Scaling up reliable evidence across Europe

June 1 - 3 2024
Rotterdam
Welcome to the European OHDSI Journey

Prof. dr. ir. Peter R. Rijnbeek
Chair Department of Medical Informatics
Erasmus MC, The Netherlands
Thank you for your support!
Open Science Community
Driving Scalability of Reliable Evidence Generation
Mission and Vision

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.
Why OHDSI Europe?

• Improving interoperability of data in Europe is very much needed and requires expertise in Europe.

• The number of projects in Europe that are building on OMOP is strongly increasing, even within institutions.

• There are >200 data sources in Europe mapped to OMOP and we just got started.
Objectives of OHDSI Europe

• Enable the generation of reliable evidence from European health data: promote the adoption of the OMOP-CDM and analytics.
• Focus on European Challenges and Opportunities.
• Community building
  – Point of contact for all stakeholders
  – Organization of European OHDSI Symposia
  – Training of stakeholders
  – Stimulate national and international collaborations in Europe -> National Nodes.
First Annual OHDSI Symposium, March 23th 2018

- 200 participants
- 24 countries
- 40 posters
- 5 software demos
- 2 full day tutorials
Second Annual OHDSI Symposium, March 29th 2019

• 250 participants
• 27 countries
• 35 posters
• 8 software demos
• 5 full day tutorials
Third Annual OHDSI Symposium, June 24th 2022

- 350 participants
- 80 posters
- 4 software demos
- 2 days with tutorials and workshops
4th Annual OHDSI Symposium, July 3rd 2023

- 350 participants
- 100 posters
- 4 software demos
- 2 days with tutorials and workshops
Meeting Goals Fifth OHDSI Symposium

- Educate and train the community through workshop (Saturday) and multiple Workgroup Meetings (Sunday) (200 participants)

- Facilitate meetings for national nodes and its leadership meeting. Data Partner meetings.

- 360 Participants, 130 submissions for collaborator showcase, 13 software demos
Thanks to all faculty!!
Breakdown of Participants: 37 Countries

2023: 28 countries
Breakdown of Participants: Stakeholders

- Academia
- Technology
- Health System
- Government
- Pharmaceutical

2023: same distribution
# Agenda

<table>
<thead>
<tr>
<th>Time</th>
<th>Title</th>
<th>Location</th>
</tr>
</thead>
</table>
| 9:10 – 9:40| **Journey of OHDSI: Where have we been and where can we go together?**  
Speaker: Patrick Ryan, PhD, Janssen Research and Development, Department of Biomedical Informatics, Columbia University Medical Center | Theatre                   |
| 9:40 – 11:00| **Selection of European Initiatives Using the OMOP CDM**  
Moderator: Renske Los, PhD, Assistant Professor of Medical Informatics, Department of Medical Informatics, Erasmus MC  
Multiple presentations of European Projects and Initiatives | Theatre                   |
| 11:00 – 11:30| **Coffee Break**                                          | Queen’s Lounge            |
| 11:30 – 12:45| **Collaborator Showcase: Rapid fire presentations**  
Moderator: Katia Verhamme, MD, Associate Professor of Use and Analysis of Observational Data, Department of Medical Informatics, Erasmus MC, Rotterdam. | Theatre                   |
| 12:45 – 13:45| **Lunch**                                                  | La Fontaine & Odyssee Room |
## Agenda (2)

<table>
<thead>
<tr>
<th>Time</th>
<th>Title</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>13:00 – 16:00</td>
<td><strong>OHDSI Collaborator Showcase</strong>&lt;br&gt;Poster presentations and open-source software demonstrations from OHDSI collaborators:&lt;br&gt;- Observational data standards and management&lt;br&gt;- Open-source analytics development&lt;br&gt;- Clinical applications&lt;br&gt;- Methodological research&lt;br&gt;- National nodes</td>
<td>La Fontaine &amp; Odyssee Room&lt;br&gt;&lt;br&gt;&lt;b&gt;Early Investigators Mentor Meetings (14:00 – 15:00)&lt;/b&gt;&lt;br&gt;Lead: Ross Williams, Assistant Professor, Department of Medical Informatics, Erasmus MC Rotterdam</td>
</tr>
<tr>
<td>16:00 – 17:10</td>
<td><strong>Large Scale Evidence Generation in EHDEN and DARWIN EU®</strong>&lt;br&gt;Moderators: Prof. Dani Prieto Alhambra and Katia Verhamme, Associate Professor, Department of Medical Informatics, Erasmus MC</td>
<td>Theatre</td>
</tr>
<tr>
<td>17:10 – 17:30</td>
<td><strong>What evidence are we going to showcase at OHDSI Europe in 2025?</strong>&lt;br&gt;Moderator: Patrick Ryan, Johnson &amp; Johnson, Columbia University</td>
<td>Theatre</td>
</tr>
<tr>
<td>17:30 – 18:00</td>
<td><strong>Closure</strong></td>
<td>Theatre</td>
</tr>
<tr>
<td>18:00 – 19:30</td>
<td><strong>Networking Reception</strong></td>
<td>Queen’s Lounge</td>
</tr>
</tbody>
</table>
Journey of OHDSI: Where have we been and where can we go together?

