Data Sharing to Accelerate Therapeutic Development for Rare Diseases

Virtual (Zoom)
August 18 & 19, 2020

Discussion Guide

Introduction

Since the passage of the Orphan Drug Act in 1983, over 600 treatments for rare diseases have been approved by the U.S. Food and Drug Administration (FDA).[1] In 2018, treatments for rare diseases constituted the majority of new drugs approved by the FDA.[2] However, only approximately five percent of rare diseases have an effective treatment.[3] The unmet need for effective therapeutics for rare conditions is due in part to challenges with disease characterization, small patient populations from which to draw clinical trial participants, and the overall rising costs of new drug development.[3,4]

One approach to address these issues and facilitate the discovery of new treatments for rare diseases is through more widespread adoption of collaborative research methods, such as pre-competitive data sharing. In an effort to speed the development of new therapeutics for rare diseases, organizations such as the FDA, the National Center for Advancing Translational Sciences (NCATS) at the National Institutes of Health (NIH), National Organization for Rare Disorders (NORD), the Critical Path Institute (C-Path), and several others have undertaken initiatives to promote pre-competitive data sharing in both preclinical and clinical settings.

There are a number of technical, operational, and regulatory considerations associated with the maintenance and use of a shared data infrastructure for rare disease. Additionally, multi-stakeholder collaboration, including input from academic, industry, patient, and regulatory stakeholders, is essential to maintaining useful shared data resources and productive collaborative research networks.

Accordingly, the Robert J. Margolis, MD, Center for Health Policy at Duke University, under a cooperative agreement with the FDA, is convening this public meeting to discuss the importance of data sharing to advance rare disease drug development. This workshop will explore topics and considerations related to:
- The importance of data sharing in clinical trials for rare diseases and current areas of unmet need in preclinical and clinical research for rare diseases
- Approaches to enhancing and maintaining an adaptable rare disease research infrastructure and to ensuring data quality, standardization, and interoperability across shared data resources
- The role of regulatory agencies and other stakeholder groups in strengthening the research infrastructure and shared data resources for rare diseases

Session 1: Benefit of Shared Data in Rare Disease Characterization and Drug Development

While data sharing can support therapeutic development in many disease areas, it is particularly beneficial for rare diseases for two reasons. First, due to the inherently small size of the population impacted by an individual rare disease, researchers face challenges in recruiting a sufficient number of patients to conduct clinical trials with the statistical power needed to draw definitive conclusions about treatment efficacy. Second, due to data limitations, disease heterogeneity, and limited investment in basic research for some rare diseases, the natural history of many rare diseases is not well understood.

Enhancements to existing shared data infrastructure and tools can support both preclinical and clinical research for rare diseases. For example, expanded contribution to centralized shared data resources (e.g., patient registries, biospecimen repositories) can facilitate better understanding and documentation of disease pathogenesis, incidence, prevalence, and response to treatment with a reduced total investment of resources.[5] This information can be used to support innovation and efficiency in trial design, inform targeted patient recruitment, and reduce initial Research & Development costs for sponsors.[3] Furthermore, enhancement of a shared data infrastructure can support international alignment in rare disease research, address some issues with small patient populations, and encourage global collaboration to mitigate research risk and burden on individuals impacted by rare diseases.[3]

Presentations in this session will discuss the specific benefits of enhancing shared data infrastructure and tools and the opportunity cost of underinvestment in the development, maintenance and use of shared resources. Speakers will highlight current areas of unmet need in basic, translational, and clinical research for rare diseases and how pre-competitive data sharing could supplement efforts to advance preclinical research and clinical development. Additionally, presenters will also provide insights from a regulatory perspective relating to application review for rare disease therapeutics.

