Clinical Trials Workgroup Community Update

November, 2022
Workgroup Objectives

Transform CDISC SDTM Formatted Clinical Trials Data to OMOP CDM
- Propose Conventions
- Generate Mappings
- Publish Guidelines
Challenges

- Data are currently collected according to CDISC SDTM Standard and require ETL definitions/mappings to convert into OMOP CDM
- Privacy concerns make it difficult to access data for testing
- Clinical Trials data are specialized type of medical record often requiring custom vocabularies
Our Approach

● Document Proposed Conventions
  ○ https://github.com/OHDSI/ClinicalTrialsWGETL/wiki

● Collect a set of SDTM formatted source data that adequately represents a breadth of data elements
  ○ Across different Phases and Therapeutic Areas

● Data Sources
  ○ PhUSE
  ○ Vivli
  ○ CPath
  ○ TrialTwin
Our Approach - Continued

- Produce ETL definitions and field mappings for source data
  - https://github.com/OHDSI/ClinicalTrialsWGETL/tree/master/docs

- Proof of Concept
  - Perform data transformations using definitions & mappings in a protected environment
Key Results & Progress

- Apply our approach to PhUSE (synthetic trial data)
  - Completed (2020-2021)
  - Validated proposed conventions
  - Initial ETL definitions and mappings generated
  - Data not comprehensive enough for large scale use

- Apply our approach to Vivli sourced data (actual trial data)
  - Ongoing since 2021 - Getting approval to access the data has taken time
  - Summer of 2022: Gained access to ~5 SDTM formatted data sets through the Vivli platform
  - Fall of 2022: Gained access to secure implementation enclave to perform data transformation
Apply our approach to Vivli sourced data (actual trial data)
  ○ Currently ongoing: Validating existing ETL mappings and generating new mappings to support conversion of Vivli sourced datasets to OMOP CDM
    ■ Team members volunteer to produce mappings for specific SDTM domains and/or OMOP CDM output tables
    ■ During Bi-Weekly Friday meetings, work on mappings as a team and work through and problems/questions

Going forward
  ○ Perform transformations for as many Vivli data sets as we can
    ■ This should get easier as our set of mappings becomes more robust
  ○ Publish/update mapping definitions on github
Join Us!

● Strong Team
  ○ 273 Registered Members
  ○ Consistent member participation every meeting
  ○ Various backgrounds represented: Pharma, Regulatory, Standards Orgs, Vendors, Technology Advocates

● Contact Us
  ○ Workgroup Leads: Mike Hamidi, Zhen Lin
    ■ Kamiar.Hamidi@pfizer.com
    ■ Zhen.Lin@ohdsi.org
  ○ We use Microsoft Teams for communication (email for an invite!)

● Resources
  ○ Github Repo: https://github.com/OHDSI/ClinicalTrialsWGETL
  ○ Proposed Conventions: https://github.com/OHDSI/ClinicalTrialsWGETL/wiki