OHDSI/OMOP Introduction

Mui Van Zandt

IQVIA

VP/Global Head, Data Strategy, Access, and Enablement

Instructors

Mui Van Zandt
IQVIA

Mukkesh Kumar
A*STAR

Erwin Tantoso
A*STAR

Evelyn Goh
National University of Singapore

Gyeol Song
IQVIA

Seng Chan You
Yonsei University

Thanh-Phuc Phan
University Medical Center HCMC

Nicole Pratt
University of South Australia

Agenda

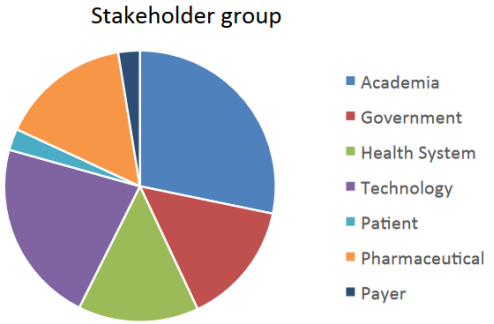
Session Name	Time	Instructor(s)
OHDSI/OMOP Intro	09:00 – 09:20	Mui Van Zandt
OMOP CDM and Vocabulary	09:20 – 10:00	Mukkesh Kumar & Erwin Tantoso
OMOP Conversion Process	10:00 – 10:30	Evelyn Goh
Energy Break	10:30 – 10:40	
ETL Exercises	10:40 – 12:00	Gyeol Song
Lunch	12:00 – 13:30	
OHDSI Analyses: Building Cohorts	13:30 – 14:50	Seng Chan You
CohortDiagnostics & Population-Level Estimation	14:50 – 15:30	Thanh-Phuc Phan
Interpreting Results	15:50 – 16:30	Nicole Pratt

OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

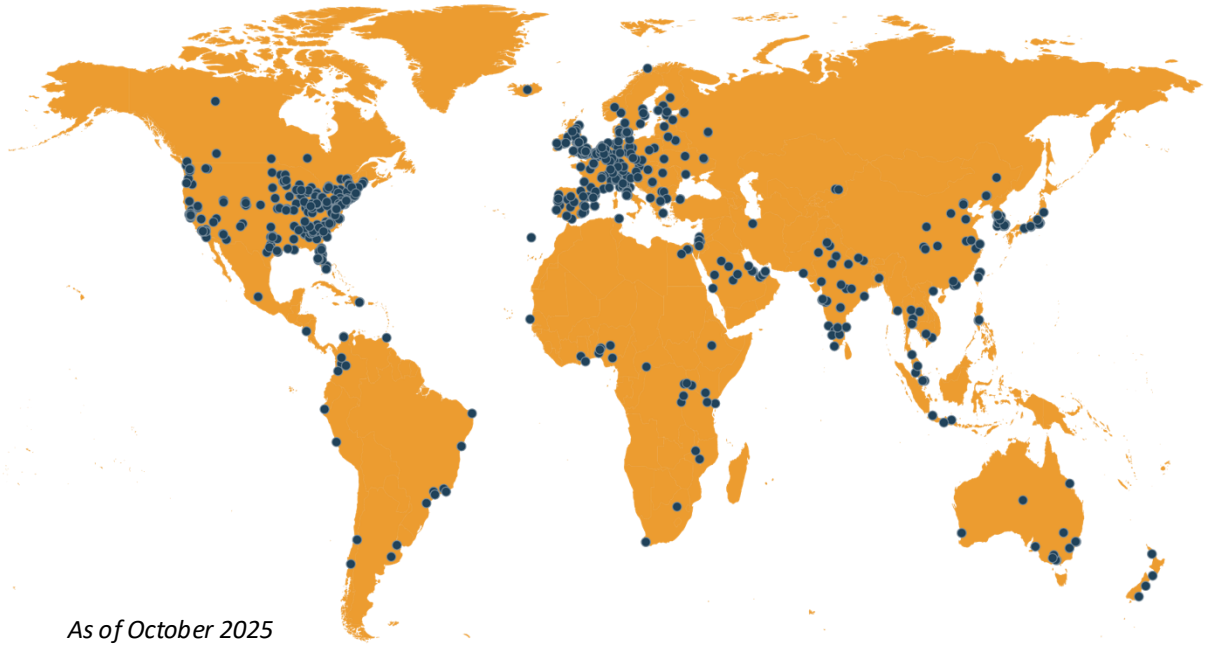
What OHDSI is:

- ✓ **Open Source**
- ✓ **Community**
- ✓ **Data**



Why Choose OHDSI/OMOP:

- ✓ **Fast, reliable** studies across a series of datasets and data types
- ✓ **Reduced cost of ownership** including understanding coding schemes, writing statistical programs across databases or developing software
- ✓ **Expanded data access** via the OHDSI network and remote multi-center database studies



As of October 2025

OHDSI Collaborators:

- 4,751 collaborators
- >1,100 organizations
- 88 countries from 6 continents

OHDSI Network:

- 544 data sources
- 974M unique patient records

https://ohdsi.org/

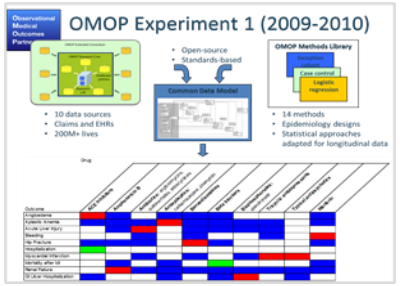
OHDSI's Mission

To improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care.

History of OMOP/OHDSI

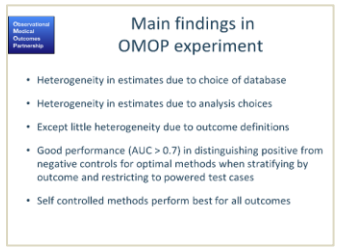
2009

OMOP Experiment #1



2013

End of OMOP Experiment



2014

Launch of OHDSI

OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Korea Chapter

2015

First OHDSI Symposium/ Network Study Published

First Hackathon at Columbia University

2016

China Chapter

2017

European Chapter

FDA Adoption (FDA BEST Launch)

2018

EHDEN Initiation (Europe)

FEEDER-NET Initiation (Korea)

