Collaborator Showcase Honorees

OHDSI Community Call
Nov. 14, 2023 • 11 am ET
<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov. 14</td>
<td>Collaborator Showcase Honorees</td>
</tr>
<tr>
<td>Nov. 21</td>
<td>Showcase Software Demos</td>
</tr>
<tr>
<td>Nov. 28</td>
<td>TBA</td>
</tr>
<tr>
<td>Dec. 5</td>
<td>Recent Publications</td>
</tr>
<tr>
<td>Dec. 12</td>
<td>Happy Birthday OHDSI! Where Have We Come In 10 Years, and in 12 Months?</td>
</tr>
<tr>
<td>Dec. 19</td>
<td>Holiday-Themed Goodbye to 2023!</td>
</tr>
</tbody>
</table>
Three Stages of The Journey

Where Have We Been?
Where Are We Now?
Where Are We Going?
Congratulations to the team of Erica Voss, Clair Blacketer, Sebastiaan van Sandijk, Maxim Moinat, Michael Kallfelz, Michel van Speybroeck, Daniel Prieto-Alhambra, Martijn Schuemie, and Peter Rijnbeek on the publication of European Health Data & Evidence Network—learnings from building out a standardized international health data network in JAMIA.

Research and Applications

European Health Data & Evidence Network—learnings from building out a standardized international health data network

Erica A. Voss @, MPH,1,2,3 Clair Blacketer @, MPH,1,2,3, Sebastiaan van Sandijk, MSc,1,4, Maxim Moinat, MSc,1,2, Michael Kallfelz, MD,4, Michel van Speybroeck, MSc,4, Daniel Prieto-Alhambra, PhD,2,3, Martijn J. Schuemie, PhD,1,2,3, Peter R. Rijnbeek, PhD,1,2

1OHDSI Collaborators, Observational Health Data Sciences and Informatics (OHDSI), New York, NY, United States 2Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, the Netherlands 3Janssen Pharmaceutical Research and Development LLC, Raritan, NJ 08869, United States 4Diagnostics Data Services, Prague, Czech Republic *Centre for Statistics in Medicine, NCEIMS, University of Oxford, Oxford, United Kingdom 5Department of Biostatistics, University of California, Los Angeles, CA 90095, United States

*Corresponding author: Erica A. Voss, MPH, Janssen Research & Development - Epidemiology, 620 US Highway 322, Raritan, NJ 08869 (evooss3i@jnj.com)

Abstract

Objective: Health data standardized to a common data model (CDM) simplifies and facilitates research. This study examines the factors that make standardized observational health data to the Observational Medical Outcomes Partnership (OMOP)/CDM successful.

Materials and methods: Twenty-five data partners (DPs) from 11 countries received funding from the European Health Data Evidence Network (EHDEL) to standardize their data. Three surveys, DataQualityDashboard results, and statistics from the conversion process were analyzed qualitatively and quantitatively. Our measures of success were the total number of days to transform source data into the OMOP/CDM and participation in network research.

Results: The health data converted to CDM represented more than 133 million patients. 100%, 99%, and 84% of DPs took Surveys 1, 2, and 3. The median duration of the 6 key extract, transform, and load (ETL) processes ranged from 4 to 115 days. Of the 25 DPs, 21 DPs were considered applicable for analysis of which 52% standardized their data on time, and 48% participated in an international collaborative study.

Discussion: This study shows that the consistent workflow used by EHDEL proves appropriate to support the successful standardization of observational data across Europe. Over the 25 successful transformations, we confirmed that getting the right people for the ETL is critical and vocabulary mapping requires specific expertise and support of tools. Additionally, we learned that teams that proactively prepared for data governance issues were able to avoid considerable delays improving their ability to finish on time.

Conclusions: This study provides guidance for future DPs to standardize to the OMOP CDM and participate in distributed networks. We demonstrate that the Observational Health Data Sciences and Informatics community must continue to evaluate and provide guidance and support for what ultimately develops the backbone of how community members generate evidence.

Key words: OMOP common data model, observational data, data standardization
OHDSI Shoutouts!