Patrick Ryan PhD
Johnson & Johnson
Columbia University Irving Medical Center
An objective evaluation on how far we’ve made it on our journey....
OHDSI’s mission

To improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care
To improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care.
Standby for the next presentation on progress on OHDSI Europe National Nodes

Regional Chapters and National Nodes

An OHDSI regional chapter represents a group of OHDSI collaborators located in a geographic area who wish to hold local networking events and meetings to address problems specific to their geographic location. The OHDSI Europe Chapter, in collaboration with the EHDEN project, recently created National Nodes to facilitate national and international collaborations.

An OHDSI Europe National Node is a collection of research institutes within a member country. The Node builds on the strengths of the stakeholders and scientific communities of that country. Each Node has a lead institute that oversees the work of that Node and assigns a lead and co-lead.

OHDSI in Africa and Partnerships with European Institutions

Check out today during the Showcase!

Other upcoming OHDSI Activities:
- OHDSI India 5Oct2024
- OHDSI Global 22-24Oct2024 (USA)
- OHDSI APAC 6-9Dec2024 (Singapore)
OHDSI Workgroups

OHDSI has a central mission to improve health globally, but there are countless areas where our community can be of service. Work around data, methods, open-source tools, and clinical applications are all pieces of the puzzle, and within OHDSI, there are opportunities to work in any or many of these areas.

Our workgroups, led by the extraordinary leaders shown on these pages, present opportunities for all community members to find a home for their talents and passions, and make meaningful contributions. We are always looking for new collaborators. See an area where you want to contribute? Please Join The Journey!

www.ohdsi.org/workgroups
OMOP Common Data Model adoption

OMOP CDM Users By The Numbers

- 534 data sources
- 49 countries
- 956 million unique patient records (12% of world’s population)
OHDSI Evidence Network

OHDSI is proud to have a global community dedicated to generating real-world evidence and which recognizes the opportunity to collaborate together as part of a distributed network based on validated and standardized data and analyses.

The OHDSI Evidence Network consists of organizations equipped with access to one or more databases standardized to the OMOP CDM who express a keen interest in participating in OHDSI network studies. Collaboratively, OHDSI Evidence Network partners share aggregate summary statistics about their databases, which are used to support Database Diagnostics, helping identify databases within the network that are fit-for-use for particular research questions. Additionally, partners have the opportunity to opt in and contribute to network studies proposed by the OHDSI community.

The recent SOS challenge serves as a compelling demonstration of the OHDSI Evidence Network’s current capabilities and its promising future potential. We wholeheartedly encourage all organizations that are adopting the OMOP CDM and aspire to apply standardized analyses for the reliable generation of real-world evidence to become part of the OHDSI Evidence Network.

A message from Common Data Model workgroup lead Clair Blacketer...

During the first community call of 2023, Patrick Ryan unveiled the strategic priorities for the OHDSI Community for the year. Among these, a key focus is on enhancing the transparency and maturity of the OHDSI network.

To address this objective, we began by considering how network studies are currently conducted, recognizing the challenges and complexities faced by collaborating organizations when contributing to the body of evidence. This investigation led to the creation of Database Diagnostics, a tool designed to answer a critical question: when booking a specific research inquiry, which data sources within the OHDSI Evidence Network are the most relevant and suitable for generating robust evidence?

This innovative approach leverages aggregated summary statistics from each data source, obtained through the open-source tool dbProfile. It evaluates data fitness-for-use across various dimensions, including patient demographics, domain coverage, longitudinal data availability, and the capture of target, comparator, and outcome variables. The overarching vision was to establish these database profiles as the foundation to enable the OHDSI Evidence Network.

On March 28, 2023, the OHDSI Global Community initiated the Save Our Statsyns (SOS) Challenge, a groundbreaking opportunity for collaborative research involving simultaneous participation in four different network studies. What made it truly remarkable was that any organization interested in joining the OHDSI Evidence Network could contribute to these studies by sharing their database profiles for the data sources they had access to. These profiles were centrally aggregated at the OHDSI Central Coordinating Center, enabling us to empirically determine which of the four study questions each data source was best suited to address. This inaugural OHDSI Evidence Network endeavor encompassed 36 diverse data sources from 16 different organizations. Not only did this foster rapid evidence generation and collaboration during the SOS Challenge, but it also positioned us for future collaborations on additional network studies as part of the OHDSI Evidence Network.