First European Symposium

2019

EMA Adoption

Australia, Japan Chapters

2020

OHDSI COVID-19 Study-a-Thon

Singapore Chapter

2022

Global Acceptance

India Chapter

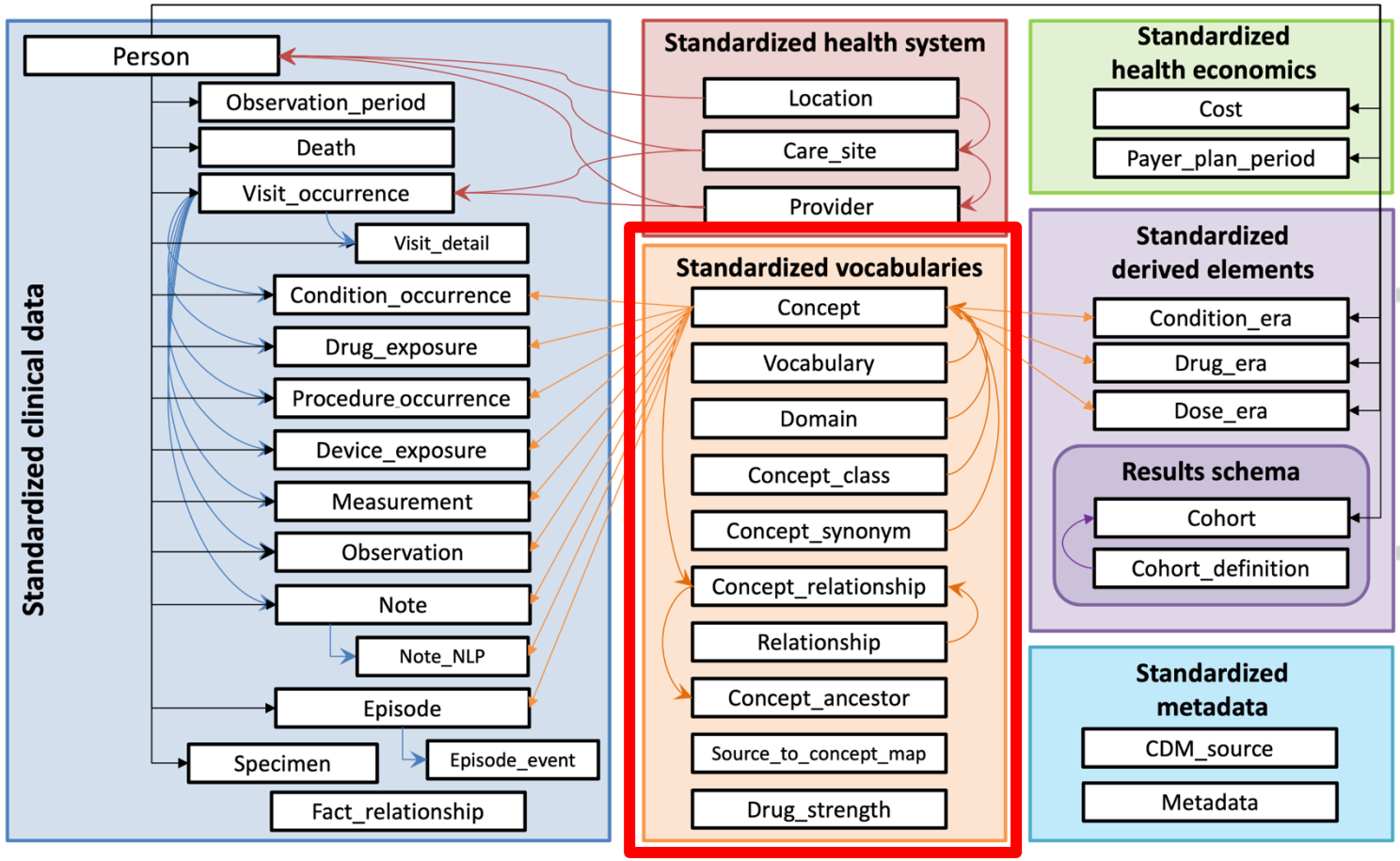
2023

OMOP in Thailand

Thailand (1)
Siriraj Hospital EHR

OMOP Common Data Model (CDM)

Ontologies are critical when designing data models

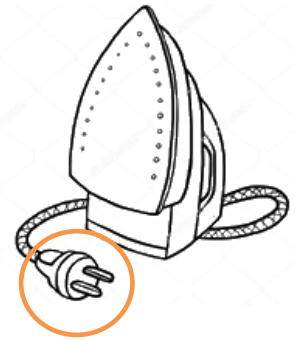


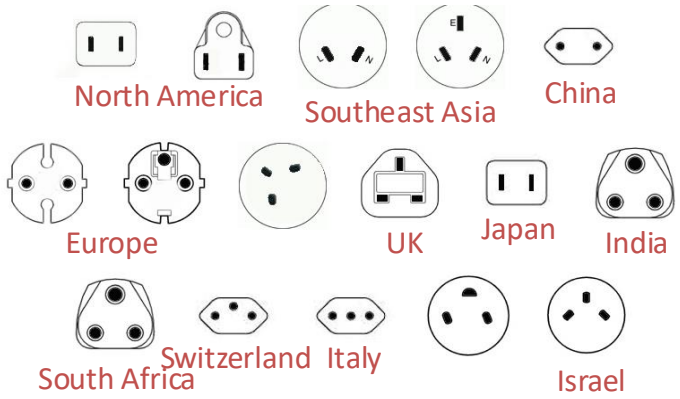
As of August 2025

- **11,804,307 concepts**
  - 3,784,263 standard concepts
  - 971,914 classification concepts
- **145 vocabularies**
- **43 domains**
- **87,948,636 concept relationships**
- **101,696,159 ancestral relationships**
- **6,028,711 concept synonyms**

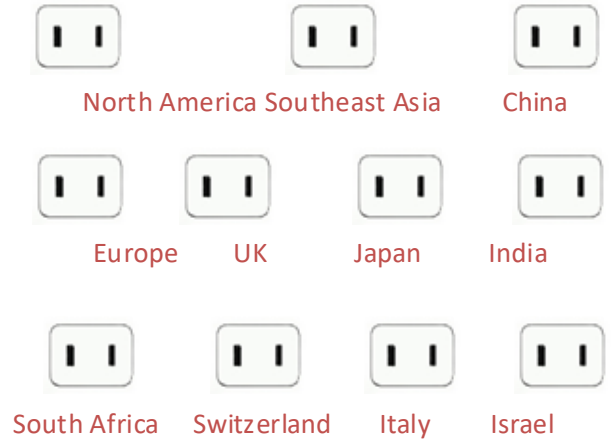
Data Standardization to OMOP Enables Systematic Research