Congratulations to the team of Soobeen Seol, Jung Ran Choi, Byungjin Choi, Sungryeal Kim, Ja Young Jeon, Ki Nam Park, Jae Hong Park, Min Woo Park, Young-Gyu Eun, Jung Je Park, Byung-Joo Lee, Yoo Seob Shin, Chul-Ho Kim, Rae Woong Park and Jeon Yeob Jang on the publication of Effect of statin use on head and neck cancer prognosis in a multicenter study using a Common Data Model in Scientific Reports.
Congratulations to the team of Jung-Yeon Choi, Sooyoung Yoo, Wongeun Song, Seok Kim, Hyunyoung Baek, Jun Suh Lee, Yoo-Seok Yoon, Seonghae Yoon, Hae-Young Lee, Kwang-II Kim on the publication of Development and Validation of a Prognostic Classification Model Predicting Postoperative Adverse Outcomes in Older Surgical Patients Using a Machine Learning Algorithm: Retrospective Observational Network Study in the Journal of Medical Internet Research.
OHDSI Shoutouts!

OHDSI RWE Revolution:
Igniting Data Modernization with Harmonized Standards for Cutting Edge Health Research

11-Nov-2023

ohdsi.org/ohdsi-news-updates/
OHDSI Shoutouts!
Three Stages of The Journey

Where Have We Been?
Where Are We Now?
Where Are We Going?
## Upcoming Workgroup Calls

<table>
<thead>
<tr>
<th>Date</th>
<th>Time (ET)</th>
<th>Meeting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuesday</td>
<td>12 pm</td>
<td>Common Data Model</td>
</tr>
<tr>
<td>Tuesday</td>
<td>3 pm</td>
<td>OMOP CDM Oncology Outreach/Research Subgroup</td>
</tr>
<tr>
<td>Tuesday</td>
<td>6 pm</td>
<td>Eyecare &amp; Vision Research</td>
</tr>
<tr>
<td>Wednesday</td>
<td>11 am</td>
<td>Perinatal &amp; Reproductive Health</td>
</tr>
<tr>
<td>Wednesday</td>
<td>12 pm</td>
<td>Health Equity Journal Club</td>
</tr>
<tr>
<td>Wednesday</td>
<td>4 pm</td>
<td>Vulcan/OHDSI Meeting</td>
</tr>
<tr>
<td>Wednesday</td>
<td>7 pm</td>
<td>OMOP CDM Oncology Vocabulary/Development Subgroup</td>
</tr>
<tr>
<td>Thursday</td>
<td>9 am</td>
<td>Phenotype Development &amp; Evaluation</td>
</tr>
<tr>
<td>Thursday</td>
<td>12 pm</td>
<td>HADES</td>
</tr>
<tr>
<td>Thursday</td>
<td>7 pm</td>
<td>Dentistry</td>
</tr>
<tr>
<td>Friday</td>
<td>9 am</td>
<td>GIS – Geographic Information System General</td>
</tr>
<tr>
<td>Friday</td>
<td>10:30 am</td>
<td>Open-Source Community</td>
</tr>
<tr>
<td>Friday</td>
<td>11 am</td>
<td>Clinical Trials</td>
</tr>
<tr>
<td>Monday</td>
<td>10 am</td>
<td>Healthcare Systems Interest Group</td>
</tr>
<tr>
<td>Monday</td>
<td>11 am</td>
<td>Data Bricks User Group</td>
</tr>
</tbody>
</table>
Global Symposium Homepage

2023 OHDSI Symposium
Oct. 20-22 • East Brunswick, New Jersey

The 2023 OHDSI Global Symposium welcomed more than 400 of our global collaborators together for three days of sharing research, forging new connections and pushing forward the OHDSI mission of improving health by empowering a community to collaboratively generate evidence to advance better health decisions and better care.