If you are interested in becoming a part of the OHDSI Evidence Network and contributing to advancing evidence-based healthcare, please use the provided QR code to complete a brief form about your organization and your data source. A member of the OHDSI Network Data Quality Working Group will reach out to you to explore this exciting opportunity further!
187 Data Partners from 29 different countries = >33% of OMOP CDM adopters and >59% of countries
## Network Dashboard

### Country Filter
- **Type or Select [Country]**

### Database Type Filter
- **Type or Select [Type]**

### Data Source Filter
- **Type or Select [Data Source]**

### Overview
- **Country:** 30
- **Data Sources:** 188
- **Patients:** 357M

### Datasource Types
- **Hospital**
- **Registry**
- **Other**

---

**EHDEN portal**
Common Data Model (CDM) mapping

CDM mapping
Has the data source been converted (ETL-ed) to a common data model?

Yes

CDM Mappings

<table>
<thead>
<tr>
<th>CDM name</th>
<th>OMOP</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDM website</td>
<td><a href="https://www.ohdsi.org/Data-standardization/">https://www.ohdsi.org/Data-standardization/</a></td>
</tr>
<tr>
<td>Data source ETL frequency</td>
<td>6,00 months</td>
</tr>
<tr>
<td>Data source ETL status</td>
<td>Completed</td>
</tr>
</tbody>
</table>
EHDEN is advancing the science of data networks

Check out today during the Showcase!

CDMOnboarding R package for data quality assessment

Lessons learned from EHDEN Data Partner Reviews: Improving ETL Processes and Data Quality in OMOP CDM Conversions

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 Kalfelz, MD\(^1,4\), Michel van Speybroeck, MSc\(^3\), Daniel Prieto-Alhambra, PhD\(^1,2,6\), Martijn J. Schuemie, PhD\(^1,3,6\), Peter R. Rijnbeek, PhD\(^1,2\)

\(^1\)OHDSI Collaborators, Observational Health Data Sciences and Informatics (OHDSI), New York, NY, United States; \(^2\)Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, the Netherlands; \(^3\)Janssen Pharmaceutical Research and Development LLC, Raritan, NJ 08869, United States; \(^4\)Oyseaus Data Services, Prague, Czech Republic; \(^5\)Centre for Statistics in Medicine, NIGMAMS, University of Oxford, Oxford, United Kingdom; \(^6\)Department of Biostatistics, University of California, Los Angeles, CA 90095, United States.

*Corresponding author: Erica A. Voss, MPH, Janssen Research & Development – Epidemiology, 920 US Highway 202, Raritan, NJ 08869 (evoss3@lts.jnj.com)

Abstract

Objective: Health data standardized to a common data model (CDM) simplifies and facilitates research. This study examines the factors that make standardizing 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 (EHDEN) 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%, 98%, 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 EHDEN proves appropriate to support the successful standardization of observational data across Europe. Over the 2 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.

Conclusion: 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.

Keywords: OMOP common data model; observational data; data standardization.
OHDSI Vocabularies Improvement Initiative

Landscape assessment

FINDINGS
- 87% of the community feels confident about Vocabularies’ integrity
- Most commonly used vocabularies: SNOMED, ICD 9/10 (US and international versions), MedDRA, LOINC, ATC, RxNorm/RxE, ICD10PCS, ICD/ProCPT4, LOINC, CVX, HPCs, UCUM, NDC, NACCR, Cancer Modifier
- Most update data annually or semi-annually

NEEDS
- Transparent release schedule
- Vocabulary changes, versioning
- Transparent QA/QC
- Better coverage and hierarchies
- More documentation and educational materials

Vocabulary committee

Vocabulary team

Release schedule and roadmap

Community contributions

Community contribution pipeline

Issue on GitLab

Template

Commit

Merge data

Checklist (QA)

Submission

Review

Release

Quality framework & documentation
Research and Applications

OHDSI Standardized Vocabularies—a large-scale centralized reference ontology for international data harmonization

Christian Reich MD1,2,3,*, Anna Ostropolets, PhD1,4,5, Patrick Ryan, PhD1,4,6, Peter Rijnbeek, PhD1,3, Martijn Schuemie, PhD1,6, Alexander Davydov, MD1,6, Dmitry Dymshyts, MD1,6, George Hripcsak, MD1,6

1Coordinating Center, Observational Health Data Sciences and Informatics, New York City NY 10032, United States; 2OHDSI Center at the Roux Institute, Northeastern University, Portland ME 04101, United States; 3Department of Medical Informatics, Erasmus University Medical Center, 3015 GD Rotterdam, The Netherlands; 4Department of Biomedical Informatics, Columbia University Medical Center, New York City NY 10032, United States; 5Odysseus Data Services, Cambridge MA 02142, United States; 6Observational Health Data Analytics, Janssen Research & Development, Titusville NJ 08560, United States

*Corresponding author: Christian Reich, MD, OHDSI Center at the Roux Institute, Northeastern University, 100 Fore St, Portland ME 04101 (reich@ohdsi.org)

Abstract

Importance: The Observational Health Data Sciences and Informatics (OHDSI) is the largest distributed data network in the world encompassing more than 331 data sources with 2.1 billion patient records across 34 countries. It enables large-scale observational research through standardizing the data into a common data model (CDM) (Observational Medical Outcomes Partnership (OMOP) CDM) and requires a comprehensive, efficient, and reliable ontology system to support data harmonization.