Analytical method:
Adherence to Drug

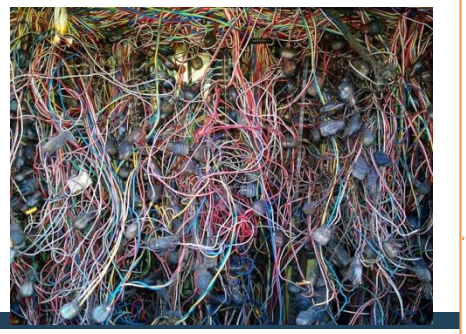




OMOP CDM

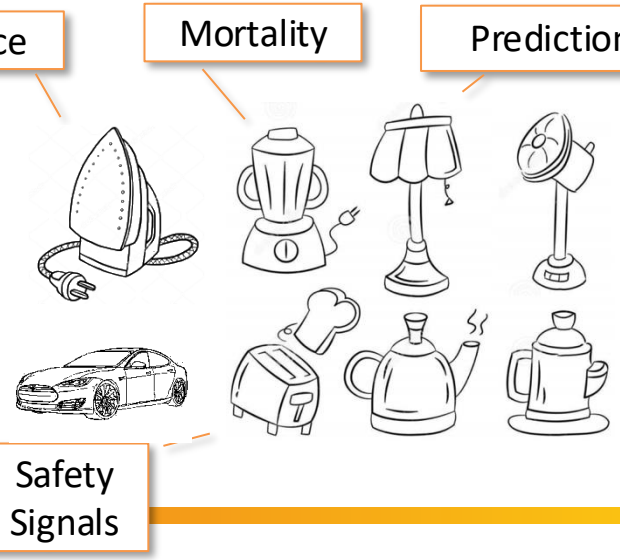


One SAS or R script for each study



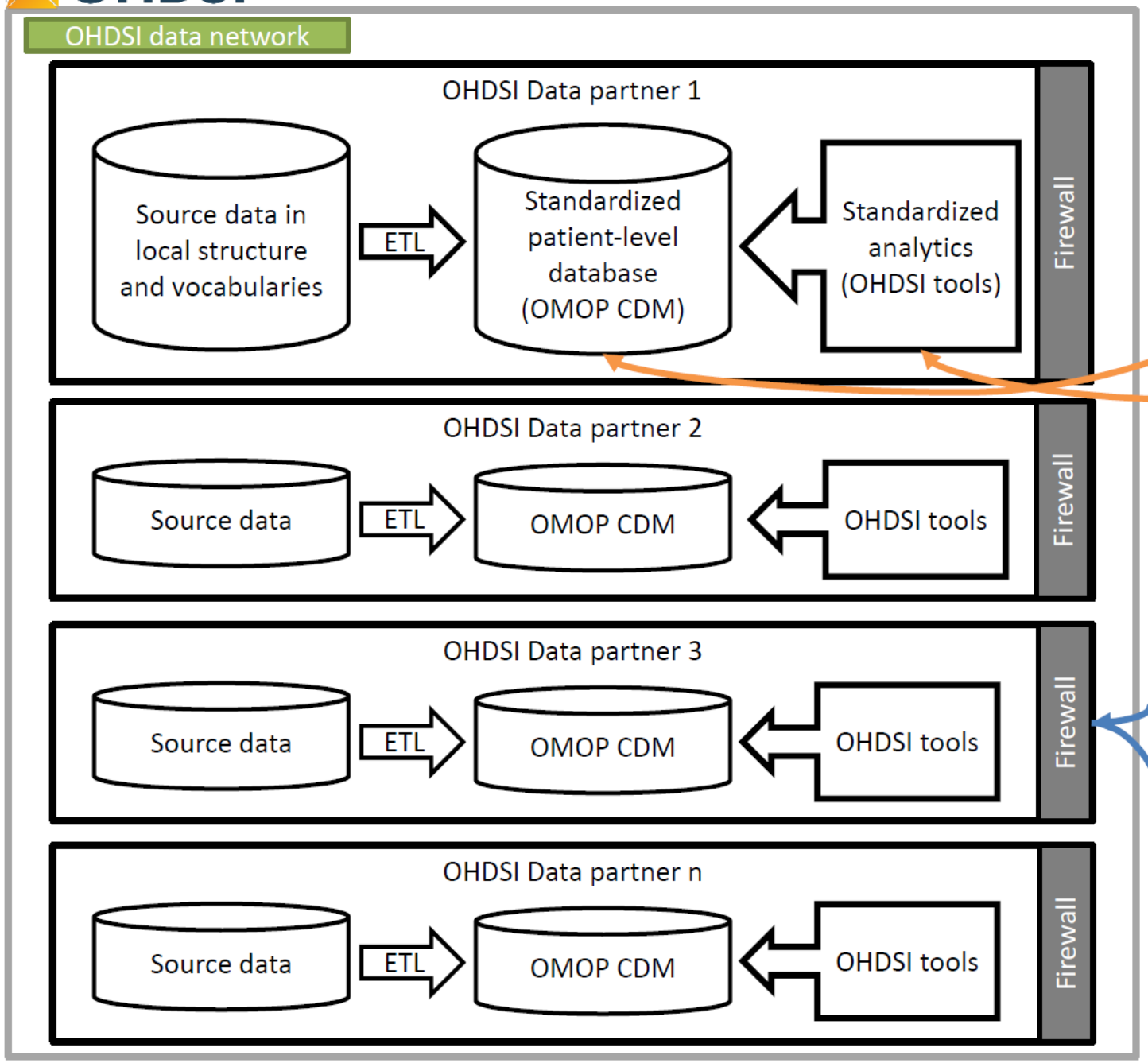
- Reliant on partner capabilities
- Not scalable
- Not transparent
- Expensive
- Slow
- Prohibitive to non-expert routine use

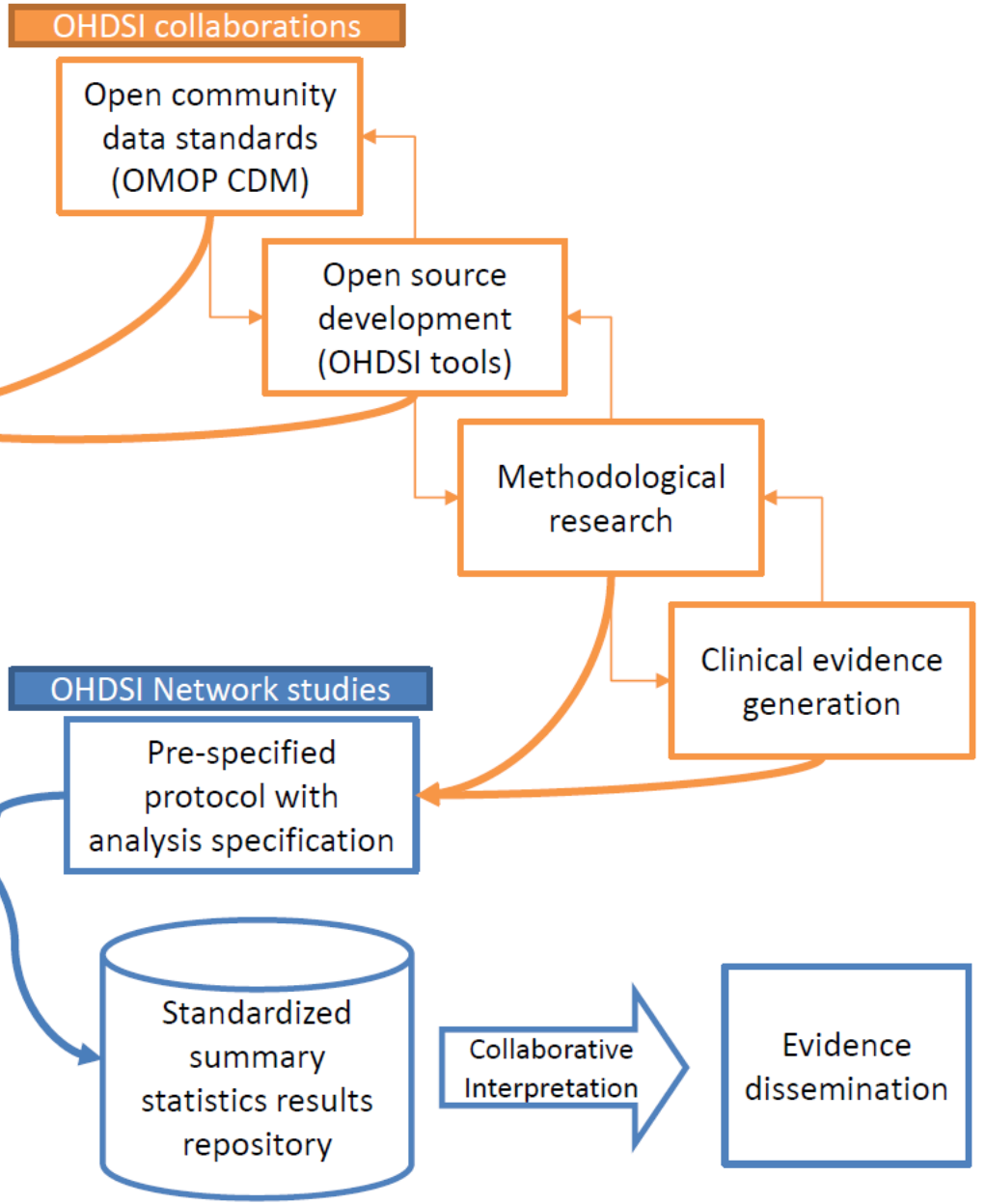
OHDSI tools



OHDSI

To improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care





Health Analytics Data-to-Evidence (HADES)

