Introduction

Rare diseases are a complex and diverse set of conditions which, when taken together, affect an estimated 30 million Americans, or 10% of the population\(^1\). A disease is defined as rare when fewer than 200,000 people in the U.S. are living with the diagnosis at any given time\(^2\), and it is estimated that 7,000 such diseases exist\(^3\). Over 80% of rare diseases are genetic\(^4\), clinically progressive, and life-limiting or life-threatening. Despite some progress in developing drugs to treat these conditions, only five percent of rare diseases have an approved treatment\(^5\), and continued advocacy from patient and research organizations has encouraged industry, regulators, and policy-makers to identify new opportunities for improving the science and process of drug development for rare disease populations.

A primary barrier to rare disease drug development are the statistical challenges posed by the small affected populations. The limited number of patients with any given rare disease makes the size of most trials necessarily small, and that in turn restricts the usage of some forms of inferential statistics (i.e., methods that rely on large sample approximations for their performance). Furthermore, small population trials employing conventional design parameters and characteristics may only be able to detect very large treatment effects and may preclude adequate replication.

Other significant challenges include:

- Geographically dispersed affected populations and limited numbers of clinical research centers that hinder clinical trial participation by patients and caretakers and increase the need for global, multi-regional clinical trials of rare diseases, which introduces additional challenges in trial design, conduct, and reporting

- Phenotypic diversity within small populations, including inter-patient variability and intra-patient (e.g. diurnal/weekly) variability and diagnostic delays that frequently must be addressed through carefully developed trial endpoints and study designs

- Lack of reliable and well-defined trial endpoints, outcome measures, and biomarkers, coupled with a small population available for tool development

- Limited availability of contemporaneous, disease-specific, natural history studies that could serve as a source for external control data or for biomarker or endpoint validation

- No precedent for drug development for some specific diseases

- Unanticipated or unaddressed ethical considerations for children participating in pediatric clinical trials, which are governed in the United States by additional regulatory requirements
Given these challenges, there is a clear need for novel and efficient trial designs and data analysis methods that can be appropriately and effectively used in rare diseases. To support the goal of expediting the development and review of novel drugs intended to address unmet medical needs, the Duke-Margolis Center for Health Policy, under a cooperative agreement with the U.S. Food and Drug Administration, is convening this public workshop to advance discussion around such designs, methods, and statistical tools for rare disease drug development. Specifically, the objectives for this workshop are to bring rare disease stakeholders together to discuss the challenges associated with the development and regulatory decision-making for rare disease treatments and to discuss promising approaches that may help overcome these challenges.

**Session I: Using Prior Data from Early Phase Trials to Inform Phase 3 Designs**

For trials with small sample sizes, standard statistical methods that rely on large sample approximations to establish their operating characteristics (e.g., Type I error probabilities) do not necessarily apply, thus making the results and conclusions of the analysis difficult to interpret. Additionally, modern trials of rare diseases often involve multiple patient subgroups, such as gene-mutation subgroups, with even smaller sample sizes. In these types of situations Bayesian analysis methods may be useful and appropriate. For example, they may allow for borrowing of information from homogeneous subgroups while discounting the information from heterogeneous subgroups when subgroup sample sizes are too small to support reliable estimation. In addition, when there is historical or external information available (e.g., through historical control groups, early phase trials, or some observational studies), Bayesian methods naturally lend themselves to incorporating the information to improve efficiency while ensuring that historical data that are obsolete or less useful can be down-weighted.

In this session, discussion will focus on the use of prior data (in particular from early phase trials) to inform registrational or phase 3 trial designs. The first presentation will discuss borrowing information from the treatment effect distribution from early phase trials. The second presentation will discuss incorporation of control group data from early phase trials.

**Discussion Questions:**

1. Under what circumstances might information about treatment effects be borrowed to improve the efficiency of a phase 3 trial design? How might the applicability of the early phase data to the phase 3 trial be determined, and how can external data about treatment effects be down-weighted in a principled manner?
2. What factors are important to consider in determining whether to borrow control arm data from early phase trials, and how do those considerations differ from borrowing information about treatment effects?