Materials and methods: We created the OHDSI Standardized Vocabularies—a common reference ontology mandatory to all data sites in the network. It comprises imported and de novo generated ontologies containing concepts and relationships between them, and the praxis of converting the source data to the CDM is based on these. It enables harmonization through assigned domains according to clinical categories, comprehensive coverage of entities within each domain, support for commonly used international coding schemes, and standardization of semantically equivalent concepts.

Results: The OHDSI Standardized Vocabularies comprise over 10 million concepts from 136 vocabularies. They are used by hundreds of groups and several large data networks. More than 6000 users have performed 500,000 downloads of the system. This open-source resource has proven to address an impediment of large-scale observational research—the dependence on the context of source data representation. With that, it has enabled efficient phenotyping, covariate construction, patient-level prediction, population-level estimation, and standard reporting.

Discussion and conclusion: OHDSI has made available a comprehensive, open vocabulary system that is unmatched in its ability to support global observational research. We encourage researchers to exploit it and contribute their use cases to this dynamic resource.

Key words: OHDSI; controlled vocabulary; common data model; observational data.
OHDSI Standardized Vocabularies roadmap

Roadmap 2023 Q1 - 2024 Q2:
HADES

HADES is a set of open source R packages for large scale analytics, including population characterization, population-level causal effect estimation, and patient-level prediction.

The packages offer R functions that together can be used to perform an observational study through the full journey from data to evidence, including data manipulation, statistical modeling, and results generation with supporting statistics, tables and figures.

Each package includes functions for specifying and subsequently executing multiple analyses efficiently. HADES supports best practices for use of observational data as learned from previous and ongoing research, such as transparency, reproducibility, as well as measuring the operating characteristics of methods in a particular context and subsequent empirical calibration of estimates produced by the methods.

Learn more about the individual HADES packages in this section.

The open-source tools that empower OHDSI research are not only available to the community, but they are DEVELOPED by the community. We thank the many developers and maintainers who empower our research initiatives around the world!

The eight HADES packages shown above have been released on CRAN and have been downloaded more than 500,000 times.
Shout out to the DARWIN-EU Development team!
Open-source development accessible on R CRAN
OHDSI scholarship

Publications & Cumulative Citations

Summary

609
PubMed Manuscripts

3613
PubMed Authors

<table>
<thead>
<tr>
<th>Year</th>
<th>Publications</th>
<th>Cumulative Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>10</td>
<td>~1,000</td>
</tr>
<tr>
<td>2011</td>
<td>20</td>
<td>~2,000</td>
</tr>
<tr>
<td>2012</td>
<td>30</td>
<td>~3,000</td>
</tr>
<tr>
<td>2013</td>
<td>40</td>
<td>~4,000</td>
</tr>
<tr>
<td>2014</td>
<td>50</td>
<td>~5,000</td>
</tr>
<tr>
<td>2015</td>
<td>60</td>
<td>~6,000</td>
</tr>
<tr>
<td>2016</td>
<td>70</td>
<td>~7,000</td>
</tr>
<tr>
<td>2017</td>
<td>80</td>
<td>~8,000</td>
</tr>
<tr>
<td>2018</td>
<td>90</td>
<td>~9,000</td>
</tr>
<tr>
<td>2019</td>
<td>100</td>
<td>~10,000</td>
</tr>
<tr>
<td>2020</td>
<td>110</td>
<td>~11,000</td>
</tr>
<tr>
<td>2021</td>
<td>120</td>
<td>~12,000</td>
</tr>
<tr>
<td>2022</td>
<td>130</td>
<td>~13,000</td>
</tr>
<tr>
<td>2023</td>
<td>140</td>
<td>~14,000</td>
</tr>
</tbody>
</table>
OHDSI collaborations in scholarship
OHDSI Europe is growing its impact in community scholarship
OHDSI Europe leading in data standardization
OHDSI Europe leading in open-source development

Research and Applications

Markov modeling for cost-effectiveness using federated health data network

Markus Haug 1, MSc 1, Marek Oja 1, PhD 1, Maarja Pajusalu, MSc 2, Kerli Mooses, PhD 1, Sulev Reisberg 3, PhD 1, Jaak Vilo 1, PhD 1, Antonio Fernández Giménez 4, Thomas Falconer, MSc 5, Ana Danilovic, MD 4, Filip Majkovic, MSc 5, Dalia Dawoud, PhD 6, Raivo Kolde, PhD 1

