FDA Center for Drug Evaluation and Research and Johns Hopkins Center of Excellence in Regulatory Science and Innovation (CERSI) Workshop

Addressing Challenges in the Design and Analysis of Rare Disease Clinical Trials: Considerations and Tools

May 2 – 3, 2023
May 2, 2023

Collection and Use of Fit-for-Purpose Data for Rare Disease Drug Development
Welcome

Kerry Jo Lee, MD
Associate Director for Rare Diseases
Division of Rare Diseases and Medical Genetics, Office of New Drugs, Center for Drug Evaluation and Research, FDA
Session 1:
How to Collect Quality and Fit-for-Purpose Data

Moderator: Scott Winiecki, MD
Team Lead
Rare Diseases Team, Division of Rare Diseases and Medical Genetics,
Office of New Drugs, Center for Drug Evaluation and Research, FDA
CDER-JHU CERSI Workshop

Regulatory Perspectives on Real-World Data

2 May 2023

John Concato, MD, MS, MPH
Associate Director for Real-World Evidence Analytics
Office of Medical Policy
Center for Drug Evaluation and Research
U.S. Food and Drug Administration
Disclaimer

- Views and opinions expressed are those of the presenter and should not be attributed to the Food and Drug Administration.
- No conflicts of interest exist related to this presentation.
- Mention of a commercial product should not be construed as actual or implied endorsement.
Objectives

- Recognize historical context leading to current use of the terms “real-world data” and “real-world evidence”
- Understand main components of FDA’s Real-World Evidence Program, emphasizing guidance development
- Identify challenges and potential contributions of using real-world data and real-world evidence
**‘Real-World’ Definitions (from 2018 FDA Framework)**

**Real-World Data (RWD)** are data relating to patient health status and/or delivery of health care **routinely collected from a variety of sources**

- electronic health records (EHRs)
- medical claims data
- product and disease registries
- data from digital health technologies in non-research setting
- other data sources that can inform on health status, such as questionnaires

**Real-World Evidence (RWE)** is clinical evidence regarding the usage and potential benefits/risks of a medical product **derived from analysis of RWD**

Generated using various study designs—including but not limited to **randomized trials (e.g., pragmatic clinical trials)**, externally controlled trials, and observational studies
Background on ‘Big Data’

**Origin**: term appeared in computer science literature during 1990s, often referring to data too large to be stored in then-conventional storage systems.

**Contemporary usage**: *Big Data* represents “[...] shorthand for advancing trends in technology that open the door to a new approach to understanding the world and making decisions” (Lohr S, *New York Times*, 11 Feb 2012)

**Perspective**: modern technology has increased quantity and forms of available data as well as the speed to merge and manipulate data, yet integration and analysis of large-scale data has always been integral to epidemiology.
21st Century Cures Act of 2016

• FDA established a program to evaluate the potential use of real-world evidence (RWE) to:
  o Support a new indication for a drug approved under section 505(c)
  o Satisfy post-approval study requirements

• Draft framework issued in December 2018:
  o Describe sources of RWE, challenges, pilot opportunities, etc.

• Draft guidance for industry issued in Sep, Oct, Nov, Dec 2021

• Standard for substantial evidence remains unchanged; commitments met for Prescription Drug User Fee Act (PDUFA) VI; new Advancing RWE initiatives in PDUFA VII
Background on ‘Real-World Evidence’

**Origin:** “real world” is a non-specific modifier; “real-world data (RWD)” and “real-world evidence (RWE)” appeared in medical literature as of the 1970s or earlier, in various contexts (*terms to be defined in subsequent slide*).

**Contemporary usage:** RWD and RWE have specific regulatory implications.

**Perspective:** Older epidemiologic terms were sufficient, but emergence of big data and enactment of 21st Century Cures has led to sometimes confusing use of different taxonomies for study design.

**Example:** “RWE study” is not synonymous with “observational study”; additional details are needed to classify study design.
FDA’s Real-World Evidence (RWE) Program

- Applies to Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), & Oncology Center of Excellence (OCE) – Note: Center for Devices and Radiological Health (CDRH) has separate program

- Multifaceted program to implement RWE:
  1) internal processes
  2) external stakeholder engagement
  3) research (“demonstration”) projects
  4) guidance development

https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence
1) Internal and 2) External Engagement

- **Real-World Evidence Subcommittee internal activities**, w/ membership comprised of FDA staff from multiple CDER and CBER Offices:
  - providing oversight of policy development on RWE (e.g., guidances)
  - offering resources and leadership (e.g., to review divisions)
  - other activities

- **RWE Subcommittee external activities** include:
  - providing feedback on early-stage proposals from sponsors, vendors, etc.
  - discussing initiatives presented to Subcommittee for consideration

- **Additional activities**, beyond the Subcommittee, include:
  - holding FDA- or Center-level public meetings on RWE-related topics
  - conducting FDA small business & industry webinars, speaking engagements
3) RWE Demonstration Projects – Examples

**Data**
- ‘OneSource’ project to improve quality of EHR data
- Collection and use of EHR data from neonatal intensive care units

**Study Design**
- RCT-DUPLICATE trial emulations
- Statistical approach for RCT designs w/ ‘hybrid’ control arms

**Tools**
- Evaluation of confounded treatment effects
- Targeted learning framework for causal effect estimation
4) FDA Draft RWE Guidance – Sep-Dec 2021

Guidance for Industry

DRAFT GUIDANCE

Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products

Data Standards for Drug and Biological Product Submissions Containing Real-World Data

Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products

https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence
Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products

Guidance for Industry

DRAFT GUIDANCE

September 2021
Real World Data/Real World Evidence (RWD/RWE)
Focus of draft guidance:

- Selection of data source(s) to appropriately address the study question
- Development and validation of definitions for exposures, covariates, outcomes
- Data provenance during accrual, curation, analysis

Note: choice of study design and method of statistical analysis are outside of guidance scope
Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products

DRAFT GUIDANCE

November 2021
Real World Data/Real World Evidence (RWD/RWE)
Focus of draft guidance:

• Registry fitness-for-use in regulatory decision-making, focusing on attributes that support collection of relevant and reliable data

• Linking a registry to other data source(s) for supplemental information, such as data from medical claims, electronic health records (EHRs), digital health technologies, or other registries

• FDA review of submissions that include registry data

Note: The guidance does not provide recommendations on choice of study design or approach to statistical analysis
Data Standards for Drug and Biological Product Submissions Containing Real-World Data Guidance for Industry

DRAFT GUIDANCE

October 2021
Real-World Data/Real-World Evidence (RWD/RWE)
Data Standards Guidance – Overview

Focus of draft guidance:

• Processes for managing RWD
• Conforming RWD to FDA data standards
• Mapping RWD to FDA submission standards
• Considerations for data transformations

Note: this guidance applies regardless of the type of RWD
Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products

DRAFT GUIDANCE

December 2021
Real World Data/Real World Evidence (RWD/RWE)
• Marketing application to support safety/effectiveness of a drug must satisfy applicable legal standards to be approved or licensed, even if 21 CFR part 312 (Investigational New Drug Application) does not apply

• Two classifications of non-interventional studies:
  1) *involve only* analysis of data on use of marketed drug in routine practice
  2) *include* ancillary protocol-specified activities or procedures (e.g., lab tests, imaging studies, questionnaires)
• FDA does not consider these types of studies to be clinical investigations under 21 CFR part 312
• Nonetheless, protection of human subjects is critical; sponsors must ensure applicable requirements met per FDA regulations 21 CFR parts 50 (Protection of Human Subjects) & 56 (Institutional Review Boards)
Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products

DRAFT GUIDANCE

February 2023
Real-World Data/Real-World Evidence (RWD/RWE)

https://www.fda.gov/media/164960/download
Externally Controlled Trials Guidance – Overview

Focus of draft guidance:

• Importance of design considerations (e.g., finalize protocol before analyzing data)

• Data considerations for the external control arm (e.g., various comparability issues)

• Analysis considerations (e.g., “FDA does not recommend a particular approach”)

• Considerations to support regulatory review (e.g., access to patient-level data)

Note: Guidance does not address external control data a) based on summary-level estimates, or b) supplementing a control arm in a traditional randomized trial
IV. CONSIDERATIONS TO SUPPORT REGULATORY REVIEW

A. Communication with FDA

Sponsors should consult with the relevant FDA review division early in a drug development program about whether it is reasonable to conduct an externally controlled trial instead of a randomized controlled trial. As part of these discussions, sponsors should provide a detailed description of the (1) reasons why the proposed study design is appropriate, (2) proposed data sources for the external control arm and an explanation of why they are fit for use, (3) planned statistical analyses, and (4) plans to address FDA’s expectations for the submission of data.
Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drug and Biological Products

Guidance for Industry

September 2022
Procedural

https://www.fda.gov/media/124795/download
## Status of FDA RWE Guidance – April 2023

<table>
<thead>
<tr>
<th>Category</th>
<th>Topic</th>
<th>Status</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data considerations</td>
<td>EHRs and claims data</td>
<td>draft published</td>
<td>Sep 2021</td>
</tr>
<tr>
<td></td>
<td>Registry data</td>
<td>draft published</td>
<td>Nov 2021</td>
</tr>
<tr>
<td>Submission of data</td>
<td>Data standards</td>
<td>draft published</td>
<td>Oct 2021</td>
</tr>
<tr>
<td>Applicability of regulations</td>
<td>Regulatory considerations</td>
<td>draft published</td>
<td>Dec 2021</td>
</tr>
<tr>
<td>Design considerations</td>
<td>Externally controlled trials</td>
<td>draft published</td>
<td>Feb 2023</td>
</tr>
<tr>
<td></td>
<td>RCTs in clinical practice settings</td>
<td>draft in development</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Non-interventional studies</td>
<td>draft in development</td>
<td>–</td>
</tr>
<tr>
<td>Procedural</td>
<td>Submitting documents</td>
<td>final published</td>
<td>Sep 2022</td>
</tr>
</tbody>
</table>
Digital Health Technologies for Remote Data Acquisition in Clinical Investigations

*DRAFT GUIDANCE*

December 2021
Clinical/Medical
Current Status of Real-World Evidence

Real-World Evidence — Where Are We Now?
John Concato, M.D., M.P.H., and Jacqueline Corrigan-Curay, J.D., M.D.