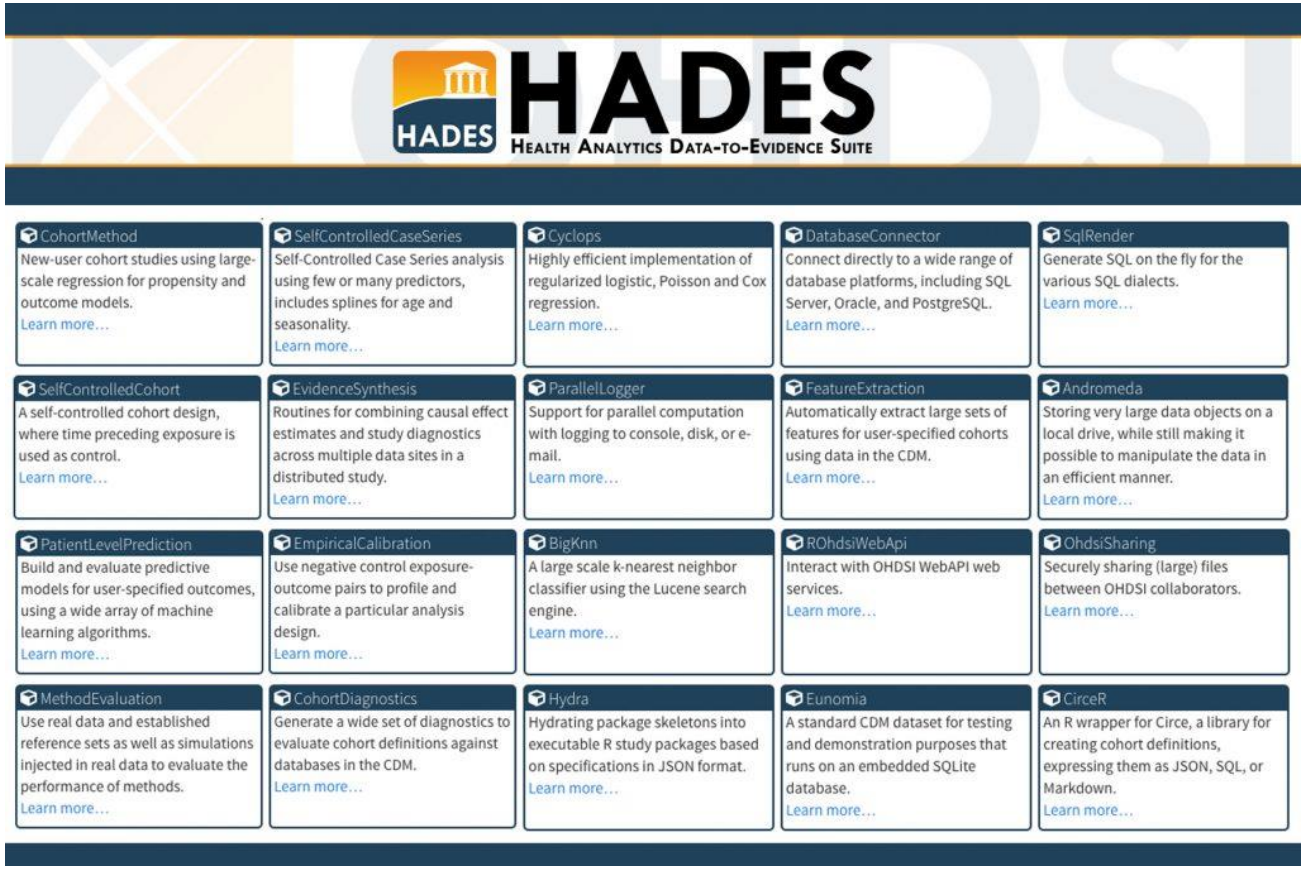
Suite of OHDSI R packages for running standardized analytics against OMOP data assets

Overview

- R packages that can run against any OMOP database

- Support R packages
  - DatabaseConnector
  - SqlRender
  - ROhdsiWebApi
  - CohortDiagnostics

- Analytical R packages
  - PatientLevelPrediction
  - CohortMethod (comparative effectiveness)
  - FeatureExtraction (characterization)
  - SelfControlledCaseSeries



https://ohdsi.github.io/Hades/index.html

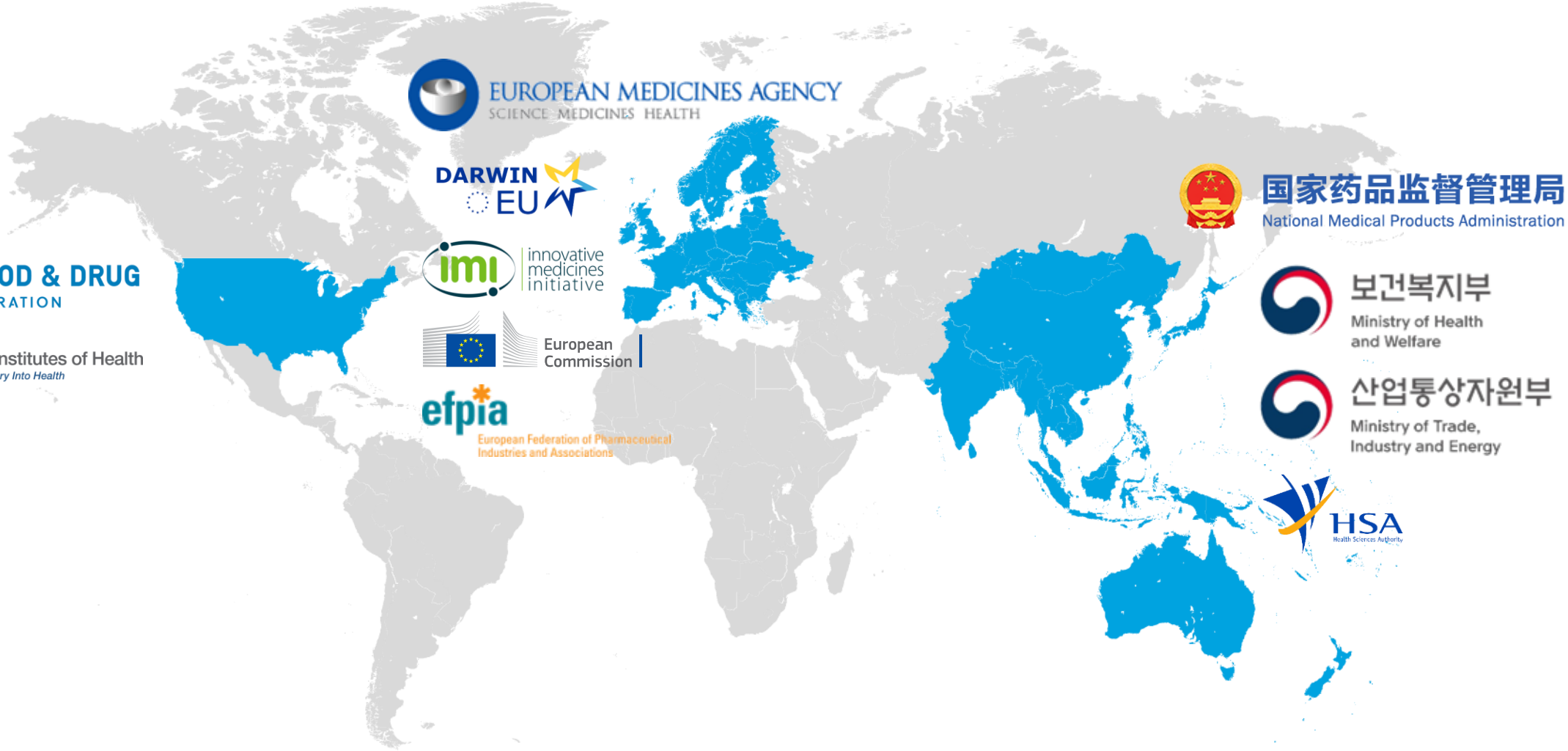
Data relevance across clinical domains

OMOP Workgroups & OHDSI Phenotype Collaborations

APAC Current Participants: 297 Lead: Mui Van Zandt	**ATLAS/WebAPI** Current Participants: 253 Lead: Anthony Sena	**Clinical Trials** Current Participants: 295 Leads: Mike Hamidi, Lin Zhen	**CDM** Current Participants: 686 Lead: Clair Blacketer	**CDM Vocab Subgroup** Current Participants: 686 Lead: Michael Kallfelz	**Data Network Quality** Current Participants: 298 Lead: Clair Blacketer	**Dentistry** Current Participants: 8 Lead: Robert Koski	**Education** Current Participants: 136 Lead: Nigel Hughes
HADES Current Participants: 295 Lead: Martijn Schuemie	**Health Equity** Current Participants: 228 Lead: Jake Gillberg	**Latin America** Current Participants: 48 Lead: Jose Posada	**NLP** Current Participants: 444 Lead: Hua Xu	**Oncology** Current Participants: 328 Lead: Asieh Golozar	**Registry** Current Participants: 175 Lead: Tina Parciak	**Steering Group** Current Participants: 82 Lead: Patrick Ryan	**Vaccine Vocabulary** Current Participants: 79 Lead: Asiyah Lin
Early-Stage Researcher Current Participants: 243 Leads: Faaizah Arshad, Ross Williams	**Eye Care & Vision Research** Current Participants: 74 Leads: Sally Baxter, Kerry Goetz	**FHIR and OMOP** Current Participants: 287 Leads: Jon Duke, Davera Gabriel, Christian Reich	**GIS** Current Participants: 157 Leads: Robert Miller, Kyle Zollo-Venecek, Andrew Williams	**Methods Research** Current Participants: 379 Leads: Martijn Schuemie, Marc Suchard	**Perinatal & Reproductive Health Group** Current Participants: 30 Leads: Alison Callahan et al.	**Psychiatry** Current Participants: 132 Leads: Dmitry Dymshyts, Andrew Williams	**Surgery & Perioperative Medicine** Current Participants: 42 Leads: Jenny Lane, Evan Minty
	Medical Imaging Current Participants: 155 Leads: Paul Nagy, Seng Chan You	**Medical Devices** Current Participants: 141 Leads: Vojtech Huser, Asiyah Lin	**Open-Source Community** Current Participants: 145 Leads: Adam Black, Paul Nagy	**Patient-Level Prediction** Current Participants: 89 Leads: Jenna Reps, Ross Williams	**Healthcare Systems** Current Participants: 471 Lead: Melanie Philofsky	**Phenotype** Current Participants: 310 Lead: Gowtham Rao	