This page will be home to all materials from the global symposium. Check back in the coming days for all video presentations from the event.
#OHDSISymposium OHDSI2023

State of the Community

Various leaders within OHDSI shared a presentation on the state of the community, with specific focuses on data standards, vocabulary development, and open-source software. Speakers included:

- George Hruscak, Columbia University
- Ciaran Blackmore, Johnson & Johnson
- Alexander Darovoy, Odyssey Data Services
- Katy Sadovsky, Boucherville Innomed
- Peter Ripplinger, Eastman MC
- Mengling ‘Harry’ Peng, National University of Singapore

State of the Community Slides

2023 Global Collaborator Showcase

Observational Data Standards & Management

- FICOMP: a consistent federated data space leverage Grace-Takhtena, Porta Kostromitova, Pia Tajmar, Bane Kukuloski, Gustavo Rangel, Anna Harnois, Pedroni Compo, Oscar Brocc, Line Febbe, Claudia Barquet, Marco Huesber, Toni Billak, Birgitta Soldén, Paolo Maffia, Gioia Rinaldi, Dorota Wronska, Johann Gmeiner, Peter Gräss, Kerstin Manteufel, Peter Spies, Claudia Cusano, Paolo Zanolla, and Wiebke Lohaus.
- OpenCQM: an open-source platform for the systematic evaluation of EHR data in epidemiological research.

Tutorial: Introduction to OHDSI

- Survey from data to evidence can be challenging alone but is greatly enhanced through collaboration. In this half-day tutorial, we will introduce newcomers to OHDSI. Specifically, about the tools, practices, and open-source approach to evidence generation that makes the OHDSI community has developed and evolved over the past decade.

Faculty will highlight the ways community members can participate as well as receive values from the community’s output. The course will include topics such as open community data standards – including the OHDSI Common Data Model and OHDSI Standardized Vocabularies, open-source analytic tools

ohdsi.org/OHDSI2023
OHDSI HADES releases: Database Connector 6.3.0

Database Connector

Introduction

This R package provides functionality for connecting to various DBMSs. Together with the SqlRender package, the main goal of DatabaseConnector is to provide a uniform interface across database platforms: the same code should run and produce equivalent results, regardless of the database backend.

Features

- Create connections to the various database platforms:
  - Microsoft SQL Server
  - Oracle
  - PostgreSQL
  - Microsoft Parallel Data Warehouse (a.k.a. Analytics Platform System)
  - Amazon Redshift
  - Apache Impala
  - Google BigQuery
  - IBM Netezza
OHDSI HADES releases: DeepPatientLevelPrediction 2.0.1

ResultModelManager

Introduction

RMM is an R package designed to handle common ohdsi results data management functions by providing a common API for data model migrations and definitions.

System Requirements

Requires R. Some of the packages used by ResultModelManager require Java.

Installation

1. See the instructions here for configuring your R environment, including Java.
2. In R, use the following commands to download and install ResultModelManager:
Integration of Clinical and Genomic Data Mapped to the OMOP Common Data Model in a Federated Data Network in Belgium

(Tatjana Jatsenko, Murat Akand, Joris Robert Vermeesch, Dries Rombaut, Michel Van Speybroeck, Martine Lewi, Valerie Vandeweerd)

**Background**

- Enriching clinical patient data with omics data can help understand and predict disease and treatment outcomes in bladder cancer.
- Omics and non-omics data is often unaligned and inaccessible due to healthcare system fragmentation and privacy concerns.
- Federated data platforms enable data accessibility, usability, and security while complying with regulations.
- The OMOP Common Data Model allows privacy-preserving, large-scale genotype-phenotype research.
- Athena is a federated data network for multiple myeloma and bladder cancer in Belgium.
- This study focuses on non-muscle invasive bladder cancer (NMIBC) and the integration of clinical and genomics data.

**Methods**

- Extraction of de-identified clinical and genomic data from different databases of multiple healthcare institutions across Belgium.
- Collection of clinical data from EHR systems and bladder cancer patients and mapping to OMOP V3.3.
- By date analyzed DNA and RNA from 102 FFPE tumor samples from UZ Leuven using T50500 assay for genomic profiling.

**Results**

- Proof of concept for combining clinical data and comprehensive genomic data using OMOP.
- A lightweight method to map all genomic data to the measurement entities in OMOP.
- Limitations:
  - Loss of granularity, specifically for the alterations of unknown clinical significance, due to limited standardized set of genomic concepts.
  - Remaining challenge of integrating the data into one single OMOP instance and matching it to the correct person ID.