Issue being addressed: More than five years after passage of the 21st Century Cures Act, the terms RWD and RWE are being used inconsistently and interchangeably.

Content of article:
- addressed two common misconceptions
- provided conceptual overview of study design
- described FDA demonstration projects and guidance
- highlighted regulatory approvals
- offered path forward
Misconceptions regarding RWD & RWE

Frequent instances of:

- **Misconception #1 – RWD & RWE are new concepts:** “In reality, sources of data and types of study design haven’t fundamentally changed, but electronic access to more detailed clinical data is evolving & the data are becoming more relevant and reliable”

- **Misconception #2 – A simple dichotomy of randomized trials vs. observational studies exists:** “In reality, clinical trials are defined by assignment of treatment according to an investigational protocol, and single-arm trials face challenges similar to those in observational studies in determining whether difference in clinical outcomes (compared to an external control group) represent actual treatment effects”
### Real-World Evidence — Where Are We Now?

John Concato, M.D., M.P.H., and Jacqueline Corrigan-Curay, J.D., M.D.

<table>
<thead>
<tr>
<th>Randomized, Interventional Study</th>
<th>Nonrandomized, Interventional Study</th>
<th>Nonrandomized, Noninterventional Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional randomized trial using RWD in planning</td>
<td>Trial in clinical practice settings, with pragmatic elements</td>
<td>Externally controlled trial</td>
</tr>
<tr>
<td>RWD used to assess enrollment criteria and trial feasibility</td>
<td>Selected outcomes identified using, e.g., health records data, claims data, or data from digital health technologies</td>
<td>Single-group trial with external control group derived from RWD</td>
</tr>
<tr>
<td>RWD used to support selection of trial sites</td>
<td>RCT conducted using, e.g., electronic case report forms for health records data or claims data</td>
<td>Cohort study</td>
</tr>
</tbody>
</table>

**Generation of RWE**

**Increasing reliance on RWD**

---

**Reliance on RWD in Representative Types of Study Design.**

RCT denotes randomized, controlled trial; RWD real-world data; and RWE real-world evidence.
RWE for Effectiveness: Overview of FDA Approach

Key considerations (from 2018 Framework):

• Whether the RWD are fit for use
• Whether the trial or study design used to generate RWE can provide adequate scientific evidence to answer or help answer the regulatory question
• Whether the study conduct meets FDA regulatory requirements
FDA Approves New Use of Transplant Drug Based on Real-World Evidence

- Prograf® (tacrolimus) approved for prophylaxis of organ rejection in patients receiving liver transplants in 1994 (later for kidney & heart) based on RCT evidence, and the drug is used widely in clinical care.

- RCTs not done for lung transplant, but sponsor (Astellas Pharma US) submitted supplemental New Drug Application to FDA with non-interventional ‘RWE’ study.

- Study data and design were evaluated according to FDA standards.

New Indication for Prograf® Based on RWE (cont’d)

Data: US Scientific Registry of Transplant Recipients (SRTR) data on all lung transplants in US during 1999–2017

Design and conduct: non-interventional (observational) treatment arm, compared to historical controls; analysis plan & patient-level data provided to FDA

Review: FDA determined this non-interventional study w/ historical controls to be adequate and well-controlled. Of note, outcomes of organ rejection and death are virtually certain without therapy, and the dramatic effect of treatment helps to preclude bias as explanation of results.

RWE – Representative Problems

Real-world data sources:
- issues related to data reliability and clinical relevance
- need for linkage to other data sources
- missing or “mistimed” data
- suitable capture of endpoints

Non-randomized study designs:
- threat of residual confounding
- problems with index date ("zero time")
- use of inappropriate comparator

Conduct of non-randomized studies:
- insufficient confirmation of pre-specified protocol and analysis plan
- issues related to FDA inspection
Summary

- “Big data” contributed to changes in how evidence generation is approached & described; research methods are also evolving.

- FDA’s RWE guidance & related efforts, along with other stakeholders, are addressing current challenges in using real-world data & evidence.

- FDA will maintain evidentiary standards while considering RWD/RWE for regulatory decision-making.
Acknowledgments

- Michael Blum, Phil Budashewitz, Jacqueline Corrigan-Curay, M. Khair ElZarrad, Tala Fakhouri, Kayla Garvin, Scott Gordon, Stefanie Kraus, Beth Kunkoski, Nahleen Lopez, Juanita Marner, Kristen Miller, Dianne Paraoan, Ken Quinto, Motiur Rahman, Leonard Sacks, Kim Smith

- Other colleagues in:
  - CDER Offices of Medical Policy, New Drugs, Surveillance & Epidemiology, Biostatistics, Regulatory Policy, Scientific Investigations, Strategic Programs, Translational Sciences
  - Center for Biologics Evaluation & Research; Oncology Center of Excellence; Center for Devices & Radiological Health
  - Office of the Commissioner
How C-Path Uses the Latest Data Management and Data Science Techniques to Maximize the Value of Data

Ramona L. Walls, Exec. Dir. of Data Science

CDER-JHU CERSI Workshop on Rare Diseases
May 2, 2023
Rare disease data are rare

• Progress toward therapies for rare diseases is hampered by poor understanding of many diseases...
• ...but there is a lot of potentially useful data out there.
• Unfortunately, those data are siloed, non-standard, and sometimes not usable due to data quality issues
Data quality concerns for reuse

- Lack of standardization (an gaps in standards)
- Siloed data sources (no access, different formats, different standards)
- Small patient populations are distributed among multiple sources without reliable methods for uniquely identifying patients
Who is C-Path and What Do We Do?
Who We Are

Mission
Critical Path Institute is a catalyst for innovation that accelerates the path to a healthier world.

Vision
C-Path is an indispensable partner of excellence in medical product development worldwide, shaping innovative scientific and regulatory pathways to accelerate delivery of therapies for patients in need.
C-Path Strengths

Core Competencies

- Biomarkers
- Clinical Outcome Assessments
- Data Management and Standards
- Modeling and Analytics
- Regulatory/Development Science

Bedrock Foundation: Unique Neutral Convener

Unmet Medical Need

DDTs and Other Solutions
How C-Path Works

• Acts as a trusted, neutral third party
• Public-Private Partnerships
• Convenes scientific consortia of industry, academia and government for sharing of data and expertise
  ✓ Active consensus building
  ✓ The best science
  ✓ Shared risk and costs
  ✓ The broadest experience
• Enable iterative FDA/EMA/PMDA participation in developing new methods to assess the safety and efficacy of medical products

Official regulatory endorsement of novel methodologies and drug development tools
Why?
Not every drug works for every patient. It is vital to target the right patients.

Who?
- Researchers
- Regulators
- Clinical trialists
- Advocacy groups

How?
Data from past clinical trials or RWD
Data standardization and integration
Informative models

What can a model do?
Information from model
Regulatory agencies

Results in
- Right Target
- Right Drug
- Right Time
- Right Patient
- SUCCESS
Data Science Advances at C-Path
C-Path Data Collaboration Center

**Mission:** Enable multiple organizations to work together in a neutral setting and share data to maximize its value to inform medical product development and regulatory decision-making

**How:**
- Creation and administration of data storage and collaboration platforms
- Planning and execution of multi-source data standardization and aggregation
- Maximize the FAIRness of data by developing and integrating standards and semantic models, tools for consumption and sharing of data, performing data transformations that increase data accessibility, and by performing analyses that transform data into information
- Utilize robust, repeatable processes to ensure data integrity, security and protect patient privacy

**Data Management**
- Data acquisition, curation, QC, standardization, aggregation
- Data analysis, queries, reports
- Data Interrogation and Datamart support
- Data privacy, provenance, governance
- Secure data access

**Data Science & Ontologies**
- Semantic data modeling and standards integration
- Data analysis and transformation pipelines
- Metadata annotation
- Analytics and Tools
- Statistical models and simulations

**Data Platform**
- Design, develop, test and maintain data collaboration platforms for project needs
- Build and Support cloud infrastructure to support data collaboration projects
- Build cloud pipelines and workflows to support data collaboration projects
- Data security and compliance

**Operations**
- Project Management and Operational support
- DCA and DUA execution and tracking
- Status and Reporting
- Auditing

**Precompetitive Neutral Environment**
FAIR Data Principles

- Findable
- Accessible
- Interoperable
- Reusable

• Apply to both human and machine-driven processes
• Humans have an innate understanding of semantics
• Machines can operate at scale with less error
• See Wilkinson et al. 2016
  https://www.nature.com/articles/sdata201618
DCC Approach to Data Management

- Data Contribution Agreement for each dataset
- Transfer anonymized data through secure link
- Curate, Standardize, Annotate Data
- Integrate into Data Sharing Platform
- Extract and Analyze Data

Datasets
Aggregate Data
Datamarts
Analysis subsets
DCC Approach to Data Management

- Data Contribution Agreement for each dataset
- Transfer anonymized data through secure link
- Curate, Standardize, Annotate Data
- Integrate into Data Sharing Platform
- Extract and Analyze Data

Innovations:
- Standard DCAs
- Machine readable DCAs
DCC Approach to Data Management

Data Contribution Agreement for each dataset

Transfer anonymized data through secure link

Curate, Standardize, Annotate Data

Integrate into Data Sharing Platform

Extract and Analyze Data

Datasets
 Aggregate Data
 Datamarts
 Analysis subsets

Innovations:
• Secure transfer via cloud services (AWS, Azure)
• Anonymization services
• Federated access
DCC Approach to Data Management

Data Contribution Agreement for each dataset

Transfer anonymized data through secure link

Curate, Standardize, Annotate Data

Integrate into Data Sharing Platform

Innovations:
- Responsive curation
- Multiple standards (CDISC, OMOP, OBO)
- Scripting and automations
- Ontology and knowledge graph development
DCC Approach to Data Management

Data Contribution Agreement for each dataset

Transfer anonymized data through secure link

Curate, Standardize, Annotate Data

Integrate into Data Sharing Platform

Innovations:
- Advanced search, discovery, visualization, subsetting
- Fine-grained, flexible access controls/sharing permissions
- Standardized DUAs

Extract and Analyze Data
DCC Approach to Data Management

Data Contribution Agreement for each dataset

Transfer anonymized data through secure link

Curate, Standardize, Annotate Data

Integrate into Data Sharing Platform

Innovations:
• Built in data preview and analytics (R, SQL, VMs)
• Enhanced security (logging, TFA, restricted download)
• Shared analyses, bring your own data

Extract and Analyze Data
# C-Path Data and Analytics Platform (DAP)

The C-Path Data and Analytics Platform (DAP) is a comprehensive online platform designed to facilitate the sharing and access to various datasets and data services. This platform is administered by the Critical Path Institute and provides tools for data management, discovery, query, and access.

