



The
Journey Begins



OHDSI

2015 ANNUAL REPORT



Table of Contents

A message from George Hripcsak	2
Our Mission & Vision	3
Our Values	4
The OHDSI Community	5
North America & Europe	6
Asia, Australia, South America & Africa	7
Vocabulary	8
Common Data Model & Data Network	8
Software Development	9
Methods Library	13
OHDSI Research	14
OHDSI Symposium	16
2015 Funding	19





A message from George Hripcsak

As head of OHDSI's coordinating center at Columbia University, it is my honor to present OHDSI's 2015 annual report. This report gives an overview of OHDSI's efforts to generate reliable evidence to inform medical decision-making.

I'm incredibly proud of the amazing progress our community has made in methodological research and open-source software development. Thanks to collaborative efforts across the OHDSI community, we've seen significant growth in the number of new collaborators and data partners. Much of this growth was bolstered by the first annual OHDSI symposium, held in October 2015. The symposium offered our community a wonderful opportunity to showcase our research and software development to date. It also generated an appetite for more collaborative research in 2016.

In response to the tremendous feedback received throughout the symposium, the OHDSI community is committed to leveraging the OHDSI data network to generate high quality evidence and support better health care decision making.

George Hripcsak, MD, MS
Vivian Beaumont Allen Professor
of Biomedical Informatics
Chair, Dept of Biomedical Informatics
Director, Medical Informatics Services





OUR MISSION

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

OUR VISION

A world in which observational research produces a comprehensive understanding of health and disease.



Our Values

- **Innovation:** Observational research is a field which will benefit greatly from disruptive thinking. We actively seek and encourage fresh methodological approaches in our work.
- **Reproducibility:** Accurate, reproducible, and well-calibrated evidence is necessary for health improvement.
- **Community:** Everyone is welcome to actively participate in OHDSI, whether you are a patient, a health professional, a researcher, or someone who simply believes in our cause.
- **Collaboration:** We work collectively to prioritize and address the real world needs of our community's participants.
- **Openness:** We strive to make all our community's proceeds open and publicly accessible, including the methods, tools and the evidence that we generate.
- **Beneficence:** We seek to protect the rights of individuals and organizations within our community at all times.



Welcome

to the

OHDSI *Community*

Over the past year, the OHDSI collaborative has grown into a vibrant community of over 140 collaborators from 16 countries. Our community comprises clinical researchers, epidemiologists, bio-statisticians, computer scientists, healthcare payers and providers, government officials, entrepreneurs and industry leaders. With such a diverse range of expertise from academia, government and industry, the OHDSI community is home to a wealth of knowledge which continues to grow everyday.

In 2015, the collaborative welcomed newcomers from Brazil, South Africa, Kenya, Japan, Germany, and Denmark. We can now proudly say our community represents each continent and look forward to strengthening our partnerships with collaborators across the world in 2016.

<http://www.ohdsi.org/who-we-are/collaborators/>



North America



Home to the OHDSI coordinating center at Columbia University in New York, North America is also home to over 110 OHDSI collaborators from across the United States and Canada.

Europe

2015 saw continuous growth within our European community. Over the past 12 months, OHDSI welcomed 20 European collaborators and data partners.

Europe is also home to two key collaboration sites, the WHO Monitoring Center in Uppsala, Sweden and Erasmus MC in the Netherlands. With their support, we look forward to building up OHDSI's presence in Europe.





Asia & Australia

OHDSI's Pacific community has seen tremendous growth and activity during the second half of 2015. We currently have over 15 collaborators working to incorporate their data into the OMOP common data model.

Africa & South America

A big goal in 2015 was to have OHDSI representation on each continent. In July, the OHDSI mission finally reached South America when our lead collaborators led several sessions at the 2015 MedInfo conference.

Thanks to the support of our strategic partnership with OpenMRS, an open source EHR system, we have successfully taken on our first African collaborators.





OHDSI Vocabulary

The Standard Vocabulary is a foundational tool initially developed by some of us at OMOP that enables transparent and consistent content across disparate observational databases, and serves to support the OHDSI research community in conducting efficient and reproducible observational research.

