OHDSI GIS OKR

2022

Overall Objective

- We support the use of place-related data in combination with clinical data to understand the impact of various geographic attributes on exposures, care-delivery, and health. To do this we create or identify the data, schemas, vocabularies, and software required to store, visualize, and analyze data on patient locations and on health-relevant social and environmental geographic attributes so that they can be used by the OHDSI tool stack an equivalent way to healthcare data. Products will minimize risks to patient privacy.

- Place-related data domains of interest include: social determinants of health, pollutants and other toxins, the built environment, political and corporate policies, distance from care. These products will be consistent with OHDSI principles for data modeling and will minimize the modification required to extend its tooling that implement best practices for analytics. We will also support the development of new HADES analytic packages that employ domain-specific features such as spatial autocorrelation.

- Finally, we will convene and support the use of these data and tools in the generation of OHDSI studies that focus on or account for geographic as well as person-level attributes.

Key Results

- Develop an **optional OMOP CDM module** that consists of a universal schema for storing geospatial data. This module supports standardized query execution on a combination of place-related and person-level data using spatiotemporal joins between tables.

- Develop a privacy-preserving **geocoding mechanism** that runs locally to relate geographic and person-level data. This process attaches longitude and latitude to patient residence data from electronic health records or others sources. With proper metadata, this can be used on arbitrarily defined sources of data on geographic attributes.

- For every place-related dataset to be ingested into the module, we will **develop metadata** that includes: basic file information needed to identify, describe, and cite the dataset, link to its documentation and license information, specifies methods of data capture and statistical transformations.

- **Drive development through use cases** to retrieve, ingest and standardize place-related datasets from multiple geographic data categories. Build frameworks and develop code to automate and standardize this ingestion process.

- **Execute analyses** with the ingested and standardized place-related data to correlate them with health outcomes when they themselves are of interest or to use them as covariates in analyses involving exposure to drugs or medical procedures.

- **Develop documentation** and tutorials that enable community members to use these resources.