European Symposium Review

OHDSI Community Call
June 28, 2022 • 11 am ET

— Catherine Cohet, European Medicines Agency
— Filip Maljkovic, Clinerion Serbia Rep
— Daniel Morales, Dundee University and HIC
— Patrick Ryan, Janssen
<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 5</td>
<td>NO MEETING</td>
</tr>
<tr>
<td>July 12</td>
<td>New Adopter Introductions and Q&amp;A</td>
</tr>
<tr>
<td>July 19</td>
<td>Workgroup Updates</td>
</tr>
<tr>
<td>July 26</td>
<td>CDM Update Process</td>
</tr>
</tbody>
</table>
## Upcoming OHDSI Community Calls

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 5</td>
<td>NO MEETING</td>
</tr>
<tr>
<td>July 12</td>
<td>New Adopter Introductions and Q&amp;A</td>
</tr>
<tr>
<td>July 19</td>
<td>Workgroup Updates</td>
</tr>
<tr>
<td>July 26</td>
<td>CDM Update Process</td>
</tr>
</tbody>
</table>
July 12: New Adopters & Community Members

Our July 12 Community Call will be focused on new adopters of the OMOP CDM or new members of the OHDSI community.

We are welcoming people to introduce themselves, share why they have joined the community and what impact they hope to make, and also ask a question to the broader community (if you wish). If you would like to take part in this event, please fill out this form to help us plan the session: https://bit.ly/3A7JNkV

Form in chat and on community calls page
Three Stages of The Journey

Where Have We Been?
Where Are We Now?
Where Are We Going?
Thank you to everybody who submitted brief reports to join our #OHDSI2022 Collaborator Showcase. We had a record amount (more than 130!) of submissions for poster presentations, software demos and oral presentations for the 2022 OHDSI Symposium, which will be held Oct. 14-16 in Bethesda, Md.

The scientific committee meets this week to begin the process of reviewing all submissions, and selected presenters will be notified by August 1.
OHDSI Shoutouts!

Any shoutouts from the community? Please share and help promote and celebrate OHDSI work!

Have a study published? Please send to sachson@ohdsi.org so we can share during this call and on our social channels. Let’s work together to promote the collaborative work happening in OHDSI!
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>Tuesday</td>
<td>3 pm</td>
<td>OMOP CDM Ontology Outreach/Research Subgroup</td>
</tr>
<tr>
<td>Wednesday</td>
<td>11 am</td>
<td>Open-Source Community</td>
</tr>
<tr>
<td>Wednesday</td>
<td>12 pm</td>
<td>FHIR and OMOP Terminologies Subgroup</td>
</tr>
<tr>
<td>Wednesday</td>
<td>7 pm</td>
<td>Medical Imaging</td>
</tr>
<tr>
<td>Thursday</td>
<td>10 am</td>
<td>Data Quality Dashboard</td>
</tr>
<tr>
<td>Thursday</td>
<td>12 pm</td>
<td>FHIR and OMOP Oncology Subgroup</td>
</tr>
<tr>
<td>Friday</td>
<td>9 am</td>
<td>GIS – Geographic Information Systems</td>
</tr>
<tr>
<td>Tuesday</td>
<td>10 am</td>
<td>Common Data Model</td>
</tr>
</tbody>
</table>

[www.ohdsi.org/upcoming-working-group-calls](http://www.ohdsi.org/upcoming-working-group-calls)
The EHDEN Portal, which provides free access to the research community, was launched at the 2022 OHDSI European Symposium. The Portal includes a Data Partner Catalogue (140 partners, >500M anonymous patient records) and Feasibility Dashboards that support data discoverability (findable under Findable, Accessible, Interoperable and Reusable (FAIR) principles).
Hello friends with EHR data,

