Tools for Adoption of OHDSI Data Standards

OHDSI Community Call
Nov. 9, 2021 • 11 am ET
## Remaining 2021 OHDSI Community Calls

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov. 9</td>
<td>Demos: Tools for Adoption of OHDSI Data Standards</td>
</tr>
<tr>
<td>Nov. 16</td>
<td>Open Network Studies</td>
</tr>
<tr>
<td>Nov. 23</td>
<td>History of OHDSI</td>
</tr>
<tr>
<td>Nov. 30</td>
<td>Collaborator Showcase Presentations</td>
</tr>
<tr>
<td>Dec. 7</td>
<td>How Did We Do This Year? Final OKR Review</td>
</tr>
<tr>
<td>Dec. 14</td>
<td>Holiday-Themed Final Meeting Of 2021</td>
</tr>
</tbody>
</table>
## Remaining 2021 OHDSI Community Calls

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov. 9</td>
<td>Demos: Tools for Adoption of OHDSI Data Standards</td>
</tr>
<tr>
<td>Nov. 16</td>
<td>Open Network Studies</td>
</tr>
<tr>
<td>Nov. 23</td>
<td>History of OHDSI</td>
</tr>
<tr>
<td>Nov. 30</td>
<td>Collaborator Showcase Presentations</td>
</tr>
<tr>
<td>Dec. 7</td>
<td>How Did We Do This Year? Final OKR Review</td>
</tr>
<tr>
<td>Dec. 14</td>
<td>Holiday-Themed Final Meeting Of 2021</td>
</tr>
</tbody>
</table>
Three Stages of The Journey

Where Have We Been?
Where Are We Now?
Where Are We Going?
OHDSI Shoutouts!

Congratulations to Patricia Biedermann, Rose Ong, Alexander Davydov, Alexandra Orlova, Philip Solovyev, Hong Sun, Graham Wetherill, Monika Brand and Eva-Maria Didden for the publication of “Standardizing registry data to the OMOP Common Data Model: experience from three pulmonary hypertension databases” in BMC Medical Research Methodology.

Abstract

Background: The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) can be used to transform observational health data to a common format. CDM transformation allows for analysis across disparate databases for the generation of new, real-world evidence, which is especially important in rare disease where data are limited. Pulmonary hypertension (PH) is a progressive, life-threatening disease, with rare subgroups such as pulmonary arterial hypertension (PAH), for which generating real-world evidence is challenging. Our objective is to document the process and outcomes of transforming registry data in PH to the OMOP CDM, and highlight challenges and our potential solutions.

Methods: Three observational studies were transformed from the Clinical Data Interchange Standards Consortium study data tabulation model (SDTM) to OMOP CDM format. OPUS was a prospective, multi-centre registry (2014–2020), and OPHEUS was a retrospective, multi-centre chart review (2013–2017); both enrolled patients newly treated with macitentan in the US. EXPOSE1 is a prospective, multi-centre cohort study (2017–ongoing) of patients newly treated with selipipraz or any PAH-specific therapy in Europe and Canada. OMOP CDM version 5.1.1 with recent OMOP CDM vocabulary was used. Imputation rules were defined and applied for missing data to avoid exclusion of data. Custom target concepts were introduced when existing concepts did not provide sufficient granularity.

Results: Of the 6622 patients in the three registry studies, records were mapped for 6457. Custom target concepts were introduced for PAH subgroups (by combining SNOMED concepts or creating custom concepts) and World Health Organization functional class. Per the OMOP CDM convention, records about the absence of an event, or the lack of information, were not mapped. Excluding these non-event records, 4% (OPUS), 3% (OPHEUS) and 1% (EXPOSE) of records were not mapped.

Conclusions: SDTM data from three registries were transformed to the OMOP CDM with limited exclusion of data and deviation from the SDTM database content. Future researchers can apply our strategy and methods in different disease areas, with tailoring as necessary. Mapped registry data to the OMOP CDM facilitates more efficient collaborations between researchers and establishment of federated data networks, which is an urgent need in rare diseases.
OHDSI Shoutouts!

The Hyve recently presented their **FAIRplus recipe** on how to make published observational research data more FAIR.

The recipe describes the work done on making data on the website of the Covid-19 EHDEN/OHDSI studyathon, held in March 2020, more findable and interoperable.

