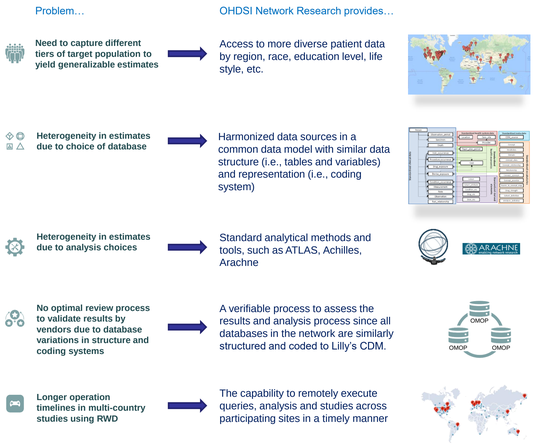
Using OHDSI Data Network for Capturing Real-World Evidence:

Our Experience with a Multi-Country Study on an Obese and Overweight Cohort

PRESENTER: **Hamed Abedtash**

INTRODUCTION:



Adoption of OMOP CDM has enabled conducting "Network Research" by simultaneously executing the same analysis or protocol on a number of harmonized data sources independent of the source coding systems and database structure.

Objectives: We aimed to assess the feasibility of conducting Network Research on OHDSI data network through a multi-country cohort characterization study on obese and overweight population using six distributed electronic health records or claims databases.

METHODS:

Data Nodes

- US Ambulatory EMR
- UK THIN EMR
- France Data Analyzer EMR
- Germany Data Analyzer EMR
- Belgium EMR
- South Korea Claims

Metrics of Interest

- Age at index date
- Gender distribution
- BMI
- Smoking status
- Systolic blood pressure
- Diastolic blood pressure
- Weight
- Height
- Waist circumference
- Liver function lab test
- Lipid panel lab test
- Glucose profile
- History of cardiovascular occurrences
- History of medications

Cohort Definition

- Age: 18 years or older
- At least 2 of the below events indicating obesity diagnosis 6 months apart during the study period. The first diagnosis date is the Index Date.
  - Obesity diagnosis report
  - Bariatric surgery procedure
  - Anti-obesity medications, such as appetite suppressants, lipase inhibitors
  - BMI ≥27 kg/m2
  - Dietary counseling encounter for obesity management
- Continuous observation period (enrolment in claims databases, continuous observation in EHR data) with drug exposure capture starting 6 months prior to and 1 year after the Index Date.

- **We provided the TPO with a requirement document (protocol) describing the rationale, objectives, study design, data coverage tests, and statistical plan.**
- **We evaluated the data coverage per each variable in each data source.**
- **We had access to the results throughout the project but not the patient-level data.**
- **Lilly scientist and TPO analyst were in constant communication throughout the study to develop analytical scripts, refine study design, and analyze the data.**
- **TPO distributed analytical scripts to data partners for data analysis execution.**
- **Subsequently, each data provider executed the scripts on their data residing on-site and returned the results for further analyses.**

We found the OHDSI Network Research a **feasible and efficient** approach to conduct multi-country observational studies. In particular, the level of **transparency** during the analysis process was encouraging.

Take a picture to download the full paper

LESSONS LEARNED:

Efficient method.
It was an exceedingly efficient method to analyze multi-database studies where the average time to analyze each data node was 2.6 weeks.

Transparent process.
Since the data was formatted in OMOP CDM, we could ask vendor further quality-check questions on existing data points, data coverage, and data transformation quality consistently across to ensure calculated measures are valid.

Ensure validity of concept mappings at the first place.
Unmapped source codes may pose selection and measurement biases; therefore, we asked the analyst to curate the list of source codes that were not mapped to OMOP standard concepts. To address this issue, we reviewed the list of concepts relevant to our study to ensure validity of mappings in each database.

RECOMMENDATIONS:

Validate the quality of data transformation.
Ensure relevant source codes are mapped to standard concepts in the OMOP repository to avoid any miscalculations and selection bias. Although OMOP CDM enables quick analysis of multiple distributed RWD repositories, there is a potential risk that not all OMOP-transformed databases have the same data transformation quality due to the complexity of embedded semantic network and OMOP-specific conventions.

Validate code lists before starting data analysis.
In one instance, the data coverage of "fasting blood glucose" variable was reported very low (2%). Upon our review of the list of LOINC codes that TPO had used in the analysis, we noticed one important LOINC code (2345-7) was missing. The code was not included in the analysis because there was no mention of "fasting" in the code description, while the code is a legitimate laboratory test code for measuring fasting glucose level in blood because one of the requirements of the test is that "patients should fast for 12 hours". Upon including this code into the analysis, the data coverage jumped to 47% from 2% for the US database, and similarly data coverage improved in other data sources.

Hamed Abedtash
Eli Lilly & Company,
Indianapolis, IN

Lilly

OHDSI