One of the Healthcare Systems group’s objectives this year is “To provide support for transforming source EHR data to the CDM”. Currently, we provide support through answering questions on the forums and during our regularly scheduled work group meetings. Another product we would like to provide to the community is a central repository of different OMOP sites, their underlying EHR system, and attributes. This will allow new OHDSI collaborators to find and reach out to sites with similar infrastructure, EHR systems, and/or research goals. Participating in this survey does NOT commit you to being a mentor, providing your ETL script, or even answering your email. However, we hope you embrace the spirit of our open source community and contribute to the cause. We all learn as we OMOP our data. I’ve been very active in the OHDSI community and digging deep into EHR data for 8 years, and I still learn something new every day. But I think all persons in any field of science continue to learn because science is continually evolving. Here’s the link to the google form.
Professor Peter Rijnbeek announced an opening for an epidemiologist to work with his team at Erasmus MC.

This position will be responsible for all aspects of observational research including protocol writing, input in the statistical analysis plan, study execution, interpretation of results and report/manuscript writing.

The application deadline is July 8, 2022.
Odysseus Data Services (Odysseus) has an exciting opening for an Epidemiologist. This role will be responsible for supporting the development, maintaining, and troubleshooting of the cutting-edge distributed solutions in the Real-World Evidence (RWE) area, utilized by the researchers in Pharmaceutical, Healthcare and Payer industries. Odysseus is looking for a self-driven individual who can hit the ground running, quick learner and wants to be a part of our dynamic global team.

**Responsibilities**

- Lead and contribute to the design of observational database analysis, including authoring protocol, reviewing and providing relevant epidemiological and project-specific comments to statistical analysis plans and analysis output
- Participate in the design and development of standardized analytic tools to generate reliable and reproducible evidence in a network of observational data
- Contribute to the execution of observational database analyses using standardized analytical tools and writing statistical packages
- Contribute to the dissemination of scientific information through technical reports and publications in peer-reviewed literature
- Work closely with healthcare and pharmaceutical customers to identify their needs
- Contribute to the development of complex phenotypes using advanced analytic approaches (i.e. machine learning, incorporating unstructured data sources using NLP, etc.)

**Qualifications**

- Graduate degree (MS, PhD, MD, etc) in epidemiology, biostatistics, pharmacy, public health or related clinical discipline plus two years' experience in observational research. PhD preferred
- Experience in designing and conducting healthcare studies and in development and applications of advanced analytics solutions
- Strong epidemiology and biostatistics background
- Experience using OHDSI tools and analytical methods is a big plus

Odysseus Data Services (Odysseus) has an exciting opening for a Healthcare/ Clinical Data Scientist. This role will be responsible for supporting the development, maintaining, and troubleshooting of the cutting-edge distributed solutions in the Real-World Evidence (RWE) area, utilized by the researchers in Pharmaceutical, Healthcare and Payer industries. Odysseus is looking for a self-driven individual who can hit the ground running, quick learner and wants to be a part of our dynamic global team.

**Responsibilities**

- Lead and contribute to the design, development and documentation of standardized analytic tools that will be executed against a network of observational data
- Lead the execution of observational database analyses using standardized analytical tools and writing statistical packages
- Provide technical support for the data and analysis infrastructure and scientific support
- Contribute to writing of protocols and statistical analysis plans, methods development, conduct of simulation studies and statistical/mathematical modeling studies
- Lead and contribute to the development of complex phenotypes using advanced analytic approaches (i.e. machine learning, incorporating unstructured data sources using NLP, etc.)
- Contribute to the dissemination of scientific information through technical reports and publications in peer-reviewed literature.
- Lead and contribute to the development of novel analytic tools and techniques to leverage the EHR data for rapid, reliable and reproducible evidence generation
Registration is OPEN for #OHDSI2022!

The 2022 OHDSI Symposium will be held Oct. 14-16 at the Bethesda North Marriott Hotel & Conference Center.

www.ohdsi.org/ohdsi2022symposium
An Introductory Journey From Data To Evidence
OHDSI2022 Tutorial • Saturday, Oct. 15 • Bethesda, Md.