The current screenshot displays a list of datasets available on the platform:

- **The PSP Rating Scale as a Prognostic Tool Study**:
  - **Description**: The purpose of this project was to formulate...
  - **Publisher**: Critical Path Institute
  - **Visibility**: Internal
  - **Created on**: September 10, 2021, 11:15:46 AM
  - **Last updated**: September 14, 2021, 2:49:10 PM

- **IAMRARE Desmoid Tumor Patient Registry**:
  - **Description**: A desmoid tumor is an abnormal growth that...
  - **Publisher**: Critical Path Institute
  - **Visibility**: Private
  - **Created on**: September 10, 2021, 2:02:59 PM
  - **Last updated**: May 4, 2022, 1:37:39 PM

- **IAMRARE National PKU Alliance Patient Registry**:
  - **Description**: Phenylketonuria is a genetic disorder inherit...
  - **Publisher**: Critical Path Institute
  - **Visibility**: Internal
  - **Created on**: September 10, 2021, 2:02:59 PM
  - **Last updated**: March 21, 2022, 2:01:07 PM

- **IAMRARE SMARD1 Patient Registry**:
  - **Description**: Spinal muscular atrophy with respiratory disease...
  - **Publisher**: Critical Path Institute
  - **Visibility**: Private
  - **Created on**: September 10, 2021, 2:06:46 PM
  - **Last updated**: May 4, 2022, 1:37:39 PM
Standards, Ontologies, and Knowledge Graph

- **OMOP Common Data Model (CDM)** is a baseline for long tail of registry data and EHR
  - Includes standard vocabularies such as SNOMED, LOINC, RXNORM
- **CDISC Study Data Tabulation Model (SDTM)** for clinical trial data
  - Many of our legacy datasets are already in SDTM
  - Standard vocabularies in NCIT are interoperable with OBO ontologies
- **OBO ontologies** for deep semantic discovery and analysis
- **Rare disease knowledge graph** of patient-level data that is interoperable with external data sources like Orphanet, Monarch, EJP-RD
Data + ontology = knowledge graph (KG)

Anatomical reference models

Clinical data (condition occurrences)

Cross-species knowledge

Diseases
How data contributors can help

Good practices for small and large data generators/contributors
Mismatch between what is shared and what is needed

https://www.appliedclinicaltrialsonline.com/view/establishing-a-basis-for-secondary-use-standards-for-clinical-trials
Mismatch between what is shared and what is needed

**Datasets and Documentation**

- AHA M (Analysis) IPO Dataset: 80%
- SOFM (Raw) IPO Dataset: 60%
- Data Dictionary: 100%
- Dataset Specifications (defines.xml): 100%
- Study Protocol: 70%
- Annotated Case Report Form: 100%
- Clinical Study Report: 100%
- Statistical Analysis Plan: 60%
- Study Data Reviewer's Guide: 100%
- Analysis Data Reviewer's Guide: 100%

**Metadata**

- Study Design: 40%
- Study Arm(s) Provided: 100%
- Patient / Sample Counts: 100%
- Study Start and End Date: 100%
- Adverse Event Encoding: 100%
- Concomitant Medication Encoding: 100%
- MedDRA Encoding: 100%
- Study in SDTM Format: 100%
- SDTM Version: 100%
- Location(s): 100%

**Variable-Level Transformation Report**

- 1 (NOT IMPORTANT): 3.9%
- 2: 2.9%
- 3: 14.4%
- 4: 38.5%
- 5 (MANDATORY FOR USE): 40.4%

**Response Scale**
Mismatch between what is shared and what is needed

Datasets and Documentation

<table>
<thead>
<tr>
<th>Datasets and Documentation</th>
<th>Tier 1</th>
<th>Tier 2</th>
<th>Tier 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw (SDTM)</td>
<td>100%</td>
<td>82%</td>
<td>83%</td>
</tr>
<tr>
<td>Analysis (ADaM)</td>
<td>92%</td>
<td>92%</td>
<td>67%</td>
</tr>
<tr>
<td>Protocol</td>
<td>100%</td>
<td>82%</td>
<td>83%</td>
</tr>
<tr>
<td>Annotated CRF</td>
<td>100%</td>
<td>73%</td>
<td>67%</td>
</tr>
<tr>
<td>Reporting and Analysis Plan / SAP</td>
<td>100%</td>
<td>82%</td>
<td>67%</td>
</tr>
<tr>
<td>CSR</td>
<td>92%</td>
<td>91%</td>
<td>33%</td>
</tr>
<tr>
<td>Dataset Specifications</td>
<td>75%</td>
<td>73%</td>
<td>50%</td>
</tr>
</tbody>
</table>

Sponsor Tiers (by employee count)
Tier 1: 25k+
Tier 2: 5 to 24.99k
Tier 3: Under 5k

(n=12)  (n=11)  (n=6)
Data contributors should:

- Follow FAIR data principles
- Ensure proper anonymization and include anonymization report
- Use standard terminology and data models where possible
  - OMOP and SDTM
  - OMOP standard vocabularies, UMLS, NCIT, NIH CDEs
  - Human Phenotype Ontology (HPO) for "phenotype" descriptions
- Follow consistent data collection practices from year to year, at least aim for backwards compatibility
- Share dictionaries, protocols, other supplemental documents
Thank you!

Critical Path Institute is supported by the Food and Drug Administration (FDA) of the Department of Health and Human Services (HHS) and is 55% funded by the FDA/HHS, totaling $17,612,250, and 45% funded by non-government source(s), totaling $14,203,111. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement by, FDA/HHS or the U.S. Government.
RARE-X

Increasing the speed and productivity of innovation in rare diseases by increasing collection and access of structured and standardized patient data.

Vanessa Vogel-Farley (Global Genes: RARE-X)
5/2/2023
CDER - JHU CERSI Workshop Addressing Challenges in the Design and Analysis of Rare Disease Clinical Trials: Considerations and Tools
The speed and productivity of innovation in rare disease is limited by cost & lack of access to standardized, structured, available patient data.

1. Data exists in silos & is unavailable for open research
2. Data is not in a structured, standardized format that is useful to research / patient communities
3. Data doesn’t yet exist; many communities are too young or don’t have the resources to collect data for research.
From Registries to Real-World Data
What Patient-Powered Registries Enable

Nominate Disease
Identify population of interest and understand where they are in the world

Launch Data Collection
Determine what data is needed
Create relevant patient-reported data collection modules
Launch DNA and clinical collection efforts (if relevant)

Design Trial
Use data to:
Inform trial enrollment criteria
Inform trial endpoints

Support Regulatory Requirements
Leverage registry to collect long-term surveillance data
Advocacy Today: Opportunities & Challenges in Rare Disease

Daily Challenges of Living with a Life-Limiting or Chronic Condition

1. Diagnosis/living with life-altering condition
2. Become disease experts [self and for clinician education]
3. Outreach community building & providing support
4. Starting a non-profit for support or research
5. Education on business, science, research, fundraising, legislative advocacy
6. Finding and funding researchers
7. Partnering with biopharma, government, global
8. Patients as Investors
9. Patients as Biotech Entrepreneurs

Legislative Engagement

Healthcare Architect
In Your Own Healthcare & For Your Community

Thrive in own family, healthcare, team, life • Drivers within your community

Patients as R&D Partners and Drivers
Enabling Patients to determine sharing their data
Data Governance is a Big Deal

Umbrella Institutional Review Board (IRB)

RARE-X provides ALL Patients and Disease communities with Governance support

Data Sharing Preferences Agreement

Patient Rights

Informed Consent Forms

Identified Data

De-Identified Data

Privacy

Country by country regulations & compliance

Data Security
You choose the **type of research** you would like your data to be used for. You must choose one of the following two types of research:

1. **General Research**
   - This is the broadest type of research. When you choose General Research, researchers may use your data for:
     - Health/Medical/Biomedical Research
     - Researchers can access and use your data to learn more about a health condition, its causes, symptoms, progression, and treatments. This type of research could include research on any health condition, even if it is not a rare disease.
     - Other kinds of studies that are not related to health such as:
       - Research on age, race, and ethnicity
       - Research studying traits such as how long people live or how easily they may get sick
       - Research about genetic traits of different populations
       - Studies to develop survey questions to improve research

2. **Health/Medical/Biomedical Research**
   - This type of research is narrower than type 1, General Research. If you choose just Health/Medical/Biomedical Research, your data may be used for fewer types of research studies than if you choose General Research.
     - Your data **may only** be used to learn more about a health condition, its cause, symptoms, progression, and treatments. (Research described in section 1.a above)
     - Your data **will not** be used for other kinds of studies not related to health described in section 1.b. above.

Survey responses are dynamic and can be updated at any time.
Leveraging Data Use Ontologies in a direct to the patient manner

FOR FASTER AND MORE EFFICIENT ACCESS TO DATA

Presentation of the data use options are shown as part of the consent process directly to the patient.

A separation of the represented data uses ontologies to enable the participant.
1. Review the potential data-sharing options multiple times
2. Update the data-sharing preferences outside of the consent document itself.
3. Use these ontologies in a machine-readable manner to speed the access to data in line with patient consent.
# RARE-X Consent Choices DRAFT work

## The Broad Consent Choice

### 2.3 – Choices for DCP

### 2.4 – Choice for Secondary Data Use Terms - Federated

1. Anyone wanting to study data associated with rare disease.

   This category includes all the researchers listed below. It also includes citizen scientists. Citizen scientists are people who research science in their spare time.

