

Our Journey in Integrating External Control Arms (ECAs) and Real-World Data (RWD) for Rare Disease Trials

Cyril Sombrin, Gautham Selvaraj, Cytel Inc., Geneva, Switzerland

Background

External Control Arms (ECAs) are increasingly applied in rare disease trials where Randomized Controlled Trials (RCTs) may be impractical. Recent guidance from the FDA, CDISC, and PHUSE highlights the importance of data reliability, bias reduction, adherence to standards (SDTM, ADaM), and the regulatory use of Real-World Data (RWD). This poster summarizes best practices and presents two case studies: constructing an ECA from natural history studies and legacy RCT data, and leveraging publicly available RWD and RCT sources such as the Critical Path for Alzheimer’s Disease (CPAD) database. We address common challenges in data integration, standardization, and statistical approaches such as propensity score matching, alongside practical strategies for regulatory communication and dataset conformance.

When and Why ECA important

Rare Disease Challenges
(RCT recruitment challenge, ethical, cost)

Alternative Data resources
(Historical, RWD, Observational)

Faster Timelines

Better Representation
(wide variety data more leads)

Beyond NDA

Driven by guidance and industry initiatives

Challenges

Data Quality

- **Diverse** data, **inconsistent** data collection
- **Missing or incomplete** data
- Endpoint and timing **varies**

Selection Bias

- **No randomisation** baseline imbalances.
- Channelling **Bias**
- Requires **advanced methods** (e.g., PSM, IPW)

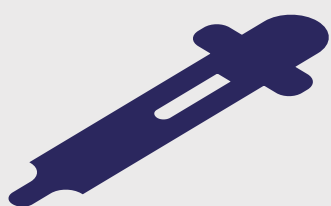
CDISC Package

- **Difficulty** with legacy data
- External data often **lacks standards**.
- Lacks detailed metadata, **complicate** define xml creation

FDA Stance

- Strict requirements for **data reliability and validity**
- Need for **high-quality, well-documented RWD sources**
- **Well-defined and consistent endpoints** are essential
- Strong emphasis on **bias mitigation** and **CDISC alignment**.
- FDA is **open-minded but cautious**, ECAs are not a shortcut to approval

Case Study 1 - ECAs from Natural History and Legacy RCT Data



Treated Pivotal Study (Randomized Controlled Trial)

Phase III, randomized, double-blind, placebo-controlled trial with a long-term open-label extension.



ECA Source A

Multicenter longitudinal observational natural history cohort of patients with a similar population in region A



ECA Source B

Multicenter longitudinal observational natural history cohort of patients with a comparable population in region B

SDTM source

Pooled in ADAM

Analysis

CDISC package and SDTM TS

Case Study 2 – ECAs from Public Databases



Treated Pivotal Study

Phase III, randomized, double-blind, placebo-controlled trial with a long-term open-label extension.



CPAD data

The Critical Path for Alzheimer's Disease (CPAD) database contains a pool of patients coming from several Randomized Controlled Trials (RCTs).

Long process to get data

Pool of Metadata from trials

ADAM like dataset.

No CDISC package.

Case studies – Integration challenges

Data Cleaning and harmonization

- **Unclean, missing, and partial** data
- **Limited** cleaning options
- **Inconsistent subject data** requiring harmonization

Standards and Traceability

- RWE source **lacked SDTM compliance** → required ADaM alignment
- **Non-traditional** SDTM structures complicated traceability
- Pinnacle 21 flagged **missing** DM/EX (expected per FDA guidance)

Data Selection & Integration

- Complex RWE databases **with unclear origins and missing variables**
- Applying pivotal trial criteria to RWE **required complex algorithms + manual review**
- **Inconsistent** outcome tools, versions, and visit windows
- Single third-party pooled database with **restricted access** (deleted post-use)

Statistical Modeling Challenges

- Propensity score matching difficult due to **covariate differences**
- Repeated model tuning; **new covariates added** after sponsor review
- Some models required **4–5 hours per figure** (runtime burden)

Regulatory Agility

- **Rapid turnaround** for regulatory queries and versioning
- Documentation updates **with minimal** lead time

Learnings and Future considerations

- Mapping real-world data to SDTM/ADaM takes time and custom work
- Design flexible CDISC structures to adapt diverse data sources
- Align external and trial data to CDISC early in the process
- Systematic handling of non-standard IDs and missing dates
- Harmonize terminology (e.g., MedDRA, lab units) across sources

- Document all assumptions, data limitations and transformation steps
- Perform pooled datasets validation
- Plan for additional sensitivity analyses
- Be prepared for additional questions from regulatory

ABBREVIATIONS

ECA = External control Arm; RWE = Real world evidence; RWD = Real world Data

Contact Details

Cyril Sombrin: cyril.sombrin@cytel.com, Gautham Selvaraj: gautham.selvaraj@cytel.com