The OHDSI Journey: Where Are We Going?
Patrick Ryan

OMOP Common Data Model and Vocabulary
Clair Blacketer

ETL – A Source Database Into OMOP CDM
Melanie Philofsky

Creating Cohort Definitions
Asieh Golozar

Phenotype Evaluations
Gowtham Rao

Characterization
Kristin Kostka

Estimation
Martijn Schuemie

Prediction
Jenna Reps

The OHDSI Journey: Where Do We Go From Here?
George Hripcsak
# Workgroup Activities

**Saturday, Oct. 15, and Sunday, Oct. 16**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>800</td>
<td>900</td>
<td>Tutorial</td>
<td>Lunch</td>
<td>800</td>
<td>900</td>
<td>All-Hands Workgroup Meeting</td>
</tr>
<tr>
<td>900</td>
<td>1000</td>
<td>HADES Hack-a-thon: Part 1</td>
<td>Lunch</td>
<td>900</td>
<td>1000</td>
<td>Phenotype Evaluation</td>
</tr>
<tr>
<td>1000</td>
<td>1100</td>
<td>Methods Research (PLE/PLP)</td>
<td>Oncology WG (continued)</td>
<td>1000</td>
<td>1100</td>
<td>HADES Hack-a-thon: Part 2</td>
</tr>
<tr>
<td>1100</td>
<td>1200</td>
<td></td>
<td>Oncology WG</td>
<td>1100</td>
<td>1200</td>
<td>Education</td>
</tr>
<tr>
<td>1200</td>
<td>1300</td>
<td></td>
<td>Natural Language Processing</td>
<td>1200</td>
<td>1300</td>
<td>Health Equity</td>
</tr>
<tr>
<td>1300</td>
<td>1400</td>
<td></td>
<td>FHIR-OMOP: Terminologies Subgroup, Part 1</td>
<td>1300</td>
<td>1400</td>
<td>CDM and Data Quality</td>
</tr>
<tr>
<td>1400</td>
<td>1500</td>
<td></td>
<td>FHIR-OMOP: Increasing the Value of Data Through a Rich Set of Attributes</td>
<td>1400</td>
<td>1500</td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td>1600</td>
<td></td>
<td>FHIR-OMOP: Data Model Harmonization Subgroup</td>
<td>1500</td>
<td>1600</td>
<td></td>
</tr>
<tr>
<td>1600</td>
<td>1700</td>
<td></td>
<td>FHIR-OMOP: Oncology Subgroup</td>
<td>1600</td>
<td>1700</td>
<td></td>
</tr>
</tbody>
</table>
Why predicting risk can’t identify ‘risk factors’: empirical assessment of model stability in machine learning across observational health databases

**Lead:** Aniek Markus

---

**MONDAY**

---

Be careful interpreting prediction models as the identified ‘risk factors’ appear to depend on study design choices.

---

Why predicting risk can’t identify ‘risk factors’: empirical assessment of model stability in machine learning across observational health databases

**Lead:** Aniek Markus
OMOP Genomic mapping capacities in conversion of comprehensive genomic profiling results

Lead: Maria Rogozhkina

TUESDAY
Perseus Design and run your own ETL to CDM

Lead: Anton Ivanov
Using geospatial approaches and machine learning for asthma and COPD outcomes: a systematic review

**Title:** Using geospatial approaches and machine learning for asthma and COPD outcomes: a systematic review

**Lead:** Daniel Jeannetot

**THURSDAY**

**Asthma/COPD research has a lot of potential to benefit from machine learning algorithms and geospatial approaches, especially if combined with observational data.**

Most papers approached spatial information in two steps: use of geospatial models to estimate an environmental factor exposure (e.g. specific air pollutants values at specific location) to then integrate values in non-spatial regression models (e.g. linear, multivariate, etc.), removing specific geospatial and geographical processes information.

