**OHDSI Network Study: Assessing Fitness of OHDSI Datasets for Clinical and Regulatory Use**

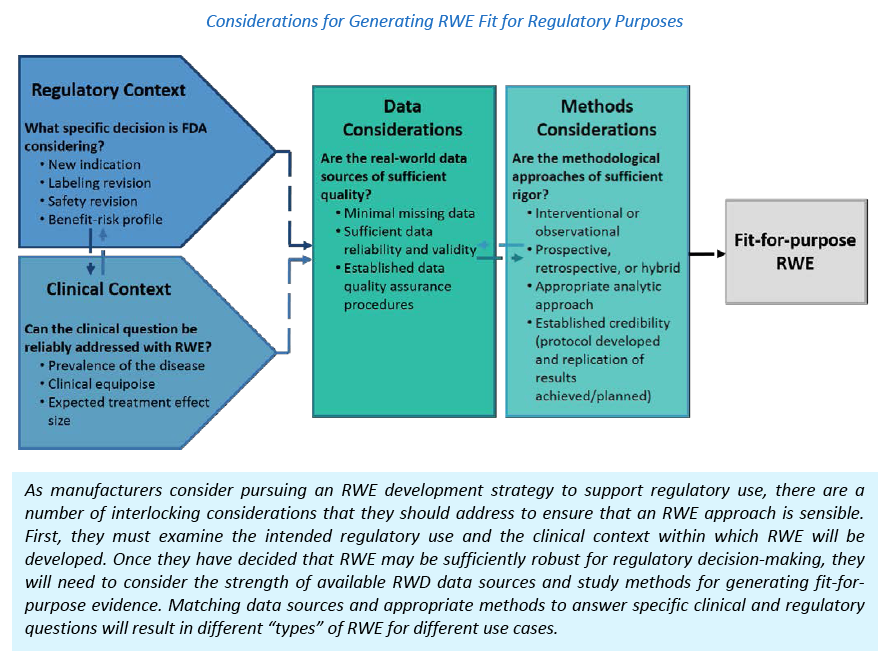
**Project Leads:** Lili Peng ([lilipeng@gmail.com](mailto:lilipeng@gmail.com))

**Collaborators:** Recruiting

**Objective:**

* Aim 1: Characterize the fitness of a set of OHDSI datasets across various diseases and medical conditions. Identify the diseases and conditions for which the OHDSI datasets are the most fit for clinical use.
* Aim 2: Characterize the fitness of a set of OHDSI datasets for across various drug-outcome relationships. Identify the drug-outcome relationships for which the OHDSI datasets are the most fit for regulatory use.

**Rationale:** Given the recent issuances of Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices1 and Prescription Drug User Fee Act VI Commitment Letter2, it is clear that FDA has recognized real-world data (RWD) as a useful source for generating real-world evidence (RWE) and has even prioritized RWE in its regulatory decision-making. In September 2017, the Duke-Margolis Center for Health Policy posited a multi-stage ‘fit-for-purpose’ RWE framework, intended for regulatory agencies and pharmaceutical manufacturers, to help guide efforts in assessing whether an RWD source(s) and complementary analyses and methodologies are appropriate, or ‘fit’ enough, to generate high-quality, comprehensive, reliable RWE evidence specific to answering clinical and regulatory questions (see Figure 1).3 We argue that an intermediate step should exist between defining the clinical or regulatory question and considering the data quality: determining the availability, scope, and type of RWD. Specifically, in the clinical context, the scope of an RWD relies on the depth and breadth of data available on the disease or medical condition, whereas in the regulatory text, this scope applies to drug-outcome relationship for conducting safety surveillance.

****

**Figure 1. Considerations for generating RWE fit for regulatory purposes**

Given that OHDSI datasets are publicly available and its analytics tools (i.e. ACHILLES) are open-source, we assess select OHDSI datasets to determine the diseases and drug-outcome relationships for which OHDSI is most fit for clinical and regulatory use, respectively. Awareness of these results can enable manufacturers and regulatory agencies about the fit-for-purpose of OHDSI datasets.

**Research Question:**

* Aim 1: What are the diseases and conditions for which the OHDSI datasets are the most fit for clinical use? Is there a significant representation of diseases that are difficult to treat or rare diseases? Examples include:
  + Oncology
  + Neurological diseases
    - Alzheimer’s disease
    - Parkinson’s
    - ADHD
  + Others… (have to research)
* Aim 2: What are the drug-outcome relationships for which the OHDSI datasets are the most fit for regulatory use? Is there a significant representation of drug-outcome relationships for drugs that have recently warranted urgent societal attention?
  + Opioids
  + Check CDC.gov for outbreaks: <https://www.cdc.gov/outbreaks/index.html>

**Research Methods:**

* **For the clinical context:**
  1. Acquire a list of all unique conditions in ACHILLES
  2. For each unique condition, run statistics on:
     + Age at First Observation
     + Number of people with a diagnosis (for all diagnoses)
     + Number of people exposed to a drug (for all drugs)
     + Number of people with an observation period for a given year (for all years)
     + Distributions of observation period durations, length of exposures, age at first observation
  3. Rank in order the diseases with the most 🡪 least OHDSI data, based on ACHILLES results.
  4. Take note of gaps in OHDSI datasets for diseases with high clinical trial failure rates (i.e. various oncological diseases) and rare diseases with few or no cures. (These are gaps that OHDSI could fill by recruiting a potential Data Partner providing the respective data that OHDSI could perform ETL on.)
* **For the regulatory context:** 
  1. Acquire a list of all unique drug-outcome relationships in ACHILLES
  2. For each unique drug-outcome relationship, run statistics on:
     + <>
     + Test case: h1n1
  3. Rank in order the drug-outcome relationships with the most 🡪 least OHDSI data, based on ACHILLES results.
  4. Take note of gaps in OHDSI datasets for drug-outcome relationships that have received critical societal attention. (These are gaps that OHDSI could fulfill by active procuring data for ETL conversion.)

**Study Design:**

**References:**

1. Administration USFaD. Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices: Guidance for Industry and Food and Drug Administration Staff. In: FDA Center for Devices and Radiological Health CfBEaR, (ed.). U.S. Food and Drug Administration, 2017, p. 1-24.

2. Administration USFaD. PDUFA Reauthorization Performance Goals and Procedures Fiscal Years 2018 through 2022 In: Administration USFaD, (ed.). 2016.

3. M Berger GD, K Frank, A Hernandez, M McClellan, S Okun, M Overhage, R Platt, M Romine, S Tunis, M Wilson. *White Paper: A Framework for Regulatory Use of Real-World Evidence*. 2017. Washington, DC: Duke-Margolis Center for Health Policy.