**Session II: Utilizing Patient Registry and Natural History Study Data to Advance Therapeutic Development for Rare Diseases**

A natural history study is a prospectively planned observational study intended to track the course of disease over time. The purpose is to identify demographic, genetic, disease characteristic,
environmental, and other variables (like treatment modalities and concomitant medications) that are associated with the disease’s development and outcomes. Natural history studies are likely to include patients receiving the current standard of care and/or emergent care, which may alter some manifestations of the disease. A registry is a collection of information about individual subjects who have a specific disease (possibly related diseases) or conditions. While data from a natural history study can be included in a registry, the two terms are not the same.

One approach to accommodate small population sizes is to identify external data sources that could be leveraged to increase the ability of the trial to meet its objectives. Patient or disease registries are one such data source, as are control groups from earlier trials of the same treatment (the topic of Session I). Observational or natural history studies that follow a cohort of patients with the rare disease over time and chart their progression provide another source to consider, when available. The ability to supplement the sample size in the control arm of a trial with external data, however, must be carefully weighted against several factors that influence the usefulness of this approach. External data sources differ in their quality and completeness, as well as in the amount of information included about the disease or diagnosis, covariates, and outcomes.

The standards of care represented in a registry or in a longitudinal cohort study may vary significantly for a rare disease, especially if the registry spans a long time period. These factors should be carefully considered when evaluating whether an external data source can be leveraged to increase the power of a planned trial.

Assuming an external data source of sufficient quality, completeness, and richness of information has been identified, there are various statistical approaches that can be utilized to incorporate the external data in the planned trial. One approach of particular interest corresponds to combining external data with concurrently controlled data in an ongoing trial, with interim assessments on their similarity guiding decisions about how the two sources should be combined.

The day’s second session will cover opportunities and challenges specific to designing studies that utilize registry, historical, or other external data. An introductory presentation will outline several topics for group discussion including the utility and pitfalls of harnessing registries to maximize the use of patients and their data.

Discussion questions:
1. How can data collection for a prospective registry be designed to include as many variables as potentially needed, but at the same time not waste resources and create significant burden for patients, caregivers, and clinical partners who are populating the registry?
2. How can patient powered networks and other worldwide organizations help improve data quality of registries? Would encouraging use of CDISC data standards for registries or natural history studies be useful?
3. What statistical methods are well-suited for use with external data sources (e.g., patient registries and natural history studies) when data from these sources are to serve as the comparator group in a clinical trial?

Session III: Leveraging Master Protocols for Trials with Small Patient Populations

Whether under traditional or adaptive designs, the ability to run a trial with the scientifically required sample sizes can frequently be impossible in rare diseases as the number of patients needed can be
larger than the number of patients available. This problem can be magnified when multiple clinical trials are running simultaneously, causing competition for the few patients that are available. These situations, where patients are scarce and multiple trials are running, can create a myriad of challenges for the scientific community: different trials invariably have different protocols, different inclusion and exclusion criteria, different visit schedules, and different endpoint collection, among other factors. In short patients are not being used efficiently to identify promising therapies in areas of unmet need.

Researchers may be able to overcome these challenges by utilizing master protocols designed to study multiple therapies, multiple diseases, or both under one over-arching protocol and with one shared infrastructure. Advantages include the ability to screen patients for multiple trials simultaneously, to determine their eligibility based on their biomarker profile (as opposed to subjecting patients to sequential screening of trials over a longer period of time), and to share control patients with similar profiles across trials of different therapies. Master protocols may be established as platform trials that run perpetually and allow for treatments and disease subtypes to be added and dropped to the platform design over time.

This session will cover opportunities and challenges to constructing master protocols in the rare disease context. An opening presentation will highlight best practices and emerging lessons from ongoing master protocol development.

**Discussion questions:**

1. What have stakeholders learned from experience with master protocols to-date in rare disease settings?
2. When multiple companies are involved, what methods have been successful to gain their buy-in and to successfully negotiate agreement on the timing of trials and protocol details?
3. What are best practices for establishing a clear business case for companies to participate in master protocols? How can stakeholders best address concerns around sustainability, funding, and long-term sponsor cooperation?

---

3 FAQs About Rare Diseases | Genetic and Rare Diseases Information Center (GARD) – an NCATS Program. (n.d.). Retrieved September 19, 2016, from https://rarediseases.info.nih.gov/diseases/pages/31/faqs-about-rare-diseases
5 Ibid