**Conclusions**

- A federated data network that combines clinical and genomic data and maps it to the OMOP Common Data Model.
- The federated approach keeps the data secure at each institution.
- The common data model enables researchers and clinicians to do research on a larger scale.
- Future efforts will aim to grow the network with more institutions and more diverse datasets to improve the representativeness and generalizability of research findings.

**References**

Estimating Observable Time in the Absence of Defined Enrollment

(Clair Blacketer, Patrick Ryan, Frank DeFalco, Martijn Schuemie, Peter Rijnbeek)

When will I see you again? Estimating observable time in databases without defined enrollment

# PRESENTER: Clair Blacketer

INTRO

- Retrospective studies that make use of existing observational data must define observable time for each patient.
- This is relatively simple in databases that have the notion of defined enrollment, like US claims, much more difficult in databases that are primarily encounter based.
- This pilot study explores multiple approaches for the creation of the observability period and the implications of these definitions.

METHODS

1. Used National Commercial Claims and Encounters (CCI) database mapped to OMOP COM v5.4, which utilizes defined health plan enrollment as the observation period (gold standard).
2. We then used observed health care events to create alternative definitions of the observation period.
3. We created observation eras by applying varying persistence and surveillance windows to the events (Figure S).

RESULTS

- To reduce bias, methods to estimate observable time should choose a balance of persistence and surveillance windows based on the types of data available.

LIMITATIONS

- Analyses were conducted in only one database covering one country and one type of data (535 claims).
- Future work includes creating eras of observational times based on age and sex and running the present analysis and conducting the analyses across multiple data sources.

#OHDSISocialShowcase

This Week

TUESDAY

To reduce bias, methods to estimate observable time should choose a balance of persistence and surveillance windows based on the types of data available.
Modeling Decisions and Heterogeneity in Defining Aortic Diseases: Implications for Observational Studies and Phenotype Characterization

(Evan Minty, Jack Janetzi, James P. Gilbert, Jung Ho Kim, Jung Ah Lee, Elsie Ross, Nicole Pratt, Gowtham Rao, Seng Chan You)

We demonstrate the utility of RWD from the Edge Tool Suite to replicate findings of a clinical trial for COVID-19. Results indicate approach has potential to be used to assess the efficacy of treatments for emerging diseases.

KEYWORDS
- Resource-intensive OHDSI implementation can exclude smaller healthcare sites, especially in disadvantaged areas. The Edge Tool suite reduces implementation time and costs, making data conversion to OHDSI more accessible.

ABSTRACT
- Exclusion criteria applied to the pilot site records included patients under 18, and pregnant or breastfeeding women. Data quality assessment was also conducted, evaluating missingness, plausibility, and outliers in laboratory values.

Evan Minty, Jack Janetzi, James P. Gilbert, Jung Ho Kim, Jung Ah Lee, Elsie Ross, Nicole Pratt, Gowtham Rao, Seng Chan You

METHODS
1. The Edge Tool Suite enables EMR ETL into the OHDSI-CDM. Propensity score matching at a 1:1 ratio was used to match patients with and without dexamethasone administration on 11 key variables (e.g., patient age).
2. A pilot healthcare site provided over 50,000 encounter COVID-19 patient records from March 2020 to March 2022, with exclusion criteria applied to align with the RECOVERY trial.
3. Logistic regression and survival analyses were performed on the matched data to assess the impact of Dexamethasone on 28-day mortality.

RESULTS
- The analysis found that dexamethasone reduced 28-day mortality in COVID-19 patients receiving oxygen alone or mechanical ventilation, aligning with the RECOVERY trial. However, without oxygen support stratifications, this result was not confirmed.
Postnatal growth deficiency and neurodevelopmental delay phenotypes to study drug safety during pregnancy

**Presented by:** Amir Sarayani

INTRODUCTION

Post-marketing studies using real-world data (RWD) to assess drug safety during pregnancy are necessary because clinical trials rarely include this patient population.