Session 2: Leveraging Shared Data Sources for Rare Disease Drug Development

There are a number of existing initiatives aimed at expanding the pre-competitive data sharing infrastructure for rare diseases in both the preclinical and clinical setting. In the preclinical space, collaborative data collection initiatives such as the database of Genotypes and Phenotypes (dbGaP) and NORD's IAMRARE registries are leveraged for disease characterization and natural history studies. The dbGaP database compiles genotypic and phenotypic data from sources such as genome wide association studies (GWAS), medical sequencing, and molecular diagnostic assays.[6] The IAMRARE registry program collects patient data to facilitate natural history studies, [7] important tools to support better understanding of disease progression and inform clinical trial design for rare conditions through the identification of disease biomarkers and the generation of data that could be used as an external control group.[8]

Initiatives such as data analytics platforms, open access clinical databases, and clinical trial networks with common trial infrastructure can further build on preclinical data collection efforts to advance clinical research in rare diseases. For example, in 2019, C-Path, FDA, and NORD leveraged the IMRARE registry data to launch the Rare Disease Cures Accelerator-Data and Analytics Platform (RDCA-DAP).[9, 10]

The RDCA-DAP uses data curation processes to standardize and link disparate data sources to generate drug development tools and de-identified aggregate data for use in rare disease research.

Similarly, the Yale Open Data Access (YODA) Project collects and standardizes existing clinical data to ensure that the data contribute to the body of evidence about the safety and efficacy of individual treatments for rare diseases.[11] Other collaborative research tools exist to facilitate increased trial enrollment, particularly important for rare diseases which impact a patient population who may be spread over a large geographical area. Collaborative research to support trial enrollment can facilitate the conduct of trials that are sufficiently powered to generate definitive results about treatment efficacy. For example, the European Cystic Fibrosis Society–Clinical Trials Network, established in 2009, is aimed at optimizing the development and evaluation of new and approved treatments for cystic fibrosis through efficient clinical studies conducted at 30 selected sites across 11 European countries.[12]

In this session, participants will describe the current landscape of shared data resources and tools available for use in research for rare disease, including, patient registries and collaborative research networks. Discussion will also focus on approaches to integrate shared data sets in clinical trials, including considerations for primary data collection and database management.

Discussion Questions:
1. How can shared resources for preclinical research be used to better support disease characterization?
2. What are the main challenges and practical considerations associated with the integration of shared data in clinical trials?
3. What are the primary impediments to pharmaceutical industry use (data contribution and extraction) of pre-competitive data sharing resources for rare disease drug development?

Session 3: Ensuring Quality in Shared Data Resources— Data Interoperability, Protection, and Management

Encouraging collection of high-quality data is imperative to the widespread adoption of a shared data infrastructure and use of share data tools and resources. Data quality can be defined based on a variety of factors including validity, completeness, and comparability. The intended use of the data should dictate the definition and measurement of data quality.[13] The purpose of the shared data resource should be clear (e.g., regulatory use), so stakeholders can develop appropriate methods to identify, collect, and standardize relevant exposures, outcomes, and covariates. Related to data quality, is data provenance, or the history of data from generation, through curation, to analysis. Good data provenance is key for transparency and necessary for any system that intends to let different groups of users access and analyze the data for different purposes. To support quality, data needs to be properly managed to protect patient privacy, database administrators should maintain data provenance for transparency, and data interoperability should be encouraged for increased utility.

Due to small patient populations which inherently limit the amount of data available for use in preclinical and clinical research, data sharing platforms for rare diseases need to be capable of collecting and integrating data from many different sources. Such data sources include patient registries, clinical trials, natural history studies, postmarket studies, and databases housing genomic and phenotypic information. These data can be clinician-reported, patient-reported, electronic health record (EHR) data, etc.[14] Harmonizing these data through the adoption of common data elements and standard terminology and definitions increases interoperability and maximizes the utility of available data. The curation of interoperable data presents an opportunity to use an adaptive rare disease data platform

3

that can be leveraged across multiple rare diseases and facilitate testing of multiple potential treatments.

Balancing patient rights for control over their personal health information with the need for robust shared data resources to support rare disease research presents challenges.[15] Protecting patient privacy is particularly challenging in rare disease research as small patient populations can make it easier to link de-identified data to patients impacted by a particular disease. Better engagement with patients and patient advocacy groups may enable more standardized and comprehensive capture of data while respecting patients' preferences related to the use and protection of personal health information. Dynamic informed consent models are one way to involve patients and encourage increased and longitudinal data collection.[16] This is especially important for rare diseases where existing historical data is sometimes insufficient to establish natural history of the disease.

