C-Path’s Data Science

Transforming data into actionable knowledge for drug development

Klaus Romero MD MS FCP
Director of Clinical Pharmacology and Quantitative Medicine
The Critical Path Institute

Impact on Regulatory Science
- 15 global, pre-competitive, public-private partnerships with
- Participation from industry, academia, advocacy groups, and regulators

Regulatory qualification of preclinical and clinical biomarkers for use in safety, efficacy, and trial enrichment
Development and qualification of clinical outcome assessment tools
Development of quantitative modeling and simulation tools
Regulatory acceptance of nonclinical tools for medical product development

Impact on regulatory science
Forming and managing large international consortia
Provision of large-scale data solutions for scientific research
Clinical data standards development
How C-Path Works: A Public-Private Partnership

• Act as a trusted, neutral third party
• Convene scientific consortia of industry, academia, and government for sharing of data/expertise
  ✓ The best science
  ✓ The broadest experience
  ✓ Active consensus building
  ✓ Shared risk and costs

• 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
What is Model-Informed Drug Development?

- Development and application of pharmaco-statistical models of drug efficacy and safety from preclinical and clinical data to improve drug development knowledge management and decision-making\(^1\)

- Quantitative framework for prediction and extrapolation, centered on knowledge and inference generated from integrated models of compound-, mechanism-, and disease-level data and aimed at improving the quality, efficiency and cost effectiveness of decision making\(^2\)

Critical questions for trial design

• How many patients should be recruited to properly power the trial?
• What should be the inclusion criteria?
• Can the control arm be optimized?
• What types of progression rates are expected for different subpopulations?
• What measures of progression are most adequate, at which stages of the disease continuum?
• How long should the trial duration be?
• How often should I assess?
• What is the time-varying probability of dropouts, and what are their predictors?

How should one go about providing sound quantitative answers to these questions?
Answer 1: Quantifying variability

Quantifying multiple sources of variability simultaneously within the patient population reduces overall unexplained variability.

Result: The ability to predict more accurate progression rates for heterogeneous subpopulations of patients in clinical trials.
Answer 2: Multiple data sources

Understanding the ‘universe’ of a given disease’s heterogeneity

Result: The ability to more accurately account for the heterogeneity in rare diseases and avoid biased conclusions on few data sources

Unexplained variability in **ALL** patients with rare diseases

Unexplained variability in patients with rare diseases from a single study
Answer 3: Drug-trial-disease modeling

- **Disease**
  - Disease progression and characterization model

- **Clinical Trial Simulator**
  - Drug effects model / Disease modifying effect
  - Placebo effects model, dropout model

- **Longitudinal observational, registries, RWD & clinical trial data**

- **RWD & clinical trial data**
Putting it altogether

• Start with an understanding of what sponsors can practically use to design clinical trials, and reverse engineer

\[ S(t) = \frac{S_0}{\left( S_0 - (1 - S_0) e^{-\beta_1 t} \right)^{1/\beta}} \]

\[ TVP_i = \theta_i \left( \frac{cont_i}{ref} \right)^{\theta_{power} \cdot \theta_{cat_i} \cdot \theta_{coeff}} \]

\[ f(S; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) + \Gamma(\beta)} \cdot S^{(\alpha-1)} \cdot (1 - S)^{(\beta-1)} \]
Clinical Data Shared with C-Path

Clinical data: 146 studies: 77,054 subjects

Nonclinical: 179 studies; 11,775 subjects.
ReSeqTB: 9,215 Individual Isolates
RDCA-DAP: A resource for the future of drug development in rare diseases
RDCA-DAP: A resource for the future of drug development in rare diseases

Interface level I: Data interrogator to help orient researchers

Interface level II: Data analysis subset generator

Interface level III: Advanced data analytics workbench
Clinical Study Simulations

Evaluate different design options

Drug-Trial-Disease Models

\[
\frac{dFVC}{dAge} = k_{in} \times (1 + \text{AgeIn}) - (1 + \text{AgeOut}) \times FVC \times k_{out}
\]

\[
P_{01} = \frac{\exp(logit_{01})}{(1 + \exp(logit_{01}))}
\]

Study Execution

• X dose
• N
• Frequency of observations
• Inclusion/exclusion criteria

Simulated results

Trial optimization through simulations

Statistical Analysis

Design selection
Disease Progression Model

Input  Modeling  Output
Disease Progression Model

Input
Patient-level data

Modeling

Output
Disease Progression Model

**Input**
Patient-level data

**Modeling**
- Baseline severity
- Age
- Sex
- Demographics
- Genetics
- Medications
- Baseline biomarkers
- Dropouts
- Longitudinal endpoints
- Longitudinal biomarkers

**Output**
Disease Progression Model

**Input**
- Clinical studies

**Modeling**
- Age
- Sex
- Demographics
- Genetics
- Medications
- Dropout
- Baseline severity
- Baseline biomarkers
- Longitudinal biomarkers
- Longitudinal endpoints

**Output**
- Understanding of disease worsening
- Trajectory
- Rate
- Predictors
- Web Clinical Trial Simulator
Disease Progression Model

Input → Modeling → Output

DATA → TRANSFORMATION → KNOWLEDGE
C-Path’s impact in MIDD

- Survival model to predict T1D diagnosis, based on islet AA positivity.
- A model that changes the landscape for RCTs for T1D prevention.
C-Path’s impact in MIDD

- Multiple endpoints over time in Duchenne Muscular Dystrophy (DMD).
Preliminary findings from the FA database

Biphasic pattern

25ft walk velocity test

Years of Age
So…

Models, trials and endpoints?

• Quantitatively understanding sources of variability, multiple measures and markers across a disease continuum can streamline the pathway towards regulatory acceptance of quantitative solutions that can inform clinical trial design questions.
AD CTS: n=50

Mild-to-Moderate Alzheimer Disease Clinical Trial Simulator (beta v2.0)
AD CTS: n=50, with genetic enrichment
AD CTS: n=50, with genetic enrichment, and baseline severity characterization
Models, trials and biomarkers?

• It’s all about the sources of variability

• Unless dealing with safety or diagnosis, biomarkers are either:
  • Covariates in a model
  • One of many endpoints in a model

• Quantitatively understanding disease progression helps improve the understanding of biomarkers and other relevant sources of variability, and can streamline the pathway towards regulatory acceptance of quantitative solutions to improve clinical trial design efficiency
RDCA-DAP: A resource for the future of drug development in rare diseases
\[ S = f(t, p) \]
Thank you!
Rare Disease Cures Accelerator Platform
Reference architecture with key components/services/tools

Search and data provisioning

Self Service Analytics Engine

Sharing and Collaboration space

Modeling and Simulation Platform

Metadata Ontology Management

Data Modeling Service

Ontology Database

Data Sources

NORD db

Patient Registries

Pharma X Clinical Studies

Academic etc.

Data Staging/Data repositories

Data Catalog, Models, Results library

Audit tracking/Versioning

Workflow engine and schedulers

Analysis Data Sets/Data Marts

Data Preparation/Presentation

Data Ingestion/Integration

User Portal to access functionalities

Admin

User accounts/Roles Access management

User authentication service

Compute governance

Cloud/Elastic Compute infrastructure

Container config library

Search and data provisioning

Self Service Analytics Engine

Sharing and Collaboration space

Modeling and Simulation Platform

Metadata Ontology Management

Data Modeling Service

Ontology Database

Data Sources

NORD db

Patient Registries

Pharma X Clinical Studies

Academic etc.
Rare Disease Cures Accelerator Platform
Reference architecture with key components/services/tools: Some proposed examples