2. All researchers with documented proof of professional standing in the research community.

   This category does not include citizen scientists. Saying yes to this category would include researchers who study conditions or symptoms that frequently occur in the general population.

3. Researchers who are known to conduct research on the rare disease that you are afflicted with.

   This group of researchers is more limited than those in number 2. This category includes only researchers who specialize in your rare disease.

4. Only researchers that have had their studies reviewed by an Institutional Review Board (IRB) based on ethical and scientific principles.

   Researchers in this category must present proof of the IRB’s approval of their study before they can access your information for their study.

5. Data repositories[DA2] operated by other organizations may have access to your de-identified information. Allowing this type of sharing helps reduce duplication of efforts. It also would make your de-identified information available to a greater number of researchers.

6. Commercial companies, such as drug companies and biotechnology for research.

   The data cannot be used by for-profit organizations nor for commercial research purposes.
Steps towards using standards at the time of data collection:

Foundation for RARE-X Data Collection Platform
Stakeholder Support:

- Individuals (n=1, undiagnosed)
- Patient Communities (small or large)
- Disease Consortium (body system or symptom): bringing together several disease communities around a symptom (ex. vision or hearing loss)
Standards and guidance consulted by RARE-X

- **Standards**
  - CDISC (Clinical Data Interchange Standards Consortium: FDA standards)
  - Human Phenotype Ontology (Monarch Initiative)
  - Other sources of standardized questions and concepts
    - NIH Metathesaurus
    - NIH Common Data Elements Repository
    - PhenX
    - LOINC, SNOMED, OrphaNet, ICD

- **Guidance**
  - FDA
  - NCATS
  - Scientific community
  - Industry partners
  - Patients
RARE-X: Data Standardization & Data Model

Provide the infrastructure to support comprehensive data collection for analysis.
Gather precise data, map it, layer it, share it.

Platform Development Strategy

1. General core
   • “Head to Toe Survey”

2. Disease core (by domains)
   • HPO-mapped domain-specific data

3. Supplemental disease data
   • Detailed disease-specific data

4. Integrated &/or federated data
   • EMR/EHR, clinical reports, custom curation

5. Exploratory study data
   • Research study-based, raw WGS data
Current RARE-X Focus

- **General Core**
  - A data element that can be consistently collected across studies in any disease or therapeutic area.
  - *RARE-X example*: Demographics
  - *Standards consulted*: CDISC, NIH CDE, NCATS
  - *Status*: RARE-X General Core available with launch

- **Disease Core**
  - A data element specific to a particular disease or therapeutic area.
  - *RARE-X examples*: Skin; Head/Neck; Kidney/Bladder
  - *Standards consulted*: Human Phenotype Ontology, CDISC, NIH CDE
  - *Status*: RARE-X basic (HPO) phenotyping disease core available with launch

- **Supplemental (Custom Surveys)**
  - A data element which is commonly collected in clinical research studies but whose relevance depends upon the study design (i.e., clinical trial, cohort study, etc.) or type of research involved.
  - *RARE-X example*: Homocystinuria-specific dietary questions
  - *Standards consulted*: CDISC, NIH CDE, NIH Metathesaurus, others
  - *Status*: Developing on a case-by-case basis
Data Use Case
Disease Overlap: Symptoms & Disease Biology

Example: Ion Channel Disorders on the RARE-X Platform

- AHC (Alternating Hemiplegia of Childhood)
- CACNA1A
- Charcot-Marie-Tooth
Domain Prioritization- Patient/CG Reported

Domain-based Standardized Modules – Machine Readable, GA4GH Compliant for Data Sharing

<table>
<thead>
<tr>
<th>Domains in RARE-X</th>
<th>Current Domain Development</th>
<th>Domain Expansion &amp; Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics – NIH/RADAR/CDSC</td>
<td>Quality of Life (Patient and Caregiver)</td>
<td>Autoimmune</td>
</tr>
<tr>
<td>General Medical - L1 &amp; L2 (ClinGen)</td>
<td>Medication</td>
<td>Dermatology</td>
</tr>
<tr>
<td>• Health &amp; Development</td>
<td>Medical Encounters</td>
<td>Respiratory</td>
</tr>
<tr>
<td>• Mother’s Pregnancy</td>
<td>Interventional or Medical Diets</td>
<td>Gastrointestinal</td>
</tr>
<tr>
<td>• Growth</td>
<td>Neurodegeneration</td>
<td>Pain</td>
</tr>
<tr>
<td>• Hormone / Endocrine</td>
<td>Neuromuscular</td>
<td>Mental Health</td>
</tr>
<tr>
<td>• Eyes &amp; Vision</td>
<td>Sleep</td>
<td>Musculoskeletal</td>
</tr>
<tr>
<td>• Behavior</td>
<td>Seizures / Epilepsy</td>
<td>Metabolic</td>
</tr>
<tr>
<td>• Skin</td>
<td>Neurodevelopmental</td>
<td>Blood</td>
</tr>
<tr>
<td>• Bone, Cartilage &amp; Connective Tissue</td>
<td>Genetic Testing Report Upload*</td>
<td>Bone</td>
</tr>
<tr>
<td></td>
<td>*Participant uploaded</td>
<td>Hearing / Hearing Loss</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Renal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vision</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rare Cancer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cardiology / Cardiovascular</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Endocrinology</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Medication usage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diet and Nutrition</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mitochondrial</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genetic Data Abstraction &amp; Curation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Surgery</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transplant</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Medical Equipment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diagnostic testing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Treatment/Effectiveness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Disease-specific validated instruments</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Electronic Health Record (EHR) linkages</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Remote Monitoring linkages</td>
</tr>
</tbody>
</table>

Expanding on General Medical Next Layers of Surveys

Mapped to HPO, HL7, OMIM, Orphanet, CDC
Prioritized and Modeled to Generate Research-Grade, Comparable Data

Example: Pediatric Neurodevelopmental Disorders

1. Expert working group formed
2. Symptom domains prioritized
3. PRO Measures landscaped & categorized
4. Measures narrowed for deep review & discussion
5. Final measures confirmed
6. License & implement on RARE-X platform
7. Publish expert working group recommendations

252 Potential Measures
50 Deep Review
13 Implemented on Platform

Multi-Disciplinary Expert Working Group
MD – Roche
PhD – COMBINEDBrain
MD – Colorado Children’s
ScM, CGC – Boston Children’s
PhD – LGS Foundation
PhD – DYRK1A Syndrome International Assn
MD, MS – Weill Cornell Medicine
MA – CACNA1A Foundation
SYNGAP Research Fund
MD, PhD – St. Jude’s
MD, MHA – NIH / NCATS
Data Collection and Use Case: Neurogenetics Clinic (NCRC)

Basket-style Natural History Study across Rare Diseases

- Clinical and research programs launched for multiple rare disorders
- COAs collected
  - Clinician-reported scales
  - Participant-reported scales
  - RARE-X platform participant-reported scales
- Clinician-reported data can be collected on site in a shared data model/map and then transfer to RARE-X to connect data sets for expanded usage
- Future integration planned to allow direct clinician entry in RARE-X
How do ‘validated instruments’ fit in?

- Validated instruments are also known as questionnaires, PROs, or CROs that have been studied extensively using specific scientific criteria and statistical methods that give us confidence that they are reliable and valid in the population used to validate the instruments.

  - Example: an instrument validated in people with cancer may not be applicable to caregivers of children with rare epilepsy.

  - See the following slides for FDA definitions

- RARE-X maintains a library of more than 20,000 validated instruments which can be filtered by domain.
Validated Instruments: Catch-22

- We need to use validated instruments for regulatory purposes

- Validated instruments often force us to use proxy reporting when true ObsRO is not possible (e.g. answering “how they feel” questions on behalf of people unable to communicate)
  - Results in data that may not represent what the patient is actually experiencing.

- Need in the rare disease space when it comes to ”validated” instruments
  - The development of validated instruments that address these challenges
  - The acceptance and qualification of more appropriate instruments into existing standards (CDISC, FDA CRO Qualification)
Can I use a questionnaire that is not ‘validated’ and still be CDISC compliant?
Yes, but tread carefully…

- CDISC has recommendations for sponsors using questionnaires not currently defined in a CDISC QSR supplement to define scales on their own.

- Outside of the context of a specific trial, the use of instruments that have not been reviewed by FDA COA qualification process can result in data that are not considered reliable or valid by the scientific community.

- A list of CRO Qualification submissions can be found here:
  https://www.fda.gov/drugs/clinical-outcome-assessment-coa-qualification-program/clinical-outcome-assessments-coa-qualification-program-submissions
Approaches to Connecting and Making Data Accessible
The Need to Interconnect and Support Other Data

Consortia & Communities

Researchers & Research Portal
[Supporting meta data access and analysis]

Other Data

Sub Studies Include:
- Biopharma-sponsored studies/surveys
- Embargoed Period
- Other Novel Disease-Specific Data
- Clinician Reported/NH Study Data

Federated Data:
[building an interconnected data ecosystem, connecting disparate data sets]
Inverting the Model of Data Sharing

**Traditional approach**
Bring data to researchers

- Discourages shared research
- Data sharing = data copying
- Few audit controls
- Huge infrastructure needed
- Siloed compute

**Cloud-centric approach**
Bring researchers to data

- Facilitates collaboration
- Cost
- Threat Detection and auditing
- Increased accessibility
- Shared & elastic compute

[Diagram showing the comparison between traditional and cloud-centric approaches.]
Data Generation, Alignment, Federation

**Partner Data & Other Data**
(eg: EHR, Academia, etc)

**Research Portal Supporting Federated Data**
(eg: Global registries, natural history studies)

**Open Science Platform Enables:**
- Collaborative analysis of previously locked data sets
- Cross-disorder comparative research
- Accelerated therapeutic research path for rare diseases

**RAREX**
Rare Disease Communities actively generate & upload data

**Ontology mapping**
**Standardization**
**Curation**

**Natural history studies**

**Researchers/Clinicians**

**Biopharma**

**Patient Communities**
Partner/Stakeholder Ecosystem

RARE-X has built a fully integrated platform to support patients as partners in research and has also developed a service model to support biopharma & researchers. A turn-key comprehensive solution for patients.