Postnatal growth deficiency and neurodevelopmental delays in infants are two critical outcomes health authorities request to post-marketing drug safety studies during pregnancy.

This study aimed to develop RWD phenotypes for postnatal growth deficiency and neurodevelopmental delay in infants.

RESULTS

- Identified 152,972 preterm births with birth weight less than 2500 grams.
- Identified 44,092 cases of developmental delay with ICD-10 codes.
- Identified 200,586 cases of postnatal growth deficiency with ICD-10 codes.

METHODS

1. Comprehensive literature review to define the clinical concept and identify previously developed phenotypes.
2. OHDSI software tools, i.e., PHQOE, ATLAS, Cohort Diagnosis, and Phenovator, facilitated this phenotype development project.
3. Data sources: Optum’s de-identified ClínicaCarrefour Data Mart Database and MesoDataMarket Commercial Claims and Encounters Database (CCAD).
4. We also created a secondary definition for developmental delay by requiring a second code 01-360 days after cohort entry to improve performance metrics.

A data-driven approach to developing computable phenotypes for postnatal growth deficiency and neurodevelopmental delay showed acceptable performance metrics for the postnatal growth deficiency phenotype while modest performance for the neurodevelopmental delay phenotype.

**Additional Results**

- The ATLAS concept set expression had 29 standard concepts for postnatal growth deficiency (20 exclusions) and 23 standard concepts for developmental delay (14 exclusions) from observation and condition domains, resulting in a total of 206 and 2238 included codes in the concept sets, respectively.
- 156,000 cases in ClínicaCarrefour and 394,000 in CCAD with postnatal growth deficiency and 57,000 and 267,000 cases with developmental delay, respectively.
- Postnatal growth deficiency cohort: 77.0% of subjects had "failure to thrive" (standard concept code: 4379961), and 33.9% had "failure to thrive in neonate" (068717004) as index event in both databases.
- Developmental delay phenotype: about 18-26% of subjects had "delayed milestones" (416233), and 51.1% had "disorder of speech and language development" (428323) as index event in both databases.

The annual incidence rate estimates in Cohort Diagnostics did not show abrupt changes for both phenotypes across the databases over time.

(Amir Sarayani, Jill Hardin, Melanie Jacobson, Rupa Makadia, Joel Swerdel, Kevin Haynes, David Kern)
Harnessing OHDSI’s Framework for a Global Real-World Evidence Master’s Degree Program

**PRESENTED:** Justin Manjourides

**FRIDAY**

**WHY ENROLL?**

A unique program designed from the ground up by active members of the OHDSI Community with a range of academic and industry experience.

A distinct focus on open science, reproducibility, and systematized analytics using best practices.

Learn in an environment that best suits your style, with online and on-ground (Portland, ME) enrollment options.

**MORE WAYS TO BE INVOLVED**

Opportunities to participate as an instructor and/or provide experiential learning opportunities.

Opportunities for your organization to partner with us to create bespoke credit-bearing learning programs that can lead to badges or certificates.

Northeastern University has launched a new MS in RWE degree program with an emphasis on the OMOP CDM and the OHDSI analytic framework.

Ask me about our actively enrolled learners and faculty who are participating at this conference!

Take a picture to learn more and get in touch with us!

**CURRICULUM**

Core Courses:
- Introduction to RWE
- Foundations of Data Models
- Methods for EHR Research I & II
- Standardization of RWD
- Data Model Transformation
- Research Skills and Ethics
- Capstone

Electives:
- Phenotyping
- Cohort Building
- Advanced Characterization
- Advanced Population Level Estimation
- Advanced Patient Level Prediction

**LEARNING OUTCOMES**

1. Describe the value and process of the ethical use of observational health data to answer clinical questions.
2. Illustrate how different forms of observational health data are collected, organized, and standardized to generate accurate, reproducible, and well-calibrated evidence.
3. Use state-of-the-art statistical software and methods to construct and analyze large-scale federated health data from diverse sources while preserving privacy.
4. Construct and take part in a team to conceptualize, analyze, and communicate the results of a study using observational health data to answer a clinical question.
5. Evaluate the strengths and weaknesses of an observational health analysis.

