Trick or Treat

How to use OHDSI tools to quickly generate insights from your OMOP CDM

OHDSI Community Call
Oct. 26, 2021 • 11 am ET
## Upcoming OHDSI Community Calls

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oct. 26</td>
<td>Trick or Treat: How to use OHDSI tools to quickly generate insights from your OMOP CDM</td>
</tr>
<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>
</tbody>
</table>
# Upcoming OHDSI Community Calls

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oct. 26</td>
<td>Trick or Treat: How to use OHDSI tools to quickly generate insights from your OMOP CDM</td>
</tr>
<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>
</tbody>
</table>
Nov. 2: Future Collaboration Opportunity Breakouts

Methods Research
- Jenna Reps
- Martijn Schuemie

Data Standards
- Clair Blacketer
- Maxim Moinat

Open-Source Development
- Adam Black
- Anthony Sena

Clinical Applications
- Talita Duarte-Salles
- Asieh Golozar

@OHDSI www.ohdsi.org #JoinTheJourney
Three Stages of The Journey

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

Congratulations to the EHDEN Consortium on welcoming 21 new SMEs to support mapping to the OMOP Common Data Model, and perform services in the ecosystem of the EHDEN federated data network.

EHDEN now has a total of 47 SMEs across 19 European nations to assist in real world evidence generation within the community.
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 Measures Subgroup (ZOOM)</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>Electronic Health Record</td>
</tr>
<tr>
<td>Friday</td>
<td>10:30 am</td>
<td>Clinical Trials</td>
</tr>
<tr>
<td>Monday</td>
<td>10 am</td>
<td>GIS-Geographic Information System</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 (OHDSI) 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.

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 www.ohdsi.org/web/wiki/doku.php?id=projects:overview)

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

7. Select the studies you want to join

#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

- KT/LAS
- 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
- HESA-Health Study Research Assessment
- PIONEER for Prostate Cancer (study is under review)
- SCYLIA (SMRI-Cov2) Large-scale Longitudinal Analyses
# 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
Association Rule and Frequent Pattern Mining using the OMOP CDM

**Presented by:** Solomon Ioannou

**Introduction**
To better understand the consequence of data elements and their sequence, association rule analysis and frequent pattern analysis are powerful tools.

An Association Rule analysis answers the question, “Given a cohort of patients, what are the most associated concepts that occur together?”

A Frequent Pattern analysis answers the question, “What are the most common sequences of concepts observed in a cohort of patients?”

Potentially, they are also promising tools to improve other data mining tasks such as patient level prediction.

We introduce here an open-source analytics framework package for performing Association Rule and Frequent Pattern mining using data in this OMOP CDM.

**Methods**
1. The AssociationRuleFMiner package makes use of the open-source SPADE Java library by Philippe Fournier-Viger that implements a large collection of association rule and frequent pattern mining algorithms.
2. Using standard R-packages the user can connect to a database, create the cohorts of interest and extract relevant covariates.
3. Functionalities within the package allow efficient preparation of the input elements and analysis using the algorithms of choice.

**Workflow Description**

1. Create a cohort using one of OHDSI's tool of choice.
2. Extract covariates using the FeatureExtraction package.
   - For Association Rule Mining, extracting the first occurrence of an event (diagnosis, drug prescription) and will suffice to perform the analysis.
   - For Frequent Pattern Mining, the order of events matters, therefore extracting temporal covariates is essential.
3. Choose an algorithm for the relevant analysis and set its parameters.
   - A required parameter to extract highly occurring patterns or frequent patterns is minimum support, which acts as the threshold for the minimum number of patients that should have the concept set in their medical history, e.g., (obesity, diabetes).
   - Algorithms that extract either association rules or frequent patterns require also to specify minimum confidence, which is the threshold for determining how often the left side of the rule occurs together with the right side, i.e., (obesity, diabetes) => (heart failure).
4. Prepare input datasets and run the analysis.
   - The package provides specific functionalities to prepare the input datasets to the necessary format and execute the algorithm.
   - Based on the size of the cohort, an iterative procedure to select the optimal value for minimum support and minimum confidence may be applied.

**Viewing and exploring the results through interactive plots.**

**Results**
1. Depending on the size of the cohort to be analyzed, the number of concepts included, and the values of predefined parameters of minimum support and minimum confidence, a large number of rules can be generated.
2. Currently, results are presented in lists for further processing and use, such as covariates in prediction models.
3. Interactive visualizations are also implemented to explore the results graphically.

**How can this tool be used?**
1. We are exploring the possibilities of using these methods for.
   - Characterization problems.
   - Another research direction is the applied practice under guidelines and the use of these methods for the prediction of clinical practice problems.

**Clinical relevance**
1. Characterizing frequent patterns and associations in health data can help to identify different types of patients who may benefit from a specific treatment.
2. Frequent pattern analysis could also help to generate new hypotheses for the pathogenesis of diseases.

The European Health Data & Evidence Network has received funding from the European Unions Horizon 2020 Research and Innovation program under grant agreement No 824252. The EU receives support from the European Unions Horizon 2020 research and innovation programme and EFPIA.
Extending the OMOP CDM to store the output of NLP pipelines

Mónica Arrúe, Sandra Pulido, Alvaro Abella, Gabriel Maeztu, Alberto Labarga

2021 OHDSI Collaborator Showcase

IOMED
Accelerating Clinical Research
www.iomed.health

TUESDAY
Extending the OMOP CDM to store the output of natural language processing pipelines
Authors: Monica Arrue (presenter), Sandra Pulido, Alvaro Abella, Gabriel Maeztu, Alberto Labarga
Title: CQL Scripting From Atlas Cohort Definition

Authors: Michael Riley, Jon Duke

WEDNESDAY

Automated Translation of Cohort Definitions from Atlas JSON to CQL-FHIR

Authors: Michael Riley, Jon Duke

- ConceptSet With SystemURI Definitions were automatically expanded into ConceptEntities using VSAC Terminology Service
- ExpressionLimitFirstLast used as a global definition applied to primaryEntity
- GroupEntity Defines Grouping of AdditionalCriteria while AdditionalCriteria collects temporal and value based filtering on the entity
- InPopulation Subsumes final definition from Primary, Additional, and Group Entities
- Patient CQL Context used, Population CQL Context feature in development
Best of intent, worst of both worlds: why sequentially combining epidemiological methods does not improve signal detection in vaccine surveillance

Authors: Faaizah Arshad, Lana YH Lai, George Hripcsak, Daniel Prieto-Alhambra, Martijn J. Schuemie, Marc A. Suchard

THURSDAY
Predicting risk of recurrence after surgery for colorectal cancer

Preoperative variables can decently predict the risk of recurrence after surgery for colorectal cancer.

Increased risk can be taken into consideration for clinical decision-making to identify patients that might benefit from:
- preoperative treatment
- smaller resections
- increased postoperative monitoring

Authors: Mikail Gögenur, Viviane Lin, Adamantia Tsouchnika, Eldar Allakherzdieev, Andreas Weinberger Rosen, Karoline Bendix Bräuner, Julie Sparholt Walbech, Ismail Gögenur
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?
On Tuesday, Oct. 26 (11 am ET), Patrick Ryan will lead a Halloween-themed interactive demonstration of how you can use the OHDSI tools to quickly generate insights from your OMOP CDM.

We hope you'll learn a TRICK or two, and that it will be a TREAT.