**Patient Advocates and Orgs**
- Patient Owned and Stewarded Data
- Technology and Platform for Data Collection and Sharing
- All Data Governance & Consents
- Robust Research Ready Surveys
- Patient Engagement Team
- Education & Marketing Support

**Researchers**
- In-Depth Engagement with Patient Organizations and development of registries
- Natural History Studies including Clinician Reported Data
- Sponsored Studies
- Federated Learning and Data Connection for deeper analysis (ie. C-Path RD-CAP)

**BioPharm**
- Sponsored Studies
- Federated Learning and Data Connection for deeper analysis
- Data sharing post-study completion
- Clinical trial readiness surveys
- Patient identification for recruitment into clinical trials

Supporting basic research to: help characterize disease, create critical baseline data, future disease concept and progression models. Building a funnel and rigorous repeatable process for patient advocacy organizations.
What Is RARE-X?

- RARE-X is a program of Global Genes created to accelerate rare disease research, treatments, and cures by removing barriers for data collection and sharing.
- RARE-X is a platform to collect, connect, and share data.

RARE-X is not a replacement for any current research or clinician-sponsored patient registries, but rather a prepared collaborator and partner. Ready to meet data where it is and enable its access, in whatever way it can compliently be used.
RARE-X: Facilitating Open Science for Progress with Patient-driven Data

RARE-X Provides

- A Platform for collecting structured patient data (including clinical, PRO, molecular, & study data)
- An open science platform to facilitate sharing of large high quality data sets to accelerate therapeutic research
- AND-
- A full-service ongoing patient engagement and program management service to ensure participation & success

RARE-X is a Nonprofit Health Technology & Patient Advocacy Company
Driving Success through Data Structure & Collaboration
Thank you.

Together, we are powering progress for rare diseases.
Q&A

John Concato, MD, MS, MPH
Ramona Walls, PhD
Vanessa Vogel-Farley, BA, BS
Break

Upcoming Virtual FDA Workshop

FDA’s CDER, CBER, and Duke-Margolis Center for Health Policy Host
Rare Disease Endpoint Advancement Pilot Program Workshop: Novel Endpoints for Rare Disease Drug Development

June 7 and 8, 2023; 1-5 pm
Link in the Chat
Session 2:
Use of Data Sources to Inform Rare Disease Drug Development

Moderator: Christine Nguyen, MD
Deputy Director
Office of Rare Diseases, Pediatrics, Urologic and Reproductive Medicine,
Office of New Drugs, Center for Drug Evaluation and Research, FDA
Advancement of Drug Development Tools for Polycystic Kidney Disease (PKD) as Told Through the PKD Outcomes Consortium Story

CDER-JHU CERSI Rare Disease Workshop | May 2, 2023

Sorin Fedele, PhD, MBA
Executive Director, Polycystic Kidney Disease Outcomes Consortium (PKDOC)
Critical Path Institute (C-Path)
Content

• C-Path Overview

• PKDOC Background and Impact

• PKDOC 2.0
How Does it Work?

Data and expertise sharing

Core Competencies
- Biomarkers
- Clinical Outcome Assessments
- Modeling and Analytics
- Regulatory/Development Science

Unmet Medical Need

Concentration Area
- Neuroscience
- Immunology and Inflammation
- Infectious Diseases
- Safety Science
- Rare/Orphan Diseases
- Pediatrics

Industry Partners
- A
- B
- C

C-Path
Precompetitive Neutral Ground

Patients

Academia

EMA
FDA
PMDA

PKD
POLYCYSTIC KIDNEY DISEASE OUTCOMES CONSORTIUM

3
Clinical Datasets Contributed to C-Path

Clinical Data
- Studies: 380
- Subjects: 456,443

Nonclinical Data
- Studies: 148
- Subjects: 11,084

Note: Studies currently undergoing curation are only counted in Total Studies until evaluated.
Regulatory Successes in Drug Development Tools

**FDA**
- 7 Qualification Decisions
- 7 Letters of Support
- 1 Fit-For-Purpose Endorsement

**EMA**
- 8 Qualification Opinions
- 8 Letters of Support

**PMDA**
- 1 Qualification Decisions

Global endorsement of actionable solutions accelerates and de-risks medical product development.
Content

• C-Path Overview

• PKDOC Background and Impact

• PKDOC 2.0
PKDOC Team

C-Path:

- **Sorin Fedele, PhD, MBA**
  Executive Director
- **Wendy Vanasco**
  Senior Project Manager
- **Kitty Bogy**
  Senior Project Coordinator

Co-Directors:

- **Frank Czerwiec, MD, PhD**
  Sparrow Pharmaceuticals
- **Ronald Perrone, MD**
  Tufts University School of Medicine
- **TBD**
  PKD Foundation
What We Do

- Foster development of new evaluation tools to inform medical product development and regulatory decision-making
- Convene scientific consortia of industry, academia, and government for sharing of data/expertise

The best science
- The broadest experience
- Active consensus building
- Shared risks and costs

- Enable iterative EMA/FDA/PMDA participation in developing new methods to assess the safety and efficacy of medical products
- Obtain official regulatory endorsement of novel methodologies and drug development tools
ADPKD: Progression of Kidney Disease

GFR = glomerular filtration rate
ADPKD

- Most common hereditary renal disease (1:400 to 1:1,000)
- Autosomal dominant inheritance
- Genetically heterogeneous
  - \(PKD1\) (16p13.3) (~77%)
  - \(PKD2\) (TRPP2) (4q21-23) (~15%)
  - No mutation detected (8%)
- Affects all nationalities and ethnic groups (~12.5 M worldwide)
- No common or recurrent mutations

Polycystic Kidney Disease: Lack of Biomarkers Discouraged Therapeutic Development

The Challenges

- Heterogeneous and slow progressing disease requires long trials and challenging endpoints
- Finding clinical endpoint(s) or an accepted surrogate for measuring disease progression early in the course of the disease where kidney function is largely preserved
- Designing a clinical trial and acceptable post marketing study to use FDAs Accelerated Approval pathway

Initial Mission of PKDOC

1. Develop standard common data elements specific to ADPKD
2. Create new integrated patient-level database from existing multiple, longitudinal, well-characterized and varied data sources
3. Develop quantitative biomarker dynamics and disease progression joint model
4. Incorporate results of contemporary trials into database
5. Generate scientific consensus on the utility and reliability of TKV as a biomarker and clinical endpoint for the progression of ADPKD
6. Submit qualification package on TKV to FDA and EMA for review and possible designation as “qualified for use” in drug development
Data Sources

<table>
<thead>
<tr>
<th>Sources of Data:</th>
<th>Information</th>
</tr>
</thead>
</table>
| **University of Colorado Registry** | • Time Frame: 1985 – 2004  
• Number of individuals: 5,684 individuals from 1228 families with ADPKD, \( \approx 107,996 \) patient-years  
• Structure: Long-term registry funded by NIH  
• Process: Structured evaluation at the University of Colorado General Clinical Research Center; irregular visit and TKV measurement interval  
• Outcomes: Information from 1112 participants, 648 women (58.3%) and 464 men (41.7%), has been mapped to the CDISC SDTM standard. There were 165 deaths and 342 ESRD events with timing information |
| **Mayo Clinic Registry** | • Time frame: 1984 - present; analysis limited to those with electronic records, after mid-90s  
• Number of individuals: 2,871 patients with ADPKD, \( \approx 34,452 \) patient-years  
• Structure: Encounter for clinical care at Mayo Clinic, Rochester, MN  
• Process: Comprehensive data collection through clinical care; irregular visit and TKV measurement interval  
• Outcomes: The Mayo Clinic has supplied CDISC SDTM mapped data on 1010 participants including 607 women (60.1%) and 403 men (39.9%). There were 68 deaths and 198 ESRD events with timing information |
| **Emory University** | • Time frame: 1998 - 2014  
• Number of individuals: 700 individuals from approximately 400 families, \( \approx 11,200 \) patient-years  
• Structure: Two day visit at GCRC as part of longitudinal observational program supported by the Polycystic Kidney Disease Foundation (COHORT Study).  
• Process: Structured evaluation at Emory University General Clinical Research Center; irregular visit and TKV measurement interval  
• Outcomes: Information from 376 participants, 229 women (60.0%) and 147 men (39.1%), has been mapped to the CDISC SDTM standard. There were eight deaths and 121 ESRD events with timing information |
| **Consortium for Radiologic Imaging Studies in PKD (CRISP I and 2)** | • Time frame: 2001 - 2010  
• Number of Individuals: 241, \( \approx 2,169 \) patient-years  
• Structure: multicenter, prospective, longitudinal study of the natural history of ADPKD  
• Process: Regular visits with TKV measurements yearly through first 3 years and less frequent thereafter. Comprehensive data collection  
• Outcomes: All data from both CRISP I and II were converted to a CDISC SDTM structure. There were no deaths or ESRD events during CRISP I. In CRISP II there were two deaths and eight ESRD events with data available regarding the start of ESRD |

→ Total of 2355 patients with at least one TKV measurement (all modalities) in the database were available. Overall, the analysis dataset included 1140 patients of which 361 (31.7%) patients had a 30% worsening of eGFR (two measurements 30% lower than baseline).
TKV Qualifications from FDA and EMA

A. Use Statement

This guidance provides qualification recommendations for the use of TKV, measured at baseline, as a prognostic enrichment biomarker to select patients with ADPKD at high risk for a progressive decline in renal function (defined as a confirmed 30% decline in the patient’s estimated glomerular filtration rate (eGFR)) for inclusion in interventional clinical trials. This biomarker may be used in combination with the patient’s age and baseline eGFR as an enrichment factor in these trials.
The Value of Enrichment

Predicted event rate in placebo arm over 3 years, **number needed to enroll** and **number needed to treat to get one event** using the best fit models with and without TKV.

