical research; access to the patient population; two other experienced investigators who have done clinical trials is important; and access to the FDA resources as part of this conference.

I think really thinking broadly about where you want to go with it. I think thinking creatively, thinking of efficacy, or efficiency, of patient evaluation to phenotype them. The longitudinal study is the most valuable resource that a lot of us have in the consortium in terms of phenotyping the patients and figuring out which subset of patients, as Matthias said, would be suitable for which types of clinical trials, especially if they're competing trials going on.

So I think having access to that and also working with more established consortia that have

had experience going forward.

DR. URV: Dr. Lee?

DR. B. LEE: I think there are two things I would highlight in terms of my experience. One is certainly a very passionate and hopefully organized and perhaps mature patient advocacy partner. In the context of the Brittle Bones Consortium, we were successful partly because we built on an infrastructure that the Osteogenesis Imperfecta Foundation invested in.

I think that can be extremely galvanizing and somewhat out of the control of that new investigator that you posited, but that certainly, I think, is critical.

I think the second are other investigators who are invested in this. In many rare diseases, I think we recognize that it is a team. Any single individual really can't achieve and get to the goal. So I would say the patient advocacy organization is absolutely critical and maybe the most important, and then having other investigators who are willing to play on the team together.

DR. URV: Thank you.

Dr. Kretzler?

DR. KRETZLER: Yes, exactly. I think it's all about the patient, and listening carefully to them; also connecting them to other patient interest groups who have significant experience -- obviously not -- in the framework DRDRI are offering can be great I said also for their learning patient interest group.

Then understanding that this is team science and that if you want to go long, you have to go together, and bringing people together who are willing to play in a team science framework, understanding that in our current time and age, there are so many research opportunities and so many different directions, that academic and private entities can benefit from the multifaceted approach as long as we generate creative solutions who will make everybody win, and most of all, our patients in the end.

This is where genomic medicine really has been a fundamental gamechanger since we started our

networks, and there are incredible resources and infrastructures from NIH. And in many instances there are local entities available, and networks of people on this screen to give you advice to whom to connect, where and when, and how to move your strategy forward most effectively together.

DR. URV: Okay.

Here is one more question that any of you could answer or all of you could answer.

How do you envision real-world evidence being used to generate data as a control arm in a clinical trial versus placebo or active control trial?

(No response.)

DR. URV: Anyone want to tackle that one?

DR. B. LEE: Maybe I'll try it. It's probably a question more for our FDA colleagues.

DR. URV: Yes.

DR. B. LEE: Really, I think we are very engaged in this topic and beginning to reach out to patients to get data at -- point of care is probably not the right term, but really more in the

community, so more, quote, "how we would think of real-world."

At this point, from what I've heard, it's certainly a very powerful tool as additional evidence to the single, adequate, well-controlled trial. I'm not sure I've seen that that alone is sufficient and, frankly, may not be such a great idea, at least in the current framework; and the FDA colleagues can comment on this. But it seems as if that's the first pivotal approval that may really impact some of the more downstream developments. So that's my take on this at this point.

DR. URV: Okay.

DR. KRETZLER: The good news is our real world is changing quickly, so even real-world evidence can be leveraged to define patients in a mechanistic term because it would be very important to keep in line what we just discussed.

DR. URV: Any final words? Dr. Gropman?

DR. GROPMAN: I think what my colleagues have said is that we haven't really gone that route

yet, but we need to think about creative approaches to studying drugs and other therapeutics in rare disease. And again, I'd be interested to hear what our FDA colleagues would think of accepting that.

Adjournment

DR. URV: I do think that they mentioned that in an earlier session, but I don't want to speak for them. So I think we can go back and replay the recording and find an answer to that.

I think if we don't have any more questions -- I don't see any more -- I'd like to thank all of our speakers today for their wonderful presentations. I'd like to thank the meeting organizers and the meeting managers who have run this meeting seamlessly today. Thank you for everyone.

Tomorrow morning, we start up again at 9 a.m., and we will have two more sessions. So thank you very much, everyone. Have a good day.

(Whereupon, at 4:00 p.m., the meeting was adjourned.)