Justin Manjourides, Kristin Kostenko, Christian Reich, Asieh Golozar

@OHDSI
dwww.ohdsi.org

#OHDSISocialShowcase

#JoinTheJourney
Job Opening: Stanford University

Open Postdoctoral position, faculty mentor Brian Bateman

How To Apply

Open Postdoctoral Positions
Finding a Faculty Mentor
Cost of Living
Housing
Fellowships at Stanford
Fellowships outside Stanford

Important Info

Faculty Sponsor (Last, First Name): Bateman, Brian
Other Mentor(s) if Applicable: Stephanie Leonard
Stanford Departments and Centers: Anesthes, Periop & Pain Med
Postdoc Appointment Term: Initial appointment is 1 year with renewal after the first year for an additional 1-2 years by mutual agreement
Appointment Start Date: Flexible start date
Group or Departmental Website:
Where Are We Going?

Any other announcements of upcoming work, events, deadlines, etc?
Three Stages of The Journey

Where Have We Been?
Where Are We Now?
Where Are We Going?
Nov. 13: Collaborator Showcase Honorees

GUSTO OMOP Data Catalogue lays the foundations for developing cross-study OMOP Data Catalogues expanded across APAC and global OHDSI data partners, enabling database level characterizations.
# AMIA Workshop

<table>
<thead>
<tr>
<th>Time</th>
<th>Title</th>
<th>Type</th>
<th>Speaker(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30 – 8:50 am</td>
<td>Why use OHDSI? How can we catch attention?</td>
<td>Presentation</td>
<td>Christian Reich</td>
</tr>
<tr>
<td>8:50 – 9:10 am</td>
<td>Standardization: data structure/model, data content, semantics, cohorts, analysis, reporting</td>
<td>Presentation</td>
<td>Ben Martin</td>
</tr>
<tr>
<td>9:10 – 9:30 am</td>
<td>Value of reusable definitions of disease for research, reusable components for research</td>
<td>Hands-on</td>
<td>Asieh Golozar, Atim Adam</td>
</tr>
<tr>
<td>9:30 – 9:50 am</td>
<td>Collaborative open-science community - Transforming RWE Research</td>
<td>Presentation</td>
<td>Paul Nagy, Ben Martin</td>
</tr>
<tr>
<td>9:50 – 10:10 am</td>
<td>BREAK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:10 – 11:10 am</td>
<td>Stump the Experts (Mui Van Zandt, Paul Nagy, Christian Reich)</td>
<td>Panel</td>
<td>Moderator: Atif Adam</td>
</tr>
<tr>
<td>11:10 – 11:30 am</td>
<td>Build concept sets and cohort building</td>
<td>Demo</td>
<td>Ben Martin</td>
</tr>
<tr>
<td>11:30 – 11:50 am</td>
<td>Power of HADES and Strategus</td>
<td>Presentation</td>
<td>Gowtham Rao</td>
</tr>
<tr>
<td>11:50 – 12:10 pm</td>
<td>ATLAS: Characterization and visualization</td>
<td>Demo</td>
<td>Mui Van Zandt, Asieh Golozar</td>
</tr>
<tr>
<td>12:10 – 1:10 pm</td>
<td>LUNCH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:10 – 1:30 pm</td>
<td>Reproducibility and trust</td>
<td>Presentation</td>
<td>Ross Williams, Atif Adam</td>
</tr>
<tr>
<td>1:30 – 2:30 pm</td>
<td>ATLAS Group Exercise</td>
<td>Hands-On</td>
<td>Ben Martin, Paul Nagy</td>
</tr>
<tr>
<td>2:30 – 2:50 pm</td>
<td>BREAK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2:50 – 3:10 pm</td>
<td>Evidence at Scale</td>
<td>Presentation</td>
<td>Asieh Golozar</td>
</tr>
<tr>
<td>3:10 – 3:30 pm</td>
<td>Closing</td>
<td>Presentation</td>
<td>Christian Reich</td>
</tr>
</tbody>
</table>