- ❑ Type 2 Diabetes Mellitus
- ❑ Type 1 Diabetes Mellitus
- ❑ Atrial Fibrillation
- ❑ Multiple Myeloma
- ❑ Alzheimer's Disease
- ❑ Hemorrhagic Events
- ❑ Neutropenia
- ❑ Parkinson's Disease and Parkinsonism
- ❑ Attention Deficit Hyperactivity Disorder
- ❑ Hypertension
- ❑ Acute Myocardial Infarction
- ❑ Heart Failure
- ❑ Cardiomyopathy
- ❑ Multiple Sclerosis
- ❑ Hidradenitis Suppurativa
- ❑ Anaphylaxis
- ❑ Depression
- ❑ Non-Small-Cell Lung Cancer
- ❑ Drug-Induced Liver Injury
- ❑ Severe Visual Impairment And Blindness
- ❑ Suicide Attempts
- ❑ Kidney Stones
- ❑ Delirium
- ❑ Systemic Lupus Erythematosus
- ❑ Triple Negative Breast Cancer
- ❑ Pulmonary Hypertension
- ❑ Prostate Cancer
- ❑ HIV

Global OHDSI Adoptions



FDA U.S. FOOD & DRUG ADMINISTRATION

NIH National Institutes of Health
Turning Discovery Into Health

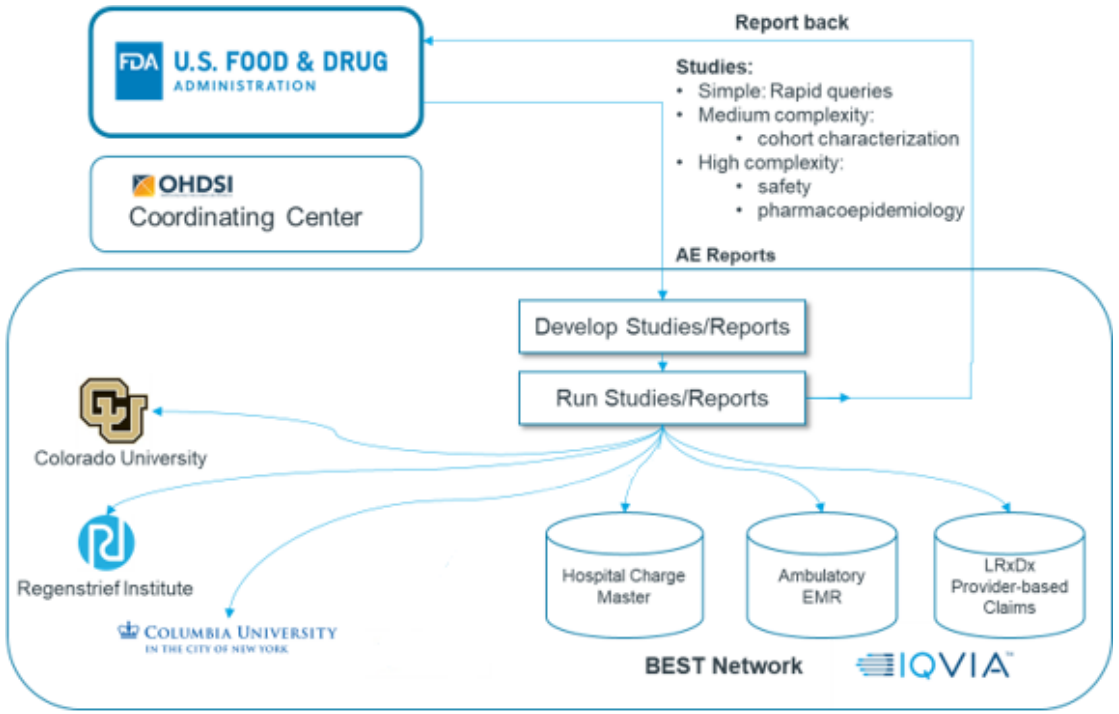
CDC
CENTERS FOR DISEASE CONTROL AND PREVENTION

DEPARTMENT OF HEALTH & HUMAN SERVICES • USA

FDA BEST – Overview

Premier, multi-center research collaborative driving large scale health analytics research

EXAMPLE OF DATA NETWORK



Network Overview

- Started in September 2017
- Today's largest distributed network of clinical data
- Collaborative research model, guided by efforts across the OHDSI community and US FDA
- Iterative sponsored studies facilitated by IQVIA and the global network of data partners

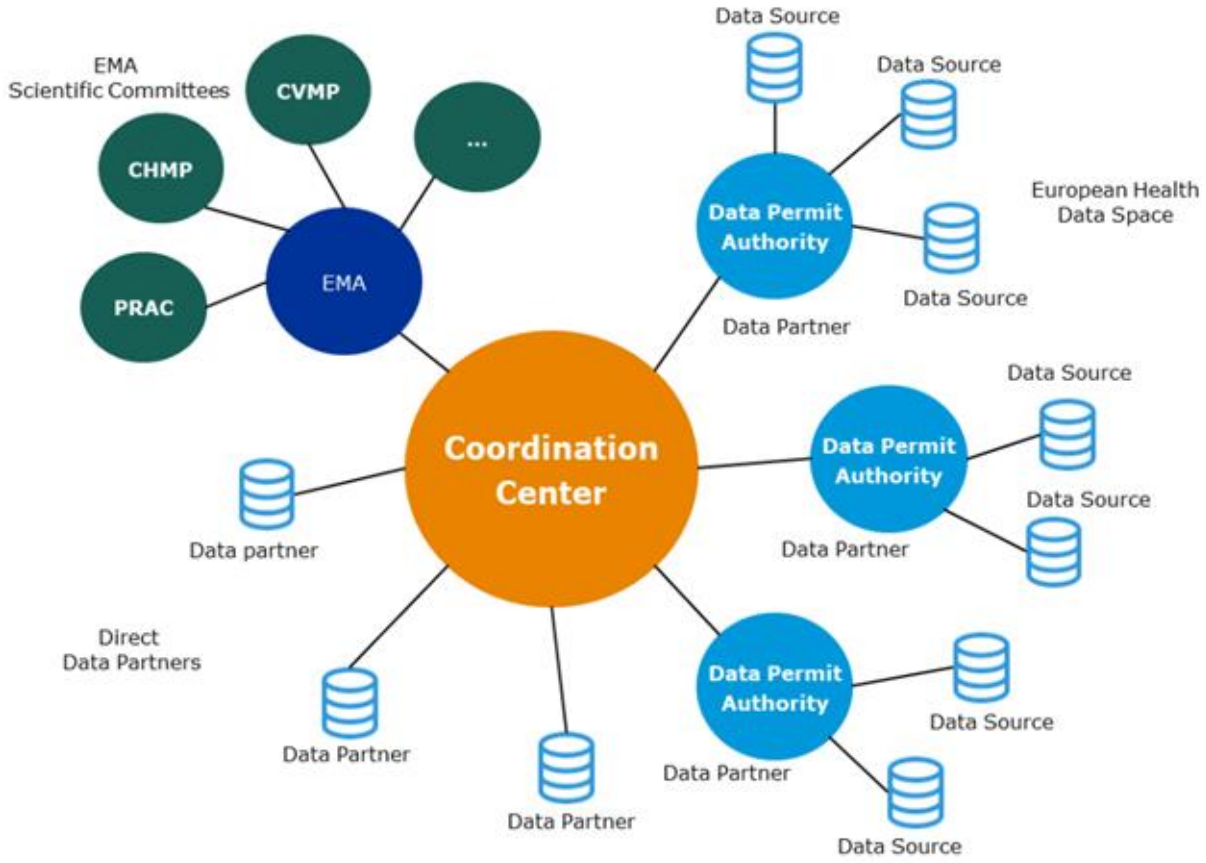
Benefits to Participating Sites

- Access to large, diverse patient populations
- Maintain direct control of your site's clinical data, share only aggregate data
- Access to IQVIA data enrichment programs to enhance site data (e.g. NLP tools, linkage services)
- Ability for researchers to externally validate single-center findings