Managing shared data resources requires a multi-disciplinary team with expertise in standardizing disparate data, safeguarding protected health information (PHI), documenting data provenance, and facilitating use.[14] Shared data resource management teams should be prepared to enhance their system over time as research needs along with expectations around informed consent, data standards, and tools for data capture, curation, and analysis evolve.

Discussion in this session will focus on common approaches to support high-quality data collection for input into shared data resources, including data standardization and protection of PHI and personally identifiable information. The group will identify some of the challenges inherent to managing a shared data resource. Participants will also discuss the role of regulatory agencies in strengthening rare disease data collection and shared data resources.

Discussion Questions:
1. What are the primary approaches to ensuring quality in the design of data collection schemes and how do these approaches support the secondary use of data to replicate or extend prior research for rare diseases?
2. What opportunities for alignment with data standards related to common data elements, definitions, and terminology exist for rare disease research?
3. What are the main challenges associated with standardizing data collection and protecting personal health information for rare disease research?
4. How can stakeholders, including regulatory agencies, strengthen the common research infrastructure and shared data resources to support research across rare diseases?

Day 2 | Collaborative Research Approaches and Analytics Platforms to Support Drug Development and Address Challenges in Rare Disease Data Sharing

Session 4: Platform Analytics Tools to Support Rare Disease Drug Development

Platforms to centralize, standardize, and analyze rare disease data provide a framework and enhanced tools for rare disease researchers to leverage in disease characterization and clinical trial design. Adaptable and disease-agnostic, shared data platforms that make use of patient-level data can facilitate improved natural history studies, the development of individual disease models and other tools to support optimized trial design, and the efficient use of data and resources to minimize research burden on the rare disease community.[10]

In 2019, C-Path, in collaboration with NORD and FDA, launched the RDCA-DAP with the intent to unify different sources of patient-level data into an investigational database.[10,17] As illustrated in Figure 1, the RDCA-DAP collects and curates data from disparate sources to produce standardized aggregate datasets which researchers can leverage to support rare disease characterization and therapeutic development for rare diseases.
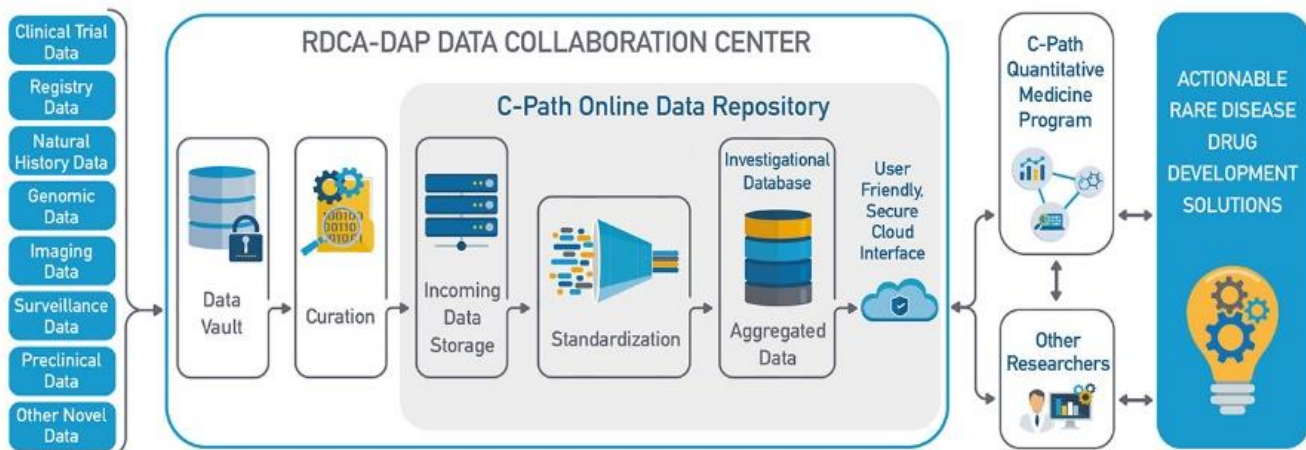

Figure 1: *Diagram of RDCA-DAP Data Collaboration Center*[10]

RDCA-DAP sources data from clinical trials, longitudinal observational studies, patient registries and real-world data (e.g. electronic health records) to support the generation of drug development solutions, such as disease models and clinical trial simulation tools, as well as expanded access to aggregated data after curation.[10] The platform has also integrated NORD's IAMRARE registries as part of its patient registry data.[10]

Participants in this session will provide an overview of the RDCA-DAP platform and discuss the utility of patient-level data in the development of disease models and other drug development tools to support innovative trial design for rare diseases. Participants will also discuss general challenges to securing multi-stakeholder collaboration and funding for the maintenance of shared data resources.