[www.thehyve.nl/articles/fairify-published-observational-studies-and-databases](http://www.thehyve.nl/articles/fairify-published-observational-studies-and-databases)
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 – Vocabulary Subgroup</td>
</tr>
<tr>
<td>Wednesday</td>
<td>10 am</td>
<td>FHIR and OMOP – Digital Quality Measurements Subgroup (ZOOM)</td>
</tr>
<tr>
<td>Wednesday</td>
<td>2 pm</td>
<td>Natural Language Processing</td>
</tr>
<tr>
<td>Thursday</td>
<td>8 am</td>
<td>Psychiatry</td>
</tr>
<tr>
<td>Thursday</td>
<td>1 pm</td>
<td>OMOP CDM Oncology – CDM/Vocabulary Subgroup</td>
</tr>
<tr>
<td>Friday</td>
<td>10 am</td>
<td>Psychiatry</td>
</tr>
<tr>
<td>Friday</td>
<td>10 am</td>
<td>Electronic Health Record</td>
</tr>
<tr>
<td>Friday</td>
<td>10:30 am</td>
<td>Phenotype Development and Evaluation</td>
</tr>
<tr>
<td>Friday</td>
<td>11 pm</td>
<td>China Chapter</td>
</tr>
<tr>
<td>Monday</td>
<td>10 am</td>
<td>GIS-Geographic Information System General Meeting</td>
</tr>
<tr>
<td>Tuesday</td>
<td>9 am</td>
<td>OMOP CDM Oncology – Genomic Subgroup</td>
</tr>
</tbody>
</table>

[www.ohdsi.org/upcoming-working-group-calls](http://www.ohdsi.org/upcoming-working-group-calls)
Get Access To Different Teams/WGs/Chapters

Welcome to OHDSI!

The Observational Health Data Sciences and Informatics (or OHDSI, pronounced “Odyssey”) program is a multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics. All our solutions are open-source.

OHDSI has established an international network of experts.

Our 2020 OHDSI Global Symposium brought together a global research community for 18 hours of open science, international collaboration, and community fun. The day included research presentations from community members, panels that brought together leaders from major healthcare organizations, as well as network sessions, the annual collaborator.

5. Select the workgroups you want to join (you can refer to the WIKI for work group objectives

- XT1AS
- Clinical Trials
- Common Data Model
- Data Quality Dashboard Development
- Early-stage Researchers
- Education Work Group
- Electronic Health Record (EHR) ETL
- Geographic Information System (GIS)
- HADES Health Analytics Data-to-Evidence Suite
- Health Equity
- Latin America
- Medical Devices
- Natural Language Processing
- OHDSI APAC
- OHDSI APAC Steering Committee
- OHDSI Steering Committee
- Oncology
- Patient-Generated Health Data
- Pharmacovigilance Evidence Investigation

6. Select the chapter(s) you want to join

- Africa
- Australia
- China
- Europe
- Japan
- Korea
- Singapore
- Taiwan

7. Select the studies you want to join

- IHEA-Health Equity Research Assessment
- PROBEHR for Prostate Cancer (study is under ended)
- SCRUA (SARS-CoV-2 Large-scale Longitudinal Analytics

@OHDSI www.ohdsi.org #JoinTheJourney
Get Access To Different Teams/WGs/Chapters

5. Select the workgroups you want to join (you can refer to the WIKI for work group objectives)

- XT,AS
- Clinical Trials
- Common Data Model
- Data Quality Dashboard Development
- Early-stage Researchers
- Education Work Group
- Electronic Health Record (EHR) ETL
- Geographic Information System (GIS)
- HADES Health Analytics Data-to-Evidence Suite
- Health Equity
- Latin America
- Medical Devices
- Natural Language Processing
- OHDSI APAC
- OHDSI APAC Steering Committee
- OHDSI Steering Committee
- Oncology
- Patient-Generated Health Data
- Pharmacovigilance Evidence Investigation

6. Select the chapter(s) you want to join
- Africa
- Australia
- China
- Europe
- Japan
- Korea
- Singapore
- Taiwan

7. Select the studies you want to join
- HEM-Self Health Data Research Assessment
- PIONEER for Prostate Cancer (study re-opened on 1/29/18)
- SCYLLA (Social-Cov-2) Large-scale Longitudinal Analytics