OHDSI provides a coherent and readily available infrastructure to help Asthma/COPD research leverage observational data, machine learning, and geospatial approaches for very large-scale analyses.

**Inclusion criteria**
- High-quality predictions
- High geospatial/geostatistical approaches
- Exposome variables include environmental (air pollution) or social-economic data
- Main outcome is COPD and/or Asthma-related
- Population should be 16 years old or above

**Search term categories**
1. Asthma and/or COPD and
2. Prediction models (ML, Modelling, Machine Learning, etc.)
3. Socio (geostatistical, prep)
4. ADULT (most children, etc.)

**Key points**
- Population varied greatly in age groups and sample size (from 300 to 1,000,000)
- Scale poorly varied but generally local
- Different machine learning algorithms
- Most geospatial approaches were 2D
- 15 papers used specific geostatistical tools
- Incorporation geospatial and machine learning tools

---

**INTRO:** Asthma & COPD are major contributors to morbidity and mortality worldwide. OHDSI COM databases provide a unique opportunity to enrich clinical health records with geospatial data and machine learning approaches to improve patient-level predictions. This systematic review shows that this is still an unexplored approach with large potential for exploration.

**METHODS**
1. Systematic review following PRISMA guideline
2. 4 databases searched
3. 2 reviewers involved in full text review
4. 12 specific characteristics for data extraction including type of models (ML, socio, HU), spatial scale, spatial approach.

**REMARKS**
- 857 papers screened.
- 523 papers fully reviewed.
- 12 Articles making use of geospatial approach and modeling to measure and predict asthma-related outcomes.
A pilot study to evaluate the feasibility of using OHDSI analytical tools for supporting safety surveillance

Lead: Ceyda Pekmez
This is ongoing research
We initiated the ‘Trials Replication through Observational study by Yonsei (TROY)’ project to generate large population-level evidence for 15 pivotal RCTs in the real world:
Type 2 diabetes mellitus, atrial arrhythmia, acute coronary syndrome, and rheumatoid arthritis

TROY: Trials Replication through Observational study by Yonsei

Lead: Jaehyeong Cho
The EHDEN Platform Roadmap

**Lead:** Michel Van Speybroeck

**TUESDAY**

#OHDSISocialShowcase This Week
A standard ETL process from REDCap to OMOP

**Lead:** Francesco Pozzoni

**WEDNESDAY**

**How to build the mapping file**
- Segment the list of EDCap objects
- Design custom functions to build a mapping file
- Group each EDCap object into an OMOP concept
- Use a mapping file to load data into the OMOP database

**AMM0 BAR**
- Load the employees into the OMOP database
- Load the products into the OMOP database
- Load the orders into the OMOP database

**RESULTS**
- Feature specific tests conducted for the ETL process
- Each phase of the ETL process has been validated

**METHODS**
1. REDCap serves as an ETL tool to extract data from REDCap to OMOP.
2. The ETL process is designed to transform data from REDCap to OMOP.
3. The ETL process is executed using a specific mapping file.
4. The ETL process is validated using specific tests.
Learning robust models from limited external statistics

Lead: Tal El Hay

Augmenting internal data with population-level statistics from external sources could improve model robustness to data-shift

External performance (AUC) of a model trained on internal data degrades faster than for models trained on pooled data or using external statistics and reweighting

Estimating Model Performance on External Samples from Their Limited Statistical Characteristics, Conference on Health, Informatics, and Learning (CHIL) 2022
FeederNet (Federated E-Health Big Data for Evidence Renovation Network) platform in Korea

Lead: Chungsoo Kim

CDM data-network of 53 hospitals and 72M patients’ medical data was established in South Korea. Currently, joint research using it have been actively conducted; 97 papers since 2017 and 55 papers in only 2021 were published by Korean first authors.
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?
June 28: The European Symposium

Presenter:
Nigel Hughes • Director, Observational Health Data Analytics at Janssen Research & Development