1Institute of Computer Science, University of Tartu, Tartu 51000, Estonia; 2Biomedical Research Institute, IIS INCLIVA, Valencia 46010, Spain; 3Columbia University Irving Medical Center, New York, NY 10032, United States; 4KBC Zvezdara, Belgrade 11000, Serbia; 5Department of Health Information Systems, Heliant, Belgrade 11000, Serbia, National Institute for Health and Care Excellence, London W1C 6NA, United Kingdom

*Corresponding author: Markus Haug, MSc, Institute of Computer Science, University of Tartu, Narva mnt 18, 51009 Tartu, Estonia (markus.haug@ut.ee)

Abstract

Objective: To introduce 2 R packages that facilitate conducting health economics research on OMOP-based data networks, aiming to standardize and improve the reproducibility, transparency, and transferability of health economic models.

Materials and Methods: We developed the software tools and demonstrated their utility by replicating a UK-based heart failure data analysis across 5 different international databases from Estonia, Spain, Serbia, and the United States.

Results: We examined treatment trajectories of 47,163 patients. The overall incremental cost-effectiveness ratio (ICER) for telemonitoring relative to standard of care was £472/€422. Country-specific ICERs were £542/€492 in Estonia, £652/€592 in Spain, £652/€592 in Serbia, and £662/€592 in the US, which surpassed the established willingness-to-pay thresholds.

Discussion: Currently, the cost-effectiveness analysis lacks standard tools, is performed in ad-hoc manner, and relies heavily on published information that might not be specific for local circumstances. Published results often exhibit a narrow focus, central to a single site, and provide only partial decision criteria, limiting their generalizability and comprehensive utility.

Conclusion: We created 2 R packages to pioneer cost-effectiveness analysis in OMOP CDM data networks. The first manages state definitions and database interaction, while the second focuses on Markov model learning and profile synthesis. We demonstrated their utility in a multisite heart failure study, comparing telemonitoring and standard care, finding telemonitoring not cost-effective.

Key words: treatment trajectories, cost-effectiveness; Markov chains; observational data; OHDSI CDM.
Comparing penalization methods for linear models on large observational health data

Egil A. Frideirsson, PhD,1, Ross Williams, PhD,1, Peter Rijnbeek, PhD,1, Marc A. Suchard, MD, PhD,2,3, Jenna M. Reps, PhD1,4

1Department of Medical Informatics, Erasmus University Medical Center, 3015 GD Rotterdam, The Netherlands, 2Department of Biostatistics, University of California, Los Angeles, Los Angeles, CA 90095-1772, United States, 3VA Informatics and Computing Infrastructure, United States Department of Veterans Affairs, Salt Lake City, UT 84148, United States, 4Observational Health Data Analytics, Janssen Research and Development, Titusville, NJ 08560, United States

*Corresponding author: Egil A. Frideirsson, PhD, Department of Medical Informatics, Erasmus University Medical Center, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands (e.frideirsson@erasmusmc.nl)

Abstract

Objective: This study evaluates regularization variants in logistic regression (L1, L2, ElasticNet, Adaptive L1, Adaptive ElasticNet, Broken-ridge (BARI), and Iterative hard thresholding (IHT)) for discrimination and calibration performance, focusing on both internal and external validation.

Materials and Methods: We use data from 5 US claims and electronic health record databases and develop models for various outcomes in a major depressive disorder patient population. We externally validate all models in the other databases. We use a train-test split of 75%/25% and evaluate performance with discrimination and calibration. Statistical analysis for comparison of performance uses Friedman’s test and critical difference diagrams.

Results: Of the 840 models we develop, L1 and ElasticNet emerge as superior in both internal and external discrimination, with a notable AUC difference. BAR and IHT show the best internal calibration, without a clear external calibration leader. ElasticNet typically has larger model sizes than L1. Methods like IHT and BAR, while slightly less discriminative, significantly reduce model complexity.

Conclusion: L1 and ElasticNet offer the best discriminative performance in logistic regression for healthcare predictions, maintaining robustness across validations. For simpler, more interpretable models, L0-based methods (IHT and BAR) are advantageous, providing greater parsimony and calibration with fewer features. This study aids in selecting suitable regularization techniques for healthcare prediction models, balancing performance, complexity, and interpretability.

Key words: logistic regression; electronic health records; regularization; discrimination; calibration.
The Lancet Respiratory Medicine

The effectiveness of COVID-19 vaccines to prevent long COVID symptoms: staggered cohort study of data from the UK, Spain, and Estonia


Summary

Background Although vaccines have proved effective to prevent severe COVID-19, their effect on preventing long-term symptoms is not yet fully understood. We aimed to evaluate the overall effect of vaccination to prevent long COVID symptoms and assess comparative effectiveness of the most used vaccines (ChAdOx1 and BNT162b2).