Common Data Model

A common data model allows for the systematic analysis of disparate observational databases. The concept behind this approach is to transform data contained within disparate databases into a common format (data model), and then perform systematic analyses using a library of standard analytic routines that have been written based on the common format. The Observational Medical Outcomes Partnership (OMOP) CDM, now entering its fifth version, offers a solution unlike any other. OMOP found that disparate coding systems can be harmonized—with minimal information loss—to a standardized vocabulary.

With the continued development of the CDM, the OHDSI data network has grown to include 84 databases from 11 countries covering over 600 million patients:

http://www.ohdsi.org/web/wiki/doku.php?id=resources:data_network



2015 Software Developments

Since OHDSI's early beginnings in 2014, the majority of our focus has been on building up our open-source software tools. These tools have been designed to run against the OMOP common data model and aim to facilitate high-quality evidence generation.

2015 software developments include CALYPSO, CIRCE, LAERTES, HERACLES, CHAOS, IRIS, APHRODITE and PENELOPE.





2015 Software Developments

CALYPSO for study population exploration

CALYPSO (Criteria Assessment Logic for Your Population Study in Observational data) is a web user interface to define, instantiate and evaluate a study population and the implications of inclusion criteria.

CIRCE for cohort definition

CIRCE (Cohort Inclusion and Restriction Criteria Expression) is a cohort definition and syntax compiler tool for the latest version of the OMOP common data model.

LAERTES Knowledge Base

LAERTES (Large-scale Adverse Effects Related to Treatment Evidence Standardization) is an open-source repository of standardized evidence about drug-outcome relationships from disparate sources, including published literature, product labeling, spontaneous adverse event reporting, and existing bio-medical ontologies. The knowledge base serves as the primary source to enable the construction of test cases (positive controls and negative controls) to facilitate systematic evaluation of method performance.



2015 Software Developments

HERACLES for quality of care

HERACLES (Health Enterprise Resource And Care Learning Exploration System) is an application that allows you to explore healthcare quality, cost, and practice patterns using the OMOP Common Data Model. HERACLES provides high-level visualization tools and deep-dive capabilities to look at standardized quality metrics (e.g., NQF) as well as utilization across a variety of patient cohorts.

PENELOPE for evaluation of product labeling

PENELOPE (Personalized Exploratory Navigation and Evaluation for Labels Of Product Effects) connects traditional health information resources with data generated by the OHDSI network. Specifically, we want to delve into the drug product labeling for prescribed medications and show how OHDSI can complement and illuminate the safety information found in these important documents.



2015 Software Developments

CHAOS Semantic Browser

CHAOS (Common_data_model Hierarchy and Attributes Overview System) provides an interactive overview of CDM tables, table fields and relationships between fields. Definitions of each table and field can be reviewed by hovering over corresponding items in the graphical representation of the CDM. For fields, two types of mouse over information can be chosen depending on radio button setting: definitions or data types.

APHRODITE for phenotype development

APHRODITE (Automated PHenotype Routine for Observational Definition Identification Training and Evaluation). Aphrodite uses OHDSI standard concept ids specific to the phenotype of interest to create “silver standard” training corpora to construct phenotype models. Aphrodite uses such silver standard corpora, in conjunction with expert knowledge codified in existing ontologies and a comprehensive representation of the patient clinical record, to learn phenotype models.

IRIS for Population Summary

IRIS provides a high-level descriptive summary of a population within an OMOP CDM-compliant database.



OHDSI

Methods Library

Over the past year, we have been developing a methods library to support our open-source tools for large-scale analytics, including population-level estimation and patient-level prediction.

Estimation methods

Cohort Method

New-user cohort studies using large-scale regression for propensity and outcome models

Self-Controlled Case Series

Self-Controlled Case Series analysis using few or many predictors, includes splines for age and seasonality.

Self-Controlled Cohort

A self-controlled cohort design, where time preceding exposure is used as control.

IC Temporal Pattern Disc.

A self-controlled design, but using temporal patterns around other exposures and outcomes to correct for time-varying confounding.