<table>
<thead>
<tr>
<th></th>
<th>Model without TKV</th>
<th>Model with TKV, using added criterion of TKV &gt; 1 L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted event rate in placebo arm over 3 years</td>
<td>0.091</td>
<td>0.110</td>
</tr>
<tr>
<td>Number needed to enroll†</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Number needed to screen</td>
<td>13</td>
<td>25</td>
</tr>
</tbody>
</table>

Assumes entry criteria of eGFR > 50 mL/min per 1.73 m² and age between 20 and 50 years.
Scoring PKD: Imaging Classification of ADPKD

- Tool for inputting htTKV and age to classify patients into groups A-E
- Classification predicts renal survival
- Useful to optimize patient selection for enrollment into clinical trials and for treatment

Maria V. Irazabal et al. JASN 2015;26:160-172
PKD1 Mutation Type Influences Renal Survival

PKD1 Mutation Type Influences Renal Survival

Cornec-Le Gall E et al. JASN 2013;24:1006-1013
Scoring PKD: PRO-PKD

Table 3. Multivariate Cox analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Patients (n)</th>
<th>HR (95% CI)</th>
<th>95% CI from Bootstrap Analysis</th>
<th>P Value</th>
<th>Points for PRO-PKD Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>541</td>
<td>1.55 (1.29 to 1.88)</td>
<td>1.27 to 1.89</td>
<td>&lt;0.001</td>
<td>0</td>
</tr>
<tr>
<td>Male</td>
<td>432</td>
<td>1.27 (1.27 to 1.89)</td>
<td>1.27 to 1.89</td>
<td>&lt;0.001</td>
<td>1</td>
</tr>
<tr>
<td>Hypertension before age 35 yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>679</td>
<td>2.11 (1.71 to 2.61)</td>
<td>1.71 to 2.62</td>
<td>&lt;0.001</td>
<td>2</td>
</tr>
<tr>
<td>Yes</td>
<td>294</td>
<td>1.73 (1.38 to 2.18)</td>
<td>1.35 to 2.24</td>
<td>&lt;0.001</td>
<td>2</td>
</tr>
<tr>
<td>≥1 urologic event before age 35 yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>734</td>
<td>2.27 (1.57 to 3.28)</td>
<td>1.61 to 3.18</td>
<td>0.002</td>
<td>2</td>
</tr>
<tr>
<td>Yes</td>
<td>239</td>
<td>4.75 (3.41 to 6.60)</td>
<td>3.63 to 6.60</td>
<td>&lt;0.001</td>
<td>4</td>
</tr>
<tr>
<td>Mutation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PKD2</td>
<td>186</td>
<td>1.73 (1.38 to 2.18)</td>
<td>1.35 to 2.24</td>
<td>&lt;0.001</td>
<td>2</td>
</tr>
<tr>
<td>PKD1 nontruncating</td>
<td>239</td>
<td>2.77 (1.57 to 3.28)</td>
<td>1.61 to 3.18</td>
<td>0.002</td>
<td>2</td>
</tr>
<tr>
<td>PKD1 truncating</td>
<td>548</td>
<td>4.75 (3.41 to 6.60)</td>
<td>3.63 to 6.60</td>
<td>&lt;0.001</td>
<td>4</td>
</tr>
</tbody>
</table>

95% CI, 95% confidence interval.

The AUC for the PRO-PKD score is 0.84; It is 0.79 for the genetic score alone.

Emilie Cornec-Le Gall et al. JASN 2016;27:942-951
PKDOC Impact

- Development of a CDISC therapeutic area user guide (TAUG) for PKD to collate data from several clinical patient registries and observational studies of ADPKD patients
- Successful qualification of total kidney volume (TKV) as prognostic biomarker to select patients for clinical trials of new therapies for ADPKD is a key milestone for the consortium
- TKV has been designated as a reasonably likely surrogate endpoint and therefore could be used in an FDA accelerated approval process, but an acceptable plan for a post-marketing confirmatory trial would be required
- Otsuka’s drug JYNARQUE® (Tolvaptan) was designated as the first FDA-approved treatment for PKD; Although it was not a direct output of PKDOC, the consortium was a significant positive influence over many years in this success story
Lessons Learned

• While TKV had been used as part of development programs, the TKV qualification effort quantified the amount of information that “was added” by using TKV to enrich a trial population

• Qualification served as a steppingstone to more meaningful discussions about the use of TKV as a reasonably likely surrogate and potential endpoints for approval

• Registry data can be critical for establishing the value of a biomarker as a tool in drug development (with inherent challenges associated with using and interpreting the data)
Content

• C-Path Overview

• PKDOC Background and Impact

• PKDOC 2.0
PKDOC 2.0 Goals and Objectives

**Biomarkers**
- Understanding of biomarker opportunities across all PKD stakeholders
- An evaluation of the maturity of the biomarkers
- Identification of biomarkers ready for qualification or IVD acceptance

**Data Sharing**
- Define mechanism for the housing of current PKD datasets to RDCA-DAP
- Obtain new data from other industry members including updated registry data and data from clinical trials

**CDISC Standards (TAUG)**
- Conduct a review of CDISC elements for standardization of data for regulatory submissions and ensure optimal clinical trial data collection

**COAs or PROs**
- PROPKD Score and other PROs
- Assess the potential of patient-focused drug development initiatives for ARPKD

**TKV Modeling**
- Further development of drug-trial-disease models and simulation tools to optimize clinical trial design
- Develop a clinical trial simulation (CTS) tool

PKDOC 2.0
Codify PKDOC as a full consortium to drive multiple drug development tools towards regulatory endorsement
Innovation Through Data Sharing

**Academia**
- Improves their research
- Understand disease course/variance
- Understand/develop biomarkers/endpoints
- Visibility of data and research, collaboration
- Publish more/better papers

**Industry**
- Design more effective trials
- Understand disease course/variance
- Understand/develop biomarkers/endpoints

**Patients/Patient Groups**
- Faster drug development
- Understand disease course/variance
- Visibility to industry
- Drive collaboration
PKD Modeling/CTS Tool Roadmap

1. Current Datasets
   - University of Colorado
   - Emory University
   - Mayo Clinic
   - CRISP
   - HALT

2. Incoming Datasets
   - ALADIN 1
   - TAME

3. Investigating strategies to leverage industry-led RCT data

Our Impact
- Novel clinical trials
- Optimize trial design, patient and endpoints selection
- Determine primary statistical analyses
- Sample size calculation
- Enable decision making

ACTIONABLE DRUG DEVELOPMENT SOLUTIONS

C-Path Data Repository
- Investigational Database
- Aggregated Data

User Friendly, Secure Cloud Interface

Data Vault
- Patient-Level Data
  - Industry Clinical Trials
  - Observational Studies
  - Natural History Studies
  - RWD
  - ...and more!

Curation
- De-identified HIPAA compliant

Quality Control
- CDISC Quality Control
- Investigational Database
- Aggregated Data

Incoming Data Storage
- Standardization

US Food and Drug Administration
European Medicines Agency
Regulatory endorsement

C-Path Quantitative Medicine Program

Approved Users
- (Contributor-permitted data sets only)

Other Data Platforms
- (Contributor-permitted data sets only)
The Envisioned Outcome: Clinical Trial Simulations

The developed model is intended to be used as a basis in a clinical trial simulation tool

Such a tool is intended to inform clinical trial design by computing trial power based on user chosen information:

1) Inclusion/exclusion criteria
2) Enrichment strategies
3) Trial duration and sample size
4) Support design of accelerated approval programs
The Envisioned Outcome: Clinical Trial Simulations

AD Trial Simulator

Mild-to-Moderate Alzheimer Disease Clinical Trial Simulator (beta v2.0)

Plot of Trial Design

Median ADAS-cog11 (95% Confidence Interval)

Time (Weeks)

Indicators
- Treatment
- Placebo
PRO-Focused Approaches

• Focus on patient-reported outcomes (PRO) as an avenue to inform medical product development

• Both ADPKD and ARPKD represent areas of unmet need for PRO development

• Use ARPKD as a case study for an externally-led patient focused drug development (EL-PFDD) project
ARPKD EL-PFDD Objectives

Broad objective of the meeting are to inform the FDA and other stakeholders (e.g., drug developers) on:

- Patients’ and families’ experiences and perspectives regarding symptoms and burdens of ARPKD and its impact on daily living
- Factors that may influence patients’ and families’ decision making on entering clinical trials, including
  - Endpoints
  - Trials conducted under Accelerated Approval Program
- Current medical management of ARPKD, patient/family experiences with treatment and their aspirations for new treatments

Who does this meeting benefit?

**FDA**
- Gain understanding of what it’s like to live with ARPKD
- Learn about side effects and risks patients are willing to accept
- Hear patients’ needs for new drugs and preferences for clinical trials

**Patients**
- Know that the FDA and industry stakeholders have heard their voices
- Hearing other patients’ experiences and needs to validate symptoms and feelings in order to better self-advocate

**Patient Advocacy Groups**
- Identify additional needs for patient education and advocacy
- Increase public awareness through gained knowledge of ARPKD
- Create greater connections with patients and their peers

**Industry**
- Gain insights into the major concerns of patients to help develop treatments and optimize clinical trial design
- Learn about symptoms and side effects to help develop drugs that matter to patients
Value to PKDOC 2.0 Stakeholders

• Regulatory acceptance
  – Better understanding of disease and application of biomarkers across all stakeholders including health authorities

• Rapid implementation of biomarkers in clinical trials
  – Accepted under IND vs qualified

• Patient stratification and disease monitoring biomarkers lead to efficient clinical trials, faster approvals

• Change patient journey—precision medicine
Thank you!
Leveraging patient engagement and real-world data to inform rare disease drug development

FDA CDER–JHU CERSI Rare Disease Workshop
2 May 2023
Despite advances in research and technology, relatively few orphan drugs are approved each year.

- **Disease Discovery**
  - Sequencing costs dropped 10x in 5 yrs;
  - 80% of rare diseases are genetic

- **Research**
  - 850+ rare disease biotech programs

- **Development**
  - 70% of rare drugs are in early development.