@OHDSI www.ohdsi.org #JoinTheJourney
# 2021 APAC Symposium • Nov. 18

<table>
<thead>
<tr>
<th>Nov. 18 (APAC Time Zone)</th>
<th>Time (Korea time)</th>
<th>Contents</th>
<th>Speaker(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morning</td>
<td>9:00 – 9:25 am</td>
<td>OHDSI State of the Community</td>
<td>George Hripcsak/Patrick Ryan</td>
</tr>
<tr>
<td></td>
<td>9:25 – 9:50 am</td>
<td>OHDSI APAC State of the Community</td>
<td>Mui Van Zandt</td>
</tr>
<tr>
<td></td>
<td>9:50 – 10:00 am</td>
<td>Energy Break</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10:00 – 10:25 am</td>
<td>EHDEN</td>
<td>Peter Rijnbeek</td>
</tr>
<tr>
<td></td>
<td>10:25 – 10:50 am</td>
<td>FHIR and OHDSI Collaboration</td>
<td>Christian Reich</td>
</tr>
<tr>
<td></td>
<td>10:50 – 11:00 am</td>
<td>Energy Break</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11:00 - 12:30 pm</td>
<td>APAC Chapter Visions for 2022</td>
<td>Chapter Leads</td>
</tr>
<tr>
<td>Lunch Break</td>
<td>12:30 – 13:00 pm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afternoon (in GatherTown)</td>
<td>13:00 – 14:00 pm</td>
<td>Workgroup Sessions (Medical Image, FHIR, CDM Tables)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14:00 – 15:00 pm</td>
<td>Collaboration Showcase</td>
<td></td>
</tr>
<tr>
<td></td>
<td>15:00 – 16:00 pm</td>
<td>APAC Study Sessions</td>
<td></td>
</tr>
</tbody>
</table>

[www.ohdsi.org/apac](https://www.ohdsi.org/apac)

@OHDSI www.ohdsi.org #JoinTheJourney
2021 APAC Symposium • Nov. 18

WG – Medical Imaging
Seng Chan You
Assist. Professor
Yonsei University Health System

WG – FHIR Collaboration
Christian Reich
VP IQVIA, OHDSI founder

Adam Chee
Chief of Smart Health Leadership Centre, NUS

WG – CDM v5.4
Clair Blacketer
Assoc. Director Janssen

APAC Study Session
Marc Suchard
Professor, UCLA

www.ohdsi.org/apac
DATA PARTNER CALL

5th Open call for data partners wanting to map their patient data to the OMOP common data model to enhance and accelerate research and healthcare decision making.

- Work with one of 26 EHDEN Certified SMEs
- Up to 100,000 € grant for mapping cycle
- Rapid evaluation & turnaround

October 13th - November 15th

EH DEN.EU
OHDSI Open-Source Governance Workshop

Nov. 29, 2021
9 am – 1 pm
The operating code for how our organization comes together and builds things.

- Software
- Book of OHDSI
- Common data model
- Conferences
- Studyathons
OHDSI Open-Source Governance

“We want our technology stack to be easy to install, use, integrate, and contribute to.”

- Would like to encourage more contributors
- Would like to see more organizations
- Would like to reduce technical debt of support
- Would like to learn how to better recognize those that contribute through code
- We would like to increase the velocity of our development cycle
Roles in an open-source project

- Coders,
- architects,
- reviewers,
- leaders,
- organizers,
- supporters,
- helpers,
- questioners,
- explainers,
- moderators

- Writing code
- Promotion of amazing people and projects
- Leading contributor communities
- Technical projection leadership
- Managing communication
- Updating documentation
- Answering questions from community members

The 4 Pillars of Successful Open-Source Communities (maximilianmichels.com)
Using the Johns Hopkins OSPO as a coordination point, half-day of tutorial/workshop (virtual) and discussion for OHDSI audience

1. The software engineering economics of consumption + contribution, and of project production.
2. The nature of success (a healthy contribution flow) and building on-ramps for users, developers, and contributors.
3. The care and feeding of culture and governance over the long term.
A manual review of vaccine mappings in the OMOP CDM identified several errors which can lead to incorrect results in research.
Covid-19 pandemic impacts on mental health related conditions via multi-database network: a Longitudinal Observational study (CERVELLO)