Methods We conducted a staggered cohort study using primary care records from the UK (Clinical Practice Research Datalink [CPRD] GOLD and AURUM), Catalonia, Spain (Information System for Research in Primary Care [SIDIAP]), and national health insurance claims from Estonia (CORIVA database). All adults who were registered for at least 180 days as of Jan 4, 2021 (the UK), Feb 20, 2021 (Spain), and Jan 28, 2021 (Estonia) comprised the source population. Vaccination exposure was defined as a time-varying exposure, staggered by vaccine rollout period. Vaccinated people were identified and compared with unvaccinated people through use of COVID-19 vaccine data (GOLD, BNT162b2) and de-identified health insurance data (CORIVA database).

Results In conclusion, we found that COVID-19 vaccines are effective in preventing long COVID symptoms, with reduced risk of post-acute heart failure, pulmonary damage, and disease activity. The effect of vaccination was consistent across countries, with lower risk of long COVID symptoms among vaccinated people compared with unvaccinated people. These findings highlight the importance of vaccination in preventing long COVID symptoms and improving health outcomes in the post-pandemic era.

In summary, we show the real-world effectiveness of COVID-19 vaccines to prevent long COVID symptoms and post-COVID thromboembolic and cardiovascular complications among the Norwegian population, consistent with previous findings from other countries (UK, Spain, and Estonia). Additionally, we show the use of federated analytics applied across national borders to analyse linked real-world data mapped to the OMOP CDM. By applying publicly available scripts, we confirm the generalisability and reproducibility of two recent publications, strengthening their findings.
Prediction of 30-day, 90-day, and 1-year mortality after colorectal cancer surgery using a data-driven approach

Karoline Bendix Bräuner, Andi Tsouchnika, Malina Mashkoor, Ross Williams, Andreas Weinberger Rosen, Morten Frederik Schlaïkjær Hartwig, Mustafa Bulut, Niclas Dohrn, Peter Rijnbeek, Ismail Gögenur

Accepted: 21 February 2024
© The Author(s) 2024

Abstract

Purpose To develop prediction models for short-term mortality risk assessment following colorectal cancer surgery.

Methods Data was harmonized from four Danish observational health databases into the Observational Medical Outcomes Partnership Common Data Model. With a data-driven approach using the Least Absolute Shrinkage and Selection Operator logistic regression on preoperative data, we developed 30-day, 90-day, and 1-year mortality prediction models. We assessed discriminative performance using the area under the receiver operating characteristic and precision-recall curve and calibration using calibration slope, intercept, and calibration-in-the-large. We additionally assessed model performance in subgroups of curative, palliative, elective, and emergency surgery.

Results A total of 57,521 patients were included in the study population, 51.1% male and with a median age of 72 years. The model showed good discrimination with an area under the receiver operating characteristic curve of 0.88, 0.878, and 0.861 for 30-day, 90-day, and 1-year mortality, respectively, and a calibration-in-the-large of 1.01, 0.99, and 0.99. The overall incidence of mortality were 4.48% for 30-day mortality, 6.64% for 90-day mortality, and 12.8% for 1-year mortality, respectively. Subgroup analysis showed no improvement of discrimination or calibration when separating the cohort into cohorts of elective surgery, emergency surgery, curative surgery, and palliative surgery.

Conclusion We were able to train prediction models for the risk of short-term mortality on a data set of four combined national health databases with good discrimination and calibration. We found that one cohort including all operated patients resulted in better performing models than cohorts based on several subgroups.
We have a lot to be proud of, but we shouldn’t be satisfied yet...
Selection of European Initiatives Using the OMOP CDM

Moderator: Renske Los, PhD, Assistant Professor of Medical Informatics, Erasmus MC
European Initiatives Using the OMOP CDM

1. **OHDSI Europe National Nodes building opportunities through collaboration**  
   Renske Los, Erasmus MC, The Netherlands

2. **ONCOVALUE: Can Real-World Data Shape the Future of Health Technology Assessment in Oncology?**  
   Andreas Henriksen, Copenhagen University Hospital, Denmark

3. **DigiONE: technical challenges and solutions to European cancer OMOP conversions from hospital EHR**  
   Piers Mahon, DIGICORE, UK

4. **The PHEMS Project: New Strategies in Health Data Sharing**  
   Sofia Bazakou, The Hyve, The Netherlands

5. **Ecraid: European Clinical Research Alliance on Infectious Diseases**  
   Ankur Krishnan, Heidelberg University Hospital, Germany

6. **PHederation – the Federated Network of Pulmonary Hypertension Registries**  
   Eva-Maria Didden, Actelion, a Johnson & Johnson Company, Switzerland
OHDSI Europe National Nodes: building opportunities through collaboration

Renske Los
Erasmus MC
What do these countries have in common?