- **Approval**
  - Only 20 rare disease drugs were approved in 2022
Rare disease drug development is uniquely challenging

- Small patient number geographically spread across the globe
- Many specialties / institutions involved in patient care
- Scarcity of high-quality data in orphan populations
- Natural history rarely understood; limited longitudinal data
- Burden of illness difficult to quantify & characterize
- Appropriate clinical outcome measures are often unclear
- Studies are clinically & ethically difficult to design & execute
- Deep engagement of patient communities is critical
Real-world evidence has the potential to address key questions across the drug development lifecycle

<table>
<thead>
<tr>
<th>Question</th>
<th>Pre-clinical</th>
<th>Ph 1 &amp; 2</th>
<th>Ph 3 &amp; Launch</th>
<th>Post-Launch</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the disease epidemiology and unmet need?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the patient journey from diagnosis to treatment?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What are the characteristics of the patient population?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How feasible is the clinical protocol?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the safety &amp; effectiveness in the real world?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How is the product used in the real world?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Real-world evidence has the potential to address key questions across the drug development lifecycle.

<table>
<thead>
<tr>
<th>Question</th>
<th>Pre-clinical</th>
<th>Ph 1 &amp; 2</th>
<th>Ph 3 &amp; Launch</th>
<th>Post-Launch</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the disease epidemiology and unmet need?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the patient journey from diagnosis to treatment?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What are the characteristics of the patient population?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How feasible is the clinical protocol?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the safety &amp; effectiveness in the real world?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How is the product used in the real world?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Integrating the patient voice is critical to a robust real-world data strategy.

Real-World Data Sources

- Claims / Billing Data
- Patient-Reported Data
- Structured EHR Databases
- Unstructured Clinical Notes
Integrating the patient voice is critical to a robust real-world data strategy.
Integrating the patient voice is key to answering the big questions in clinical trial planning.

**Who?**
- Characterize the population
- Evaluate I/E criteria feasibility

**What & When?**
- Characterize unmet need
- Determine appropriate outcomes and endpoints

**Where?**
- Evaluate recruitment approaches
- Identify suitable trial sites
AllStripes serves as the nexus of patient engagement and real-world data generation

Patients and caregivers can **sign up and e-consent in minutes**; accounts may be created for deceased patients.

The umbrella research consent allows use of de-identified data for **minimal risk research, survey, and recontact of patients over time**.

AllStripes collects, structures, and analyzes multimodal clinical data from across the patient journey **at no cost to participants**.

**Ongoing engagement, insights, and communications** shared about research programs.

Records and data collected from over 4,000 healthcare facilities in the US, Canada, and UK.
Who, What, & When: Characterizing Unmet Need and the Patient Journey
Case study: Genetic epilepsy natural history

SPONSOR: Sponsor A, a biopharmaceutical company

STAGE: Pre-IND

CONDITION: Condition B, a rare, severe epilepsy characterized by seizures that begin in infancy

CHALLENGE: Lack of understanding of natural history and progression of Condition B. Sponsor A needed to better characterize the patient journey to inform clinical trial design.

OUR SOLUTION: Natural history study to better understand needs of the patient community and inform clinical trial outcome and endpoint selection.

METHODS: Participant surveys & clinical data abstracted from patient medical records

RESULTS:

<table>
<thead>
<tr>
<th></th>
<th>250+</th>
<th>16,200+</th>
<th>235+</th>
<th>12,600+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>facilities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>documents</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Years</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>of clinical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>follow-up</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data points</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>abstracted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Longitudinal history with detailed context is critical to understand the complete patient journey.
Partnering with families is key to understanding unmet need

Condition B, n = 22

First symptom
- Developmental delays
- Gastroesophageal reflux
- Hypotonia
- Visual impairment

Symptom that most greatly affects quality of life
- Hypotonia
- Other
- Seizures
- Developmental delays
Current AllStripes symptoms database

831 completed surveys across 46 conditions

Example: Dermatomyositis (n = 52)
Who, What, & When: Characterizing the Patient Population
Case study: Characterizing a rare metabolic syndrome

SPONSOR: Sponsor C, a research institution exploring commercialization

STAGE: Pre-clinical

CONDITION: Condition D, a rare inborn error of metabolism

CHALLENGE: Lack of understanding of Condition D manifestations, including neurological signs and behavioral symptoms beginning in childhood. Future trials will require appropriate instruments for measuring these symptoms.

OUR SOLUTION: Natural history study designed in partnership with Sponsor C and Advocacy Group E

METHODS: Surveys & clinical data abstracted from patient medical records to capture longitudinal disease manifestations.

RESULTS:

- 250+ Medical facilities
- 13,500+ Clinical documents
- 6800+ Data points abstracted
- 2500+ Survey data points collected
Involving all stakeholders in instrument development is key to success

**Sponsor + Advocate KOL**
Co-develop comprehensive list of behavioral symptoms and associated data of interest

**AllStripes Research Team**
Develop and test survey instrument on proprietary patient platform, with feedback from sponsor and advocate KOL

**Pilot Participant Group**
Complete draft instrument on AllStripes platform and provide feedback on content, language, and presentation

**All Participants**
Complete final instrument longitudinally to track response consistency and disease progression
Caregiver surveys collected extensive data on Condition D behavioral symptoms

<table>
<thead>
<tr>
<th>Behavior Categories</th>
<th># Behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Aggression</td>
<td>4</td>
</tr>
<tr>
<td>Behavior Category 2</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 3</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 4</td>
<td>11</td>
</tr>
<tr>
<td>Behavior Category 5</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 6</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 7</td>
<td>2</td>
</tr>
<tr>
<td>Behavior Category 8</td>
<td>2</td>
</tr>
<tr>
<td>Behavior Category 9</td>
<td>2</td>
</tr>
<tr>
<td>Other Behaviors (free-text)</td>
<td>–</td>
</tr>
</tbody>
</table>
Caregiver surveys collected extensive data on Condition D behavioral symptoms

<table>
<thead>
<tr>
<th>Behavior Categories</th>
<th># Behaviors</th>
<th>Behaviors Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Aggression</td>
<td>4</td>
<td>● Hitting / kicking</td>
</tr>
<tr>
<td>Behavior Category 2</td>
<td>3</td>
<td>● Scratching</td>
</tr>
<tr>
<td>Behavior Category 3</td>
<td>3</td>
<td>● Biting</td>
</tr>
<tr>
<td>Behavior Category 4</td>
<td>11</td>
<td>● Grabbing</td>
</tr>
<tr>
<td>Behavior Category 5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Behavior Category 6</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Behavior Category 7</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Behavior Category 8</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Behavior Category 9</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Other Behaviors (free-text)</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>
Caregiver surveys collected extensive data on Condition D behavioral symptoms

<table>
<thead>
<tr>
<th>Behavior Categories</th>
<th># Behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Aggression</td>
<td>4</td>
</tr>
<tr>
<td>Behavior Category 2</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 3</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 4</td>
<td>11</td>
</tr>
<tr>
<td>Behavior Category 5</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 6</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 7</td>
<td>2</td>
</tr>
<tr>
<td>Behavior Category 8</td>
<td>2</td>
</tr>
<tr>
<td>Behavior Category 9</td>
<td>2</td>
</tr>
<tr>
<td>Other Behaviors (free-text)</td>
<td>–</td>
</tr>
</tbody>
</table>

Data Points Collected
- Age of onset
- Consistency
- Triggers
- Frequency
- Intensity
- Severity
- Mitigation strategies
Caregivers reported additional behaviors not assessed in the survey

<table>
<thead>
<tr>
<th>Behavior Categories</th>
<th># Behaviors</th>
<th># Additional Behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Aggression</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Behavior Category 2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Behavior Category 3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Behavior Category 4</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Behavior Category 5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Behavior Category 6</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Behavior Category 7</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Behavior Category 8</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Behavior Category 9</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Other Behaviors (free-text)</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>
Caregivers reported additional behaviors and behavior categories not assessed in survey

<table>
<thead>
<tr>
<th>Behavior Categories</th>
<th># Behaviors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Aggression</td>
<td>4</td>
</tr>
<tr>
<td>Behavior Category 2</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 3</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 4</td>
<td>11</td>
</tr>
<tr>
<td>Behavior Category 5</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 6</td>
<td>3</td>
</tr>
<tr>
<td>Behavior Category 7</td>
<td>2</td>
</tr>
<tr>
<td>Behavior Category 8</td>
<td>2</td>
</tr>
<tr>
<td>Behavior Category 9</td>
<td>2</td>
</tr>
<tr>
<td>Other Behaviors (free-text)</td>
<td>–</td>
</tr>
</tbody>
</table>

“Other” Findings

- Additional behavior category involving eating / feeding identified
- 3+ additional behaviors identified that do not fit cleanly into an established category
Case study: Characterizing a rare metabolic syndrome

**SPONSOR:** Sponsor C, an academic research institution with interests in commercialization

**STAGE:** Pre-clinical

**CONDITION:** Condition D, a rare inborn error of metabolism

**CHALLENGE:** Lack of understanding of Condition C manifestations, including neurological signs and behavioral symptoms beginning in childhood. Future trials will require appropriate instruments for measuring these symptoms.

**OUR SOLUTION:** Natural history study designed in partnership with Sponsor C and Advocacy Group E

**METHODS:** Custom behavioral survey & clinical data abstracted from patient medical records to capture longitudinal disease manifestations.

**RESULTS:**

<table>
<thead>
<tr>
<th>250+</th>
<th>13,500+</th>
<th>6800+</th>
<th>2500+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical facilities</td>
<td>Clinical documents</td>
<td>Data points abstracted</td>
<td>Survey data points collected</td>
</tr>
</tbody>
</table>
Who:
Evaluating I/E Criteria
Case study: Recruiting for a pivotal trial in adult-onset autoimmune neuropathy

SPONSOR: Sponsor F, a biopharmaceutical company

STAGE: Pivotal trial

CONDITION: Condition G, a rare immune-related neurological condition that causes weakness and reduced sensation in the arms and legs

CHALLENGE: Recruiting participants for a large, multi-site pivotal trial

APPROACH: Pre-screen patients using data collected from medical records

RESULTS:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>132</td>
<td>112</td>
<td>&lt; 5</td>
</tr>
<tr>
<td>Consented participants</td>
<td>Participants pre-screened</td>
<td>Patients connected to site</td>
</tr>
</tbody>
</table>
Sponsors should carefully consider the characteristics of a population when selecting I/E criteria.