Data Analysis and Real-World Interrogation Network (DARWIN EU®)

A paradigm shift for the use of real-world health data for regulatory purpose in the EU

DARWIN EU® is a federated **network** of **data**, **expertise**, and services that supports better decision-marking throughout the product lifecycle by generating reliable **evidence from real world healthcare data**



To learn more: https://www.ema.europa.eu/en/about-us/how-we-work/big-data/data-analysis-real-world-interrogation-network-darwin-eu

China Government's Guides on RWE & RWD

From Center for Drug Evaluation (CDE), National Medical Products Administration (NMPA)

- **1st guide** was released in Jan 2020, introducing the definition, data source requirement, design, and evaluation of using RWE for drug effectiveness study and safety monitoring.

- **2nd guide** was released in Aug 2020, focusing on the details and importance of the source, safety, curation, quality assurance and maintenance of RWD, so that reliable RWE could be produced

国家药品监督管理局药品审评中心
CENTER FOR DRUG EVALUATION, NMPA

此页面上的内容需要较新版本的 Adobe Flash Player。

当前位置：新闻中心>>工作动态>>通知公告>>新闻正文

关于公开征求《用于产生真实世界证据的真实世界数据指导原则（征求意见稿）》意见的通知

发布日期：20200803

为进一步指导和规范申办者利用真实世界数据生成真实世界证据支持药物研发，我中心组织起草了《用于产生真实世界证据的真实世界数据指导原则（征求意见稿）》，现在中心网站予以公示，以广泛听取各界意见和建议，欢迎各界提出宝贵意见和建议，并请及时反馈给我们。

征求意见时限为自发布之日起2个月。

您的反馈意见请发到以下联系人的邮箱：

联系人：高丽丽、赵骏

联系方式：gaoll@cde.org.cn，zhaojun@cde.org.cn

感谢您的参与和大力支持。

国家药品监督管理局药品审评中心

2020年8月3日

附件 1：	《用于产生真实世界证据的真实世界数据指导原则（征求意见稿）》.docx
附件 2：	《用于产生真实世界证据的真实世界数据指导原则（征求意见稿）》起草说明.doc

China Government's Guides on RWE & RWD

CDM & OHDSI Citations in the 2nd Guide, Section 4 – Real World Data Curation

CDM Introduction in Guide:

- Under multidisciplinary collaboration, CDM was created with standardized structure, format and vocabulary, to achieve multi-center data integration and collaboration.

References in Guide:

- EMA. A Common Data Model for Europe – Why? Which? How? https://www.ema.europa.eu/en/events/common-data-model-europe-why-which-how
- OHDSI – Observational Health Data Sciences and Informatics, https://www.ohdsi.org

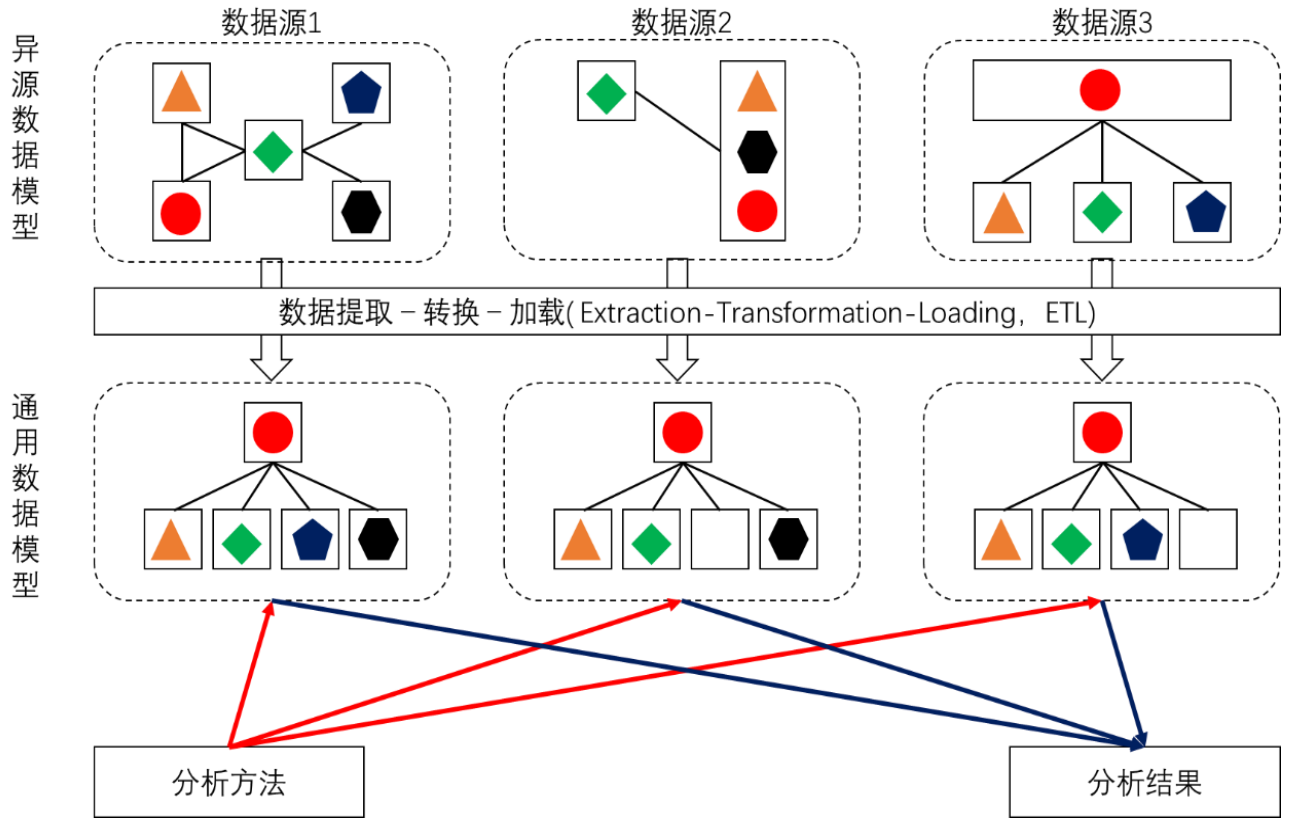
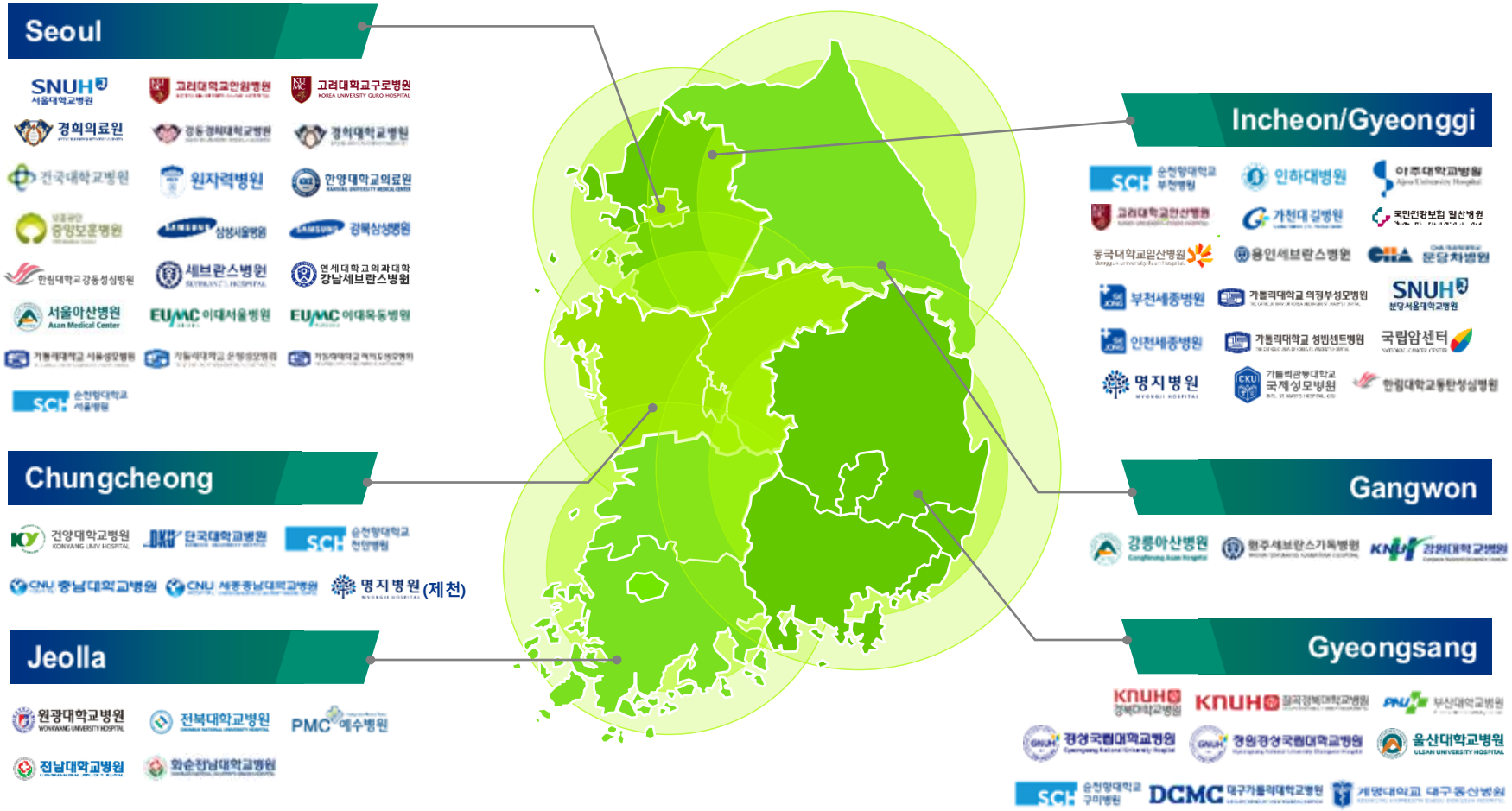