Discussion Questions:
1. What are the benefits and tradeoffs associated with unifying disparate sources of patient-level data (e.g., do you lose outcomes captured in certain datasets if they are not captured in all datasets?)
2. What are the most important operational considerations for maintaining a disease-agnostic data analytics platform and what are the challenges to integrating disease-specific needs into platform analysis and tools?
3. How are data protection and intellectual property concerns addressed in the design and implementation of RDCA-DAP and other shared databases?
4. How can shared data analytics platforms support improvements in future data collection for rare diseases?

Session 5: Collaborative Research Networks to Support Rare Disease Drug Development

A shared clinical research infrastructure promotes collaboration amongst researchers and can lead to more efficient clinical trials with reduced risk and burden for patients. For example, clinical trial networks can facilitate increased enrollment in clinical trials, which is especially important for rare diseases given the small size of the patient populations. Furthermore, clinical trial networks enable the comparison of results across trials, allowing each trial to contribute to the totality of evidence on rare disease therapeutics. There are currently a few national and international efforts to create and expand collaborative clinical trial networks specifically for rare disease research.

One such effort is the Rare Disease Clinical Research Network (RDCRN) designed to engage researchers, scientists, patients, and advocacy groups across centers around the world in rare disease research.[18] The network consists of 23 consortia members which collaborate to improve both the availability of existing rare disease data as well as the collection and storage of additional data.[18,19] The Data Management Coordinating Center (DMCC), located at Cincinnati Children's Hospital Medical Center, is tasked with maintaining the data produced by the consortia's research.[19] In addition to maintaining a collaborative clinical research network, the RDCRN program places significant emphasis on promoting future rare disease research. RDCRN supports pilot studies and efforts to train clinical investigators and researchers in conducting rare disease research in order to promote future therapeutic development.[20]

Participants in this session will discuss examples of clinical trial networks, including the RDCRN, and discuss the value of such collaborations in promoting and supporting rare disease drug development. Panelist discussion will focus on practical challenges to the implementation of rare disease clinical trial networks, including any issues engaging stakeholders, centralizing and standardizing data, and eliciting and integrating patient input.

Discussion Questions:
1. How can collaborative research networks best support clinical trial readiness for rare diseases?
2. What are key impediments to the increased adoption of shared data collection protocols for rare disease clinical trials?
3. What are the most important lessons learned from administering trial design and conduct through a centralized trial coordination center?

Session 6: Synthesis and Next Steps

Discussion in this session will address next steps and stakeholder roles for the enhancement of an adaptable data sharing infrastructure to support rare disease research and therapeutic development. Participants will discuss feasible approaches to improving quality and availability of shared data resources to support disease characterization. Participants will also discuss next steps for maximizing the utility of these shared resources to reduce research burden on individuals impacted by rare diseases and to support the design of innovate clinical trials that produce definitive answers to researchable questions about therapeutic safety and efficacy.

Discussion Questions:
1. What are the near-term, next steps to improve the quality and availability of aggregate data to support preclinical research?
2. What are the near-term, next steps to facilitate increased integration of shared data in clinical trials for rare diseases?
3. What is the role of researchers, industry, patient groups, and regulatory stakeholders in facilitating the enhancement of shared data infrastructure and increased data contributions to shared resources?