<table>
<thead>
<tr>
<th>Reasons for Failing Pre-screening</th>
<th># Patients (%) / 112</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes diagnosis</td>
<td>9 (8%)</td>
</tr>
<tr>
<td>History of malignancy</td>
<td>8 (7%)</td>
</tr>
</tbody>
</table>

- 1 in 10 Americans
- 15–20% of individuals with Condition G
Sponsors should carefully consider the characteristics of a population when selecting I/E criteria.

<table>
<thead>
<tr>
<th>Reasons for Failing Pre-screening</th>
<th># Patients (%) / 112</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes diagnosis</td>
<td>9 (8%)</td>
</tr>
<tr>
<td>History of malignancy</td>
<td>8 (7%)</td>
</tr>
</tbody>
</table>

~ 1 in 2 people over a lifetime
Where:
Identifying trial sites
Most participants are interested in future clinical trials

Previous trial participation

- No: 342 (73.4%)
- Yes: 109 (23.4%)
- Not sure: 15 (3.22%)

Interest in clinical trial participation

- Extremely interested: 231 (49.6%)
- Very interested: 116 (24.9%)
- Moderately interested: 64 (13.7%)
- Slightly interested: 33 (7.08%)
- Not at all interested: 22 (4.72%)

n = 466
Distance to study sites is participants’ most common concern about potential clinical trial enrollment

Barriers to trial participation

- Previous negative experiences with clinical trials
- None of the above
- Time required for clinical visits
- Concern about being placed in the placebo group
- Failure to meet inclusion criteria
- Potential need to stop current treatment in order to participate
- Potential risks of study participation
- Financial burden of trial participation
- Potential negative side effects of experimental treatment
- Distance to study center

Percent of patients

n = 466
Average distance to nearest trial site illustrates potential travel burden for participants

<table>
<thead>
<tr>
<th>Condition Category</th>
<th>Median Distance (miles)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vascular</td>
<td>919</td>
</tr>
<tr>
<td>Other systemic</td>
<td>812</td>
</tr>
<tr>
<td>Neurological</td>
<td>965</td>
</tr>
<tr>
<td>Lysosomal storage</td>
<td>333</td>
</tr>
<tr>
<td>Tumor/lymphatic</td>
<td>539</td>
</tr>
<tr>
<td>Other metabolic</td>
<td>431</td>
</tr>
<tr>
<td>Hematological</td>
<td>474</td>
</tr>
<tr>
<td>Mitochondrial</td>
<td>106</td>
</tr>
<tr>
<td>Other systemic</td>
<td>281</td>
</tr>
<tr>
<td>Mitochondrial</td>
<td>348</td>
</tr>
<tr>
<td>Other systemic</td>
<td>347</td>
</tr>
<tr>
<td>Mitochondrial</td>
<td>342</td>
</tr>
<tr>
<td>Other systemic</td>
<td>242</td>
</tr>
<tr>
<td>Mitochondrial</td>
<td>135</td>
</tr>
</tbody>
</table>

*distance not available for condition categories with 0 interventional or observational studies
**distance not shown for categories with fewer than 10 patients with conditions covered by available studies

n = 903
Case study: Recruiting for a pivotal trial in adult-onset autoimmune neuropathy

**SPONSOR:** Sponsor F, a biopharmaceutical company

**STAGE:** Pivotal trial

**CONDITION:** Condition G, a rare immune-related neurological condition that causes weakness and reduced sensation in the arms and legs

**CHALLENGE:** Recruiting participants for a large, 8-site pivotal trial

**APPROACH:** Pre-screen patients using data collected from medical records

**RESULTS:**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>132</strong></td>
<td></td>
<td><strong>112</strong></td>
<td></td>
</tr>
<tr>
<td>Consented</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>participants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>&lt; 5</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>connected to</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>site</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sponsors should select trial sites with patient geography in mind.
Sponsors should select trial sites with patient geography in mind.

Minimum Distance between Participants and Any Condition G Trial Site

November 2019  71 patients, 6 trials, 15 trial sites

<table>
<thead>
<tr>
<th>Distance Range</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;50 miles</td>
<td>11.3%</td>
</tr>
<tr>
<td>51-200 miles</td>
<td>32.0%</td>
</tr>
<tr>
<td>201-500 miles</td>
<td>43.7%</td>
</tr>
<tr>
<td>&gt;500 miles</td>
<td>12.7%</td>
</tr>
</tbody>
</table>

Nov 2019 – Feb 2020

- Targeted recruitment within 200 mi of sites for large trial
- 10 trial sites added
Sponsors should select trial sites with patient geography in mind.

Nov 2019 – Feb 2020
• Targeted recruitment within 200 mi of sites for large trial
• 10 trial sites added

Minimum Distance between Participants and Any Condition G Trial Site

**November 2019**  71 patients, 6 trials, 15 trial sites

- <50 miles: 11.3%
- 51-200 miles: 32.0%
- 201-500 miles: 43.7%
- >500 miles: 12.7%

**February 2020**  111 patients, 6 trials, 25 sites

- <50 miles: 22.5%
- 51-200 miles: 44.1%
- 201-500 miles: 21.6%
- >500 miles: 11.7%
Sponsors should select trial sites with patient geography in mind.

### Minimum Distance between Participants and Any Condition G Trial Site

**November 2019**  
71 patients, 6 trials, 15 trial sites
- <50 miles: 11.3%  
- 51-200 miles: 32.0%  
- 201-500 miles: 43.7%  
- >500 miles: 12.7%

**February 2020**  
111 patients, 6 trials, 25 sites
- <50 miles: 22.5%  
- 51-200 miles: 44.1%  
- 201-500 miles: 21.6%  
- >500 miles: 11.7%

**Nov 2019 – Feb 2020**
- Targeted recruitment within 200 mi of sites for large trial
- 10 trial sites added

**57% increase**
Geographic distribution of US lysosomal storage disorder (LSD) cohort vs. prospective COEs

9 LSDs, 151 participants
Geographic distribution of US lysosomal storage disorder (LSD) cohort vs. prospective COEs

9 LSDs, 151 participants

Multidisciplinary care
Peer-reviewed publications
Clinical trial participation
Presence of a metabolic genetics clinic
Travel to prospective COEs would entail a substantial burden.

- Median time to nearest prospective COE: 1.68 hours
- Median distance to nearest prospective COE: 93.4 miles
Travel time to prospective LSD COEs varies by region

Median distance from nearest prospective COE

- West
- East
- South
- Central
- North
- Mountain
- East
- North
- Central
- Pacific
- South
- Atlantic
- New
- England
- Middle
- Atlantic
- West
- South
- Central

US Census Divisions

Miles

0 20 40 60 80 100 120 140 160

West North Central
East South Central
Mountain
East North Central
Pacific
South Atlantic
New England
Middle Atlantic
West South Central

US Census Divisions (number of patients)
Participants’ preference for telehealth may indicate an openness to future siteless clinical trials.

Would the patient prefer telehealth appointments in the future, if offered?

<table>
<thead>
<tr>
<th>Category</th>
<th>Yes, whenever possible</th>
<th>Yes, but only for some types of appointments</th>
<th>No</th>
<th>I don’t know</th>
<th>Total (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>73 (26%)</td>
<td>97 (34%)</td>
<td>205 (78%)</td>
<td>1 (0.4%)</td>
<td>731</td>
</tr>
<tr>
<td>Complex neurodevelopmental</td>
<td>59 (26%)</td>
<td>97 (34%)</td>
<td>205 (78%)</td>
<td>3 (1.5%)</td>
<td>59</td>
</tr>
<tr>
<td>Hematological / immunological</td>
<td>82 (26%)</td>
<td>97 (34%)</td>
<td>205 (78%)</td>
<td>3 (1.5%)</td>
<td>82</td>
</tr>
<tr>
<td>Lysosomal storage</td>
<td>102 (26%)</td>
<td>97 (34%)</td>
<td>205 (78%)</td>
<td>3 (1.5%)</td>
<td>102</td>
</tr>
<tr>
<td>Mitochondrial</td>
<td>53 (26%)</td>
<td>97 (34%)</td>
<td>205 (78%)</td>
<td>3 (1.5%)</td>
<td>53</td>
</tr>
<tr>
<td>Neurological</td>
<td>94 (26%)</td>
<td>97 (34%)</td>
<td>205 (78%)</td>
<td>3 (1.5%)</td>
<td>94</td>
</tr>
<tr>
<td>Neuromuscular</td>
<td>217 (26%)</td>
<td>97 (34%)</td>
<td>205 (78%)</td>
<td>3 (1.5%)</td>
<td>217</td>
</tr>
<tr>
<td>Other metabolic</td>
<td>66 (26%)</td>
<td>97 (34%)</td>
<td>205 (78%)</td>
<td>3 (1.5%)</td>
<td>66</td>
</tr>
<tr>
<td>Tumor / lymphatic</td>
<td>68 (26%)</td>
<td>97 (34%)</td>
<td>205 (78%)</td>
<td>3 (1.5%)</td>
<td>68</td>
</tr>
</tbody>
</table>
Takeaways

- Real-world data can help address the challenges inherent to orphan drug development
- Integrating the voice of the patient can help answer the big questions in clinical trial planning:
  - Who?
  - What & When?
  - Where?
Power to the patients
Thank you!
Q&A

Sorin Fedele, PhD, MBA, MS
Caitlin Nichols, PhD
Aliza Thompson, MD, MS

Deputy Director of Division of Cardiology and Nephrology,
Office of New Drugs, Center for Drug Evaluation and Research, FDA
Concluding Remarks

Kerry Jo Lee, MD
Associate Director for Rare Diseases
Division of Rare Diseases and Medical Genetics, Office of New Drugs,
Center for Drug Evaluation and Research, FDA