Prediction methods

Patient Level Prediction

Use a large set of features derived from data in the CDM to build predictive models for user-specified outcomes.

Method characterization

Empirical Calibration

Use negative control exposure-outcome pairs (where relative risk is assumed to be 1) to profile and calibrate a particular analysis design.

Method Evaluation

Use real data and established reference sets as well as simulations injected in real data to evaluate the performance of methods.

Supporting packages

Database Connector

Connect directly to a wide range of database platforms, including SQL Server, Oracle, and PostgreSQL.

Sql Render

Generate SQL on the fly for the various SQL dialects.

Cyclops

Highly efficient implementation of regularized logistic, Poisson and Cox regression.

Ohdsi R Tools

Support tools that didn't fit other categories, including tools for maintaining R libraries.

Our population-level estimation work is focused on developing open-source software for safety surveillance and comparative effectiveness. Also released in 2015 were tools for patient-level prediction and other study designs such as self-controlled case series and self-controlled cohorts, as well as tools for evaluating and calibrating population-level estimation methods.



OHDSI Research

2015 OHDSI research titles include:

- Treatment pathways in chronic disease
- Transforming the national department of veterans affairs data warehouse to the OMOP common data model
- Converting the data in the U.S. CMS virtual research data center to the OHDSI common data model version
- Applying the OMOP common data model to survey data
- Mapping Korean national insurance billing code to OMOP code for drugs used in a Korean tertiary teaching hospital
- An algorithm for mapping local measurement concepts into OMOP vocabulary using logical observation identifiers names and codes (LOINC®) terminology
- Establishing interoperability standards between OMOP CDM v4, v5, and PCORnet CDM v1
- Using semantic queries for cohort discovery across research networks
- Lift your anchors and begin the OHDSI with APHRODITE
- Discovering the hidden risk factors: An empirical evaluation of incorporating feature-learning methods into a risk model framework using the OMOP CDM
- Accuracy of an automated knowledge base for identifying adverse drug reactions



OHDSI Research

2015 OHDSI research titles include:

- How high can we go? Evaluating massively high-dimensional propensity score and outcome models in large-scale observational studies
- Size comparison of 17 CDM datasets using IRIS tool
- Identifying and understanding data quality issues in a pediatric distributed research network
- Lessons from CIRCE implementation of eMERGE phenotype definitions into actionable CDM v5 SQL queries
- Storing, sharing, and using algorithms for implementing clinical studies: The jigsaw algorithm repository
- A climate wide journey to explore mechanisms underlying birth month disease risk associations: A call for collaboration
- Determination of pregnancy episodes and outcomes within a distributed network of observational databases
- Trajectories in diabetes mellitus type II treatment intensification with massive observational data
- Exploration of the epidemiology of endometriosis



2015 OHDSI Symposium

The most anticipated event of 2015 was the first annual OHDSI symposium. The symposium was a full day event with morning and afternoon sessions. The morning was comprised of three presentations which gave an overview of the OHDSI program, a presentation of the first OHDSI network research study and the unveiling of OHDSI's latest open-source software tools. The aim of these presentations was to inform participants about OHDSI's history, mission and future goals. Each presentation included a Q&A session to engage the broader community and ask participants how their own research or policy questions could be answered through collaboration with OHDSI.

The afternoon was more hands on. It included demonstrations of OHDSI software tools, a poster session and two panel discussions designed to facilitate audience feedback. The demonstrations aimed for maximum impact by providing tangible examples of how OHDSI conducts collaborative research. The poster session gave OHDSI collaborators a chance to showcase their individual work and highlight the diverse range of research topics covered across the community.

The symposium informed not only the next symposium, but also the direction of OHDSI research. 2015 was a year of tremendous development for the OHDSI community and the core analytics infrastructure. The community expressed shared interest in focusing further on applying the open-source tools that have been developed to addressing real clinical questions in 2016 and beyond. We believe that this is a natural progression, where early work has to build the infrastructure to get started, but the collaborative must quickly emphasize concrete results to provide value and remain focused on the ultimate goal.