_____

1. Developing Products for Rare Diseases & Conditions. FDA. Published May 28, 2020. Accessed July 15, 2020. https://www.fda.gov/industry/developing-products-rare-diseases-conditions
2. Orphan Drug Act Resolution Introduced in the House of Representatives. NORD (National Organization for Rare Disorders). Published March 18, 2019. Accessed July 17, 2020. https://rarediseases.org/orphan-drug-act-resolution-introduced-in-the-house-of-representatives/
3. Kaufmann P, Pariser AR, Austin C. From scientific discovery to treatments for rare diseases – the view from the National Center for Advancing Translational Sciences – Office of Rare Diseases Research. *Orphanet Journal of Rare Diseases*. 2018;13(1):196. doi:10.1186/s13023-018-0936-
4. Why does early drug development cost so much and can it become more efficient? Published March 22, 2019. Accessed July 15, 2020. https://www.pcisynthesis.com/want-to-know-why-early-drug-development-costs-so-much/
5. Boat TF, Field MJ, eds. 4. Discovery Research for Rare Diseases and Orphan Product Development. In: *Rare Diseases and Orphan Products: Accelerating Research and Development*. Washington, DC: National Acad. Press; 2010. https://www.ncbi.nlm.nih.gov/books/NBK56191/
6. dbGaP. National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/about.cgi. Accessed August 12, 2020.
7. IAMRARE™ Registry Program. NORD (National Organization for Rare Disorders). Accessed July 15, 2020. https://rarediseases.org/iamrare-registry-program/
8. U.S. Food and Drug Administration. Rare Diseases: Natural History Studies for Drug Development Guidance for Industry. *Rare Diseases*. Published online 2019:22. Accessed August 12, 2020 https://www.fda.gov/media/122425/download
9. Funded by FDA, C-Path and NORD to Launch Rare Disease Data and Analytics Platform | Critical Path Institute. Accessed July 14, 2020. https://c-path.org/c-path-and-nord-to-launch-fda-funded-rare-disease-data-analytics-platform/
10. RDCA-DAP | Critical Path Institute. Accessed July 15, 2020. https://c-path.org/programs/rdca-dap/

11. Policies & Procedures to Guide External Investigator Access to Clinical Trial Data. The YODA Project | Policies & Procedures to Guide External Investigator Access to Clinical Trial Data. https://yoda.yale.edu/policies-procedures-guide-external-investigator-access-clinical-trial-data. Accessed August 12, 2020.

12. Introduction | European Cystic Fibrosis Society (ECFS). Accessed August 6, 2020. https://www.ecfs.eu/ctn

13. Kodra Y, Posada de la Paz M, Coi A, et al. Data Quality in Rare Diseases Registries. *Adv Exp Med Biol*. 2017;1031:149-164. doi:10.1007/978-3-319-67144-4_8

14. Kodra Y, Weinbach J, Posada-de-la-Paz M, et al. Recommendations for Improving the Quality of Rare Disease Registries. *Int J Environ Res Public Health*. 2018;15(8). doi:10.3390/ijerph15081644

15. Mascalzoni D, Paradiso A, Hansson M. Rare disease research: Breaking the privacy barrier. *Appl Transl Genom*. 2014;3(2):23-29. Published 2014 Apr 18. doi:10.1016/j.atg.2014.04.003 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4882025/

16. Dankar FK, Gergely M, Malin B, Badji R, Dankar SK, Shuaib K. Dynamic-informed consent: A potential solution for ethical dilemmas in population sequencing initiatives. *Computational and Structural Biotechnology Journal*. 2020;18:913-921. doi:10.1016/j.csbj.2020.03.027

17. Research C for DE and. Rare Disease Cures Accelerator. *FDA*. Published online June 1, 2020. Accessed July 15, 2020. https://www.fda.gov/drugs/regulatory-science-research-and-education/rare-disease-cures-accelerator

18. About the RDCRN. National Center for Advancing Translational Sciences. Published March 11, 2015. Accessed July 14, 2020. https://ncats.nih.gov/rdcrn/about

19. About Us. Rare Diseases Clinical Research Network. Accessed July 15, 2020. https://www.rarediseasesnetwork.org/about

20. Groft SC, Gopal-Srivastava R. A model for collaborative clinical research in rare diseases: experience from the Rare Disease Clinical Research Network program. *Clinical Investigation*. 2013;3(11):1015-1021. doi:10.4155/cli.13.101

Duke-Margolis Center for Health Policy | 🐦@DukeMargolis #RareDataShare | healthpolicy.duke.edu