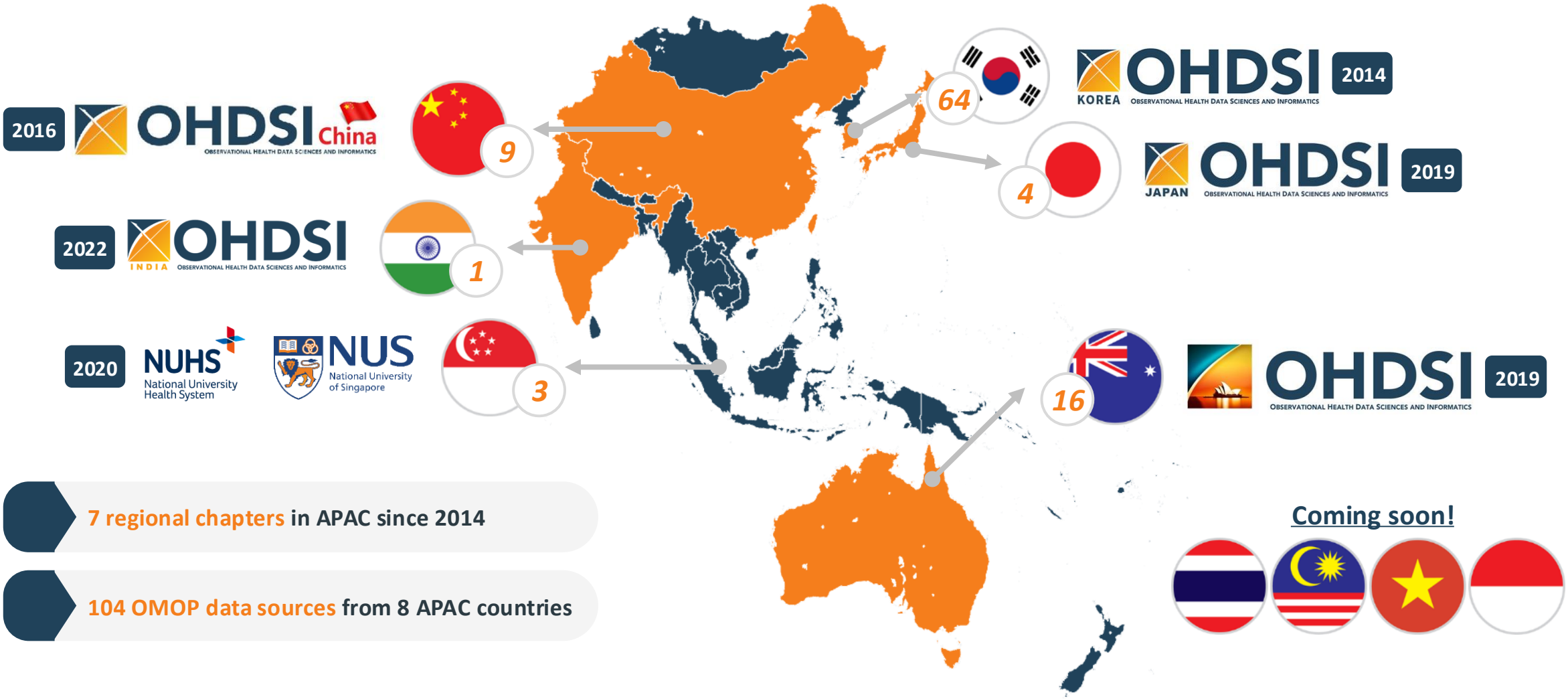
Fig. 2 in Guide – Diagram on Converting Source Data to CDM

Korean Government Initiatives



- Federated OMOP network of 62 hospitals and >76M patients
- Participation from secondary and tertiary hospitals nationwide
- Good representation of each of the provinces of South Korea
- Funded by series of grants from the Korean government
- Foster collaborative research and evidence-sharing ecosystem