2015 OHDSI Symposium

"An essential event for introduction for newcomers"



"Loved the live demos"

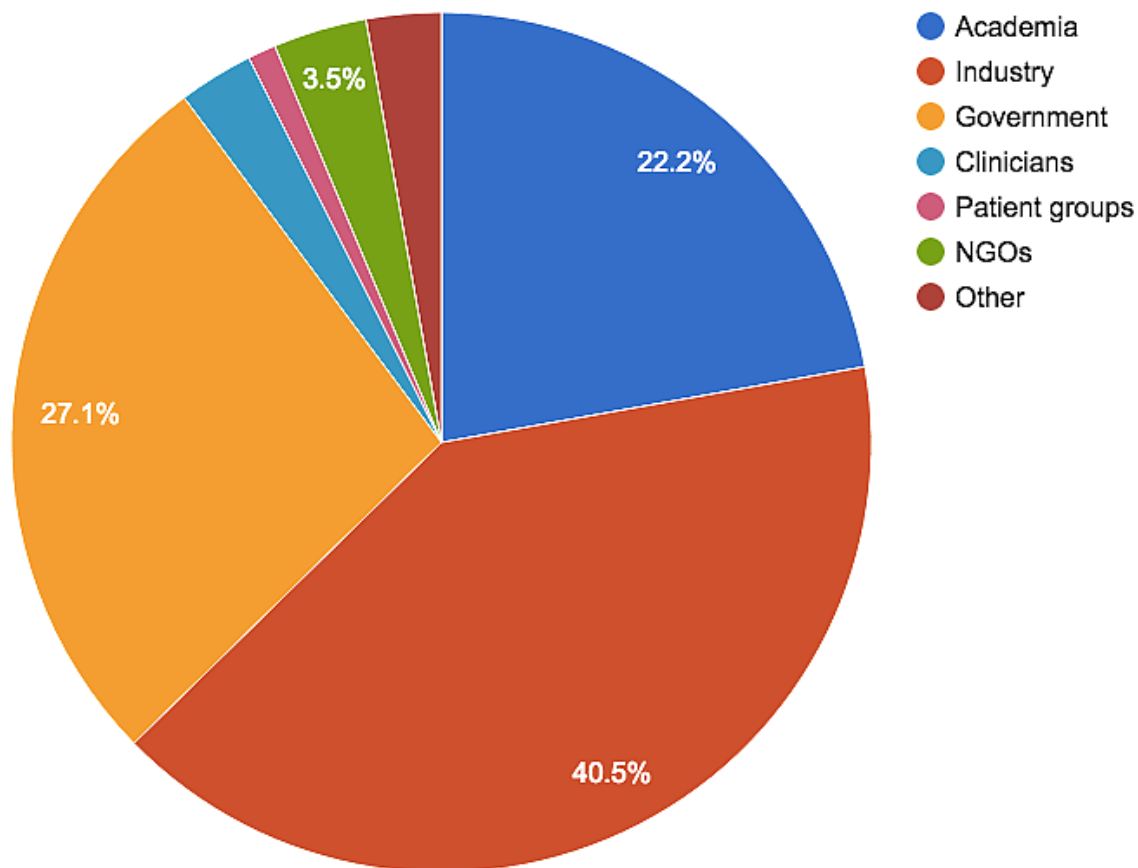


"An excellent opportunity to meet people driving research in the community"



2015 OHDSI Symposium

The 2015 symposium was attended by almost 300 participants from:



The symposium also included a collaborator showcase of:

- 28 Poster Presentations
- 6 Software Demonstrations
- ETL 101: How to convert data to the CDM
- OHDSI 101: How to 'Join the Journey'

<http://www.ohdsi.org/2015-ohdsi-symposium-materials/>



Thank-you to all OHDSI funders for a successful 2015

- Amgen
- AstraZeneca
- AHRQ
- Bayer
- IMS Health
- Janssen R&D
- Korea Health Industry Development
Institute
- Merck
- NIH
- NLM
- NSF
- PCORI
- Pfizer
- Regenstrief Institute
- Takeda