Hao Luo¹ Carmen Olga Torre²,³ Kenneth Man³ Wallis Lau³ Ian Wong¹,³

¹The University of Hong Kong
²University of Groningen
³UCL School of Pharmacy

Sep 12-13, 2021
Development of a Machine-learning Model to Predict Antibiotic Resistance using Urine Culture and Antibiotic Susceptibility Data
Authors: Chungsoo Kim, Sandy Jeong Rhie, Rae Woong Park
#OHDSISocialShowcase This Week

**ARES: A Research Exploration System**

**Authors:** Frank J DeFalco, Alan Andryc, Anthony Molinaro, Clair Blacketer

**THURSDAY**
Phenotype Development and Evaluation of Heart Failure: A Case Study in using Patient Level Prediction to Improve Phenotype Validity

Authors: Pooja M. Desai, Anna Ostropolets, Lauren R. Richter, Harry Reyes Nieva, Matthew Spotnitz, Victor A. Rodriguez, Tony Y. Sun, Karthik Natarajan

INTRODUCTION
- Establishing robust, high-fidelity phenotypes to isolate incident cases is a challenging task. A common reason is index date misclassification, which occurs at the earliest time frame.
- The proposed OHDSI pipeline for rigorous phenotype development and evaluation comprises the following steps:
  - **Data Preparation**: Data is processed to ensure consistency and quality.
  - **Feature Engineering**: Features are created to capture relevant aspects of the phenotype.
  - **Model Training**: Machine learning models are trained to predict the phenotype.
  - **Validation**: Models are validated using internal and external data sets.
  - **Final Model Selection**: The best model is selected based on performance metrics.
- The proposed pipeline was used to develop the phenotype for heart failure.

CASE STUDY: HEART FAILURE & TZD USE
- Thiazolidinediones (TZDs) are a class of anti-diabetic drugs that can increase the risk of heart failure (HF) exacerbations, even among patients without a known HF diagnosis.
- **Case Study Aim**: Predict the risk of HF-related hospitalization among TZD mono-therapy patients with and without a prior history of HF using both Columbia University Irving Medical Center (CUIMC) and Commercial Claims and Encounters (CCE) data.
- Following literature review, we defined HF as having (1) a condition occurrence of HF diagnostic codes, (2) observation of HF related condition codes, (3) condition occurrence of HF-related codes, or (4) ECG ejection fraction measurement under 50%.
- A total of 30,261 TZD mono-therapy patients hospitalized for HF were identified (CUIMC: N=1,291, CCE: N=29,070), 96.3% of patients (N=29,251) appeared to have no prior HF.

FINDINGS
- Initial ATLAS cohort characterization suggested low validity and appeared to clearly distinguish between TZD patients with and without prior HF.
- **TZD Patients with HF**: Top drug occurrence was glimepiride, hydrochlorothiazide, and losartan.
- **Non-TZD Patients with HF**: Top drug occurrence was ramipril, losartan, and spironolactone.
- **TZD Patients without HF**: Top drug occurrence was no specific drugs. Top condition occurrences were diabetes and hypertension.
- **PLP models** appeared able to predict HF hospitalization for patients with HF (CUIM C: AUC=0.55; CCE: AUC=0.57) and without HF (CUIM C: AUC=0.52; CCE: AUC=0.58).
- However, a review of top PLP features showed drug and condition features specific to HF were predictive of future hospitalization, even among patients without a history of HF.
- **Drug Features**: (Rosiglitazone, N=0.03, CCE: Rosiglitazone, N=0.03, CCE: N=0.03, CCE: N=0.03, CCE: N=0.03, CCE: N=0.03, CCE: N=0.03, CCE: N=0.03).
- **Condition Features**: (Heart Failure, N=0.03, CCE: N=0.03, CCE: N=0.03, CCE: N=0.03, CCE: N=0.03, CCE: N=0.03, CCE: N=0.03, CCE: N=0.03).

LESSONS FOR THE OHDSI COMMUNITY
- Be cautious and consider multiple time windows when establishing a look-back period for phenotype definition if calculating incidence rates.
- Leverage existing OHDSI tools (e.g., Cohort Definition) to understand potential inconsistencies, particularly ones that are temporal in nature.
- Utilize the PLP package as a second round of validation of outcomes and target phenotype definitions by reviewing inconsistencies features that appear most predictive in the models produced.
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. 9: Tools For Adoption Of OHDSI Data Standards

Introduction To CDM 5.4

OHDSI Vocabularies

ETL Inspection Report