OHDSI APAC



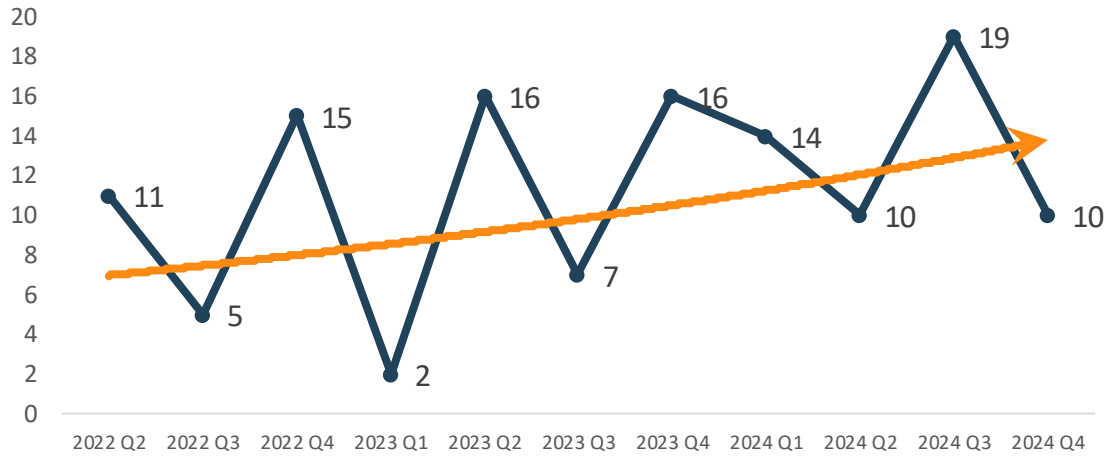
7 regional chapters in APAC since 2014

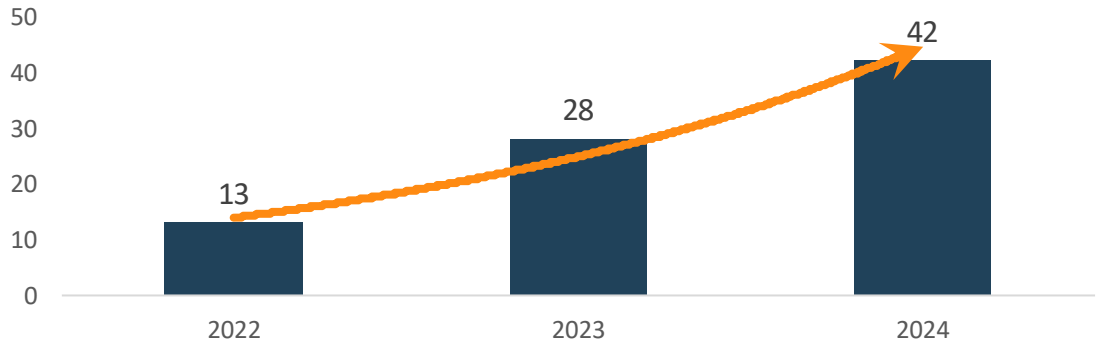
104 OMOP data sources from 8 APAC countries

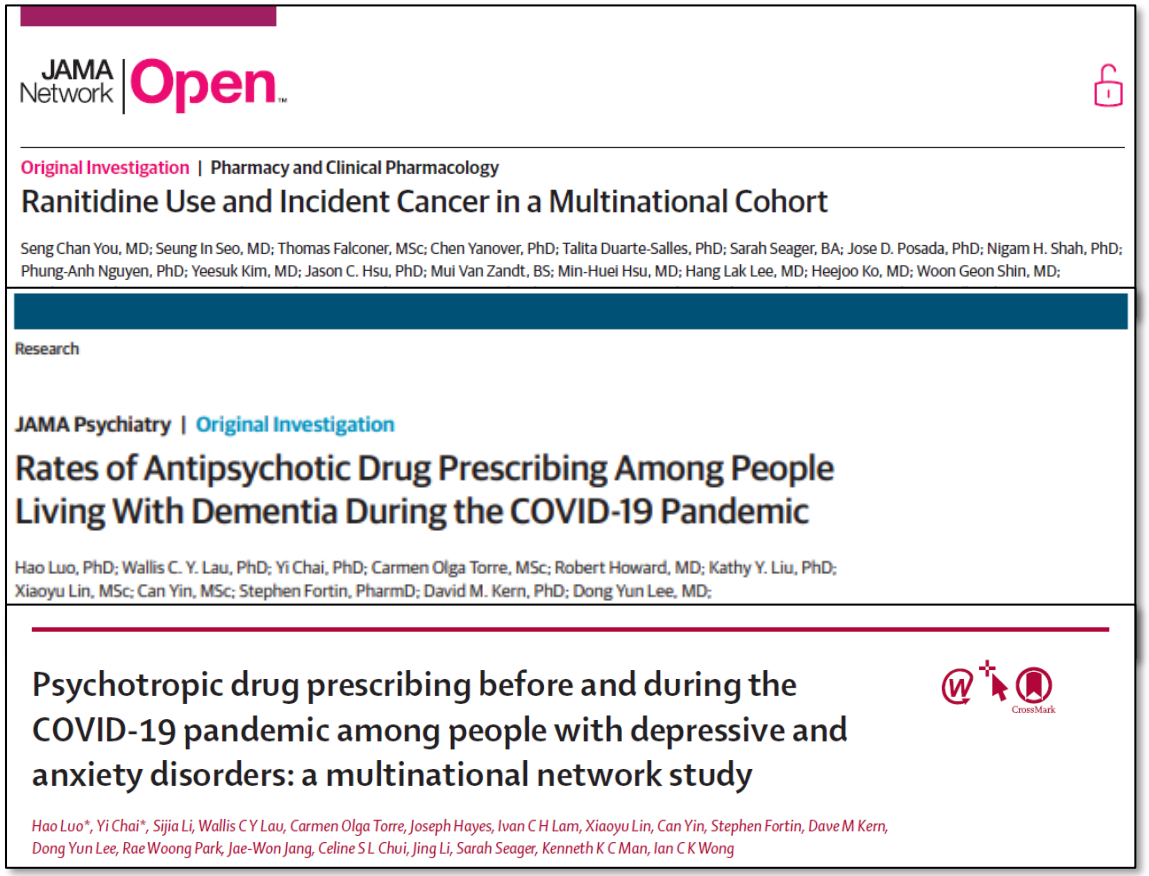
Coming soon!

APAC Studies

Increasing trend of publications and abstract submissions from APAC







JAMA Network Open

Original Investigation | Pharmacy and Clinical Pharmacology

Ranitidine Use and Incident Cancer in a Multinational Cohort

Seng Chan You, MD; Seung In Seo, MD; Thomas Falconer, MSc; Chen Yanover, PhD; Talita Duarte-Salles, PhD; Sarah Seager, BA; Jose D. Posada, PhD; Nigam H. Shah, PhD; Phung-Anh Nguyen, PhD; Yeesuk Kim, MD; Jason C. Hsu, PhD; Mui Van Zandt, BS; Min-Huei Hsu, MD; Hang Lak Lee, MD; Heejoo Ko, MD; Woon Geon Shin, MD;

Research

JAMA Psychiatry | Original Investigation

Rates of Antipsychotic Drug Prescribing Among People Living With Dementia During the COVID-19 Pandemic

Hao Luo, PhD; Wallis C. Y. Lau, PhD; Yi Chai, PhD; Carmen Olga Torre, MSc; Robert Howard, MD; Kathy Y. Liu, PhD; Xiaoyu Lin, MSc; Can Yin, MSc; Stephen Fortin, PharmD; David M. Kern, PhD; Dong Yun Lee, MD;

Psychotropic drug prescribing before and during the COVID-19 pandemic among people with depressive and anxiety disorders: a multinational network study

Hao Luo, Yi Chai*, Sijia Li, Wallis C Y Lau, Carmen Olga Torre, Joseph Hayes, Ivan C H Lam, Xiaoyu Lin, Can Yin, Stephen Fortin, Dave M Kern, Dong Yun Lee, Rae Woong Park, Jae-Won Jang, Celine S L Chui, Jing Li, Sarah Seager, Kenneth K C Man, Ian C K Wong*

100+ publications authored by APAC collaborators including numerous papers in high-impact journals

2025 APAC Study

Association Between Fasting Plasma Glucose Levels and Annual Hospitalization Days *(Fudan University)*

Gastrointestinal Risk of GLP-1 Receptor Agonists versus SGLT-2 and DPP-4 Inhibitors in Type 2 Diabetes *(Peking University)*

Studying the Disease Trajectory of Type 2 Diabetes with Transformer-based Model *(University of Science and Technology of China)*

- Studies initiated and led by China
- Ongoing work to align research questions to OHDSI's standardized methods and analytical frameworks

Titans in APAC

Summary

1. **Open Source**
   CDM, tools, methods, and documentation all publicly available
2. **Standardization**
   Standard CDM, vocabulary/ontology, tools, methods, data quality, and documentation
3. **Research Community**
   Large research community with multiple stakeholders and disciplinaries
4. **Multi-country/multi-center research**
   Large scale research using standardized tools and methods


OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Thank you!