National **OHDSI** nodes established in 2024!
facilitate national and international collaboration

National Nodes
National Nodes

An OHDSI Europe National Node is a collection of organizations within a member country.

The Node builds on the strengths of the stakeholders and scientific communities of that country.

https://ohdsi-europe.org/index.php/national-nodes
National Nodes
International collaboration

• Twice yearly meeting of node leads

• Current initiatives on European level:
  – The Drug Challenge: Inventory of drug vocabulary mappings
  – Cost Action Call proposal
What do these countries have in common?
OHDSI BELGIUM
Unifying Health Data Harmonization with Open Science

- June 2023
- 140 individuals
- 64 organizations
- 15 datapartners
Denmark

Start date: 2024-04-30

Leading Organization:
Center for Surgical Science, Zealand University Hospital

Members: 18
Organizations / data partners: 10
Sulev Reisberg, Raivo Kolde
University of Tartu

- Members: 23
- Organizations: 6
- Data partners: 1
- Promoting OMOP in a number of high-level events in Estonia
The German Node

Leading organization
University of Technology Dresden
Faculty of Medicine Carl Gustav Carus
Institute of Medical Informatics and Biometrie

Start Date
founded in Spring 2021

Members
52

Michele Zoch
Research Associate, University of Technology Dresden

Ines Reinecke
Head of Data Integration Center, University Hospital Dresden

Save the Date
October, 28th and 29th

German Studyathon

... and many more
• 5 data providers supported by EHDEN funding
• 3 universities/research centres
• 4 hospitals
• 11 private companies
• 5 certified organizations/SMEs by EHDEN
• 2 research programmes focusing on OMOP-CDM
• 1 innovation programme currently supporting ETL process
• 3 multi-national industries local branches on board
• Many dissemination activities (MSc programmes, online activities etc.)

Start date: July 2021
Lead: Pantelis Natsiavas, Institute of Applied Biosciences, Centre for Research and Technology Hellas (INAB | CERTH)
Node lead: Chen Yanover, KI Research Institute
Kick-off meeting: May 20, 2024
We had >50 attendees from most Israeli health organizations (hospitals, HMOs), government, academia, and industry
Data partners: Kineret – OMOPed data of six government medical centers; more to come
OHDSI Italia

Active since **JUN 22**
Coordinated by **University of Pavia & BIOMERIS**

55+ people
33 institutions (21 data partners)
Plenary meeting every 3 months + target meetings (e.g. DPO, protocols)

- **DPIA (Data protection Impact Assessment)** template for OMOP setup in a center
- **1st OHDSI Italia project**: two studies to evaluate whether partners are ready to participate in a study using OHDSI tools

1. Characterization of the national reality of the Italian node of OHDSI (using DQD and Achilles) ➔ 14 data partners involved
2. Correlates of hospital sepsis. A longitudinal cohort study of the Italian OHDSI Node (using Atlas and HADES) ➔ 10 data partners involved
OHDSI Luxemburg

Our mission is to redefine healthcare by harnessing real-world clinical data and experiences to empower clinicians and enhance patient well-being. We aim to shape the future of medical knowledge by also focusing on OMOP Common Data Model Medical Device Extension and on Patient Reported Outcome Measures (PROMs)

**Start date:** kicked-off December 12th, 2023

**Leading organizations:**
- Information Technology for Translational Medicine S.A. (ITTM S.A.)
- Luxembourg Institute of Health (LIH)

**Members:** Andreas Kremer (ITTM S.A.), Sebastiano La Ferla (ITTM S.A.), Rachelle Krajnc (ITTM S.A.), Loic Marx (ITTM S.A.), Romain Tching (ITTM S.A.), Nils Christian (ITTM S.A.), Maria Quaranta (ITTM S.A.), Maximilian Füngfeld (LIH); Claudine Backes (LIH), David Marcic (LIH), Michaël Schnell (LIH), Guy Fagherazzi (LIH), Sophie Couffignal (LIH), Vanessa Pereira (LIH)

**Data partners:** National Cancer Register
OHDSI NL

• 61 members (+110%)
• 8 data partners
• ≈ 3 monthly meetings ('niet lullen maar poetsen')

Topics of interest:

- Mapping
  - Harmonization of Dutch code systems to OMOP
  - Collaboration on FHIR to OMOP mapping
  - Mapping Epic/HX/MetaVision to OMOP
  - Multi-source data mapping (time series data)
  - Preparing data from multiple EHRs for mapping to OMOP
  - Dutch EHR data / BGZ mapping
  - Comparability of mappings
  - Mapping questionnaires
    - PROMS in OMOP
  - NL core data
  - Novel mapping standards
    - (Dutch) Vocabulary mappings
  - Mapping cancer patient journey
- Study-a-thon
  - Multi-database validated & updated prediction model
  - Study in federated data network
  - Work towards a study-a-thon
  - Feasibility assessment of a health (broad sense) study
- Access
  - Access to data for analysis
  - Harvesting metadata into NL catalogue & EHDS
- Data linkage
  - Link/combine CDMs for more complete patient data
  - Integration of multiple data sets
- Data quality
  - Compare data quality & discuss discrepancies
    - Pros & Cons of NL data sources
    - Completeness of data (sources)
  - Comparability of data [data quality & mapping]
    - Benchmark across data sources
- Workshops
- Education

Currently in progress:

Mapping BGZ (Dutch Int’l Patient Summary) from FHIR to OMOP

Study-a-thon characterization of breast cancer patients
NorOMOP

Espen Enerly, Siri Larønningen
Cancer Registry of Norway, part of Norwegian Institute of Public Health
Start date: 10.10.2023

Data partners: 3
- Oslo University, Department of Pharmacy
- Oslo University Hospital
- Cancer Registry of Norway, Norwegian Institute of Public Health

Members organizations: 5
- Norwegian Directorate of Health
- Oslo University, Department of Pharmacy
- Oslo University Hospital
- Norwegian Institute of Public Health
  - Cancer Registry of Norway
  - Helsedata.no (Health Data Access Body)
  - Various departments

Cool to share: Node diversity
OMOP-database: (Research project, Hospital data warehouse, Registry)
Institution: Academic, University, National Registry/Public health Institute, Health Data Access Body
Network partners: Ehden network, Darwin-EU, DigiCore
Research areas: Pharmacovigilance, perinatal research and cancer research.
Portugal is at the forefront of this movement, and the formalization of the National OHDSI Node lays the foundation for a thriving community of Portuguese healthcare data partners and SMEs. This community aims to foster interdisciplinary collaboration to discuss and address national healthcare challenges through real-world evidence. The Node will also serve as a gateway for international innovation, research partnerships, and knowledge sharing around OMOP-CDM.

Leading Organization: ULS Coimbra
OHDSI SPAIN
Activities and Plans

More than 25 data partners and SMEs (established in 2023)

**Working Groups:**
- Medications (led by the Spanish Medicines Agency)
- Costs (led by the Navarra Health Service and Hospital del Mar)
- LOINC (led by Hospital del Mar)

**Meetings and Conferences:**
- 2 ftf meeting and 2 online meetings
- Spanish Node presentations: Rotterdam 2023, INNODATA 23’, and European Big DataVF 23’
- Study-a-thon planned for Q4 2024 in Barcelona

**National Initiatives:**
- RWD Cardiovascular Diseases: environmental + clinical data network – Carlos III Health Institute (Spanish Ministry of Health) (submitted)
- OHSIRIS - Open Health Space Infrastructure for Research and Industrial Services, Next Generation EU, Spanish Ministry of Digital Transformation (submitted)
- ELADAIS Project (Clinical and Omics data standardization using OMOP) – UNICO R&D Cloud Programme, Spanish Ministry of Digital Transformation (ongoing)

Coordinators: Talita Duarte-Salles (IDAPjordiGol) and Miguel Angel Mayer (Hospital del Mar)
Coordinator Assistant: Angela Leis (Hospital del Mar Research Institute)
Key achievements and directions

- We have a logo!
- Successful f2f annual meetings – 2\textsuperscript{nd} event scheduled for September in London, with >100 attendees confirmed
- Formal support and partnership with HDRUK
- NHS England Secure Data Environments will be mapped to OMOP!
- National initiatives ‘brewing’:
  - HDRUK OMOP RWE Pilot Network – Oxford-led Coordination Centre funding agreed pending contract signature
  - Innovate UK Centre for Regulatory Sciences and Innovation – Partnership with 7 national partners shortlisted for funding to work with MHRA
Visit us!

La Fontaine – deck B
13.00 – 16.00
ONCOVALUE: Can Real-World Data Shape the Future of Health Technology Assessment in Oncology?

Andreas Henriksen
Copenhagen University Hospital
ONCOVALUE: Implementing value-based oncology care at European cancer hospitals

Mads Andersen¹, Juho Lähteenmaa², Johanna Mattson², Ulrik Lassen¹, Andreas Bjerrum¹

¹Department of Oncology, Rigshospitalet, Denmark, ²IT Management and Comprehensive Cancer Center, Helsinki University Hospital, Finland

Andreas Henriksen
MD & PhD student
Department of Oncology, Rigshospitalet, Copenhagen, Denmark

Conflicts of interest:
Travel with Daiichi Sankyo
Consortium of cancer centers & technology companies – funded by Horizon Europe

1.12.2022 – 30.11.2026
7 M€ total budget
Oncovalue Mission

**Aim**
- Make available high quality Real-World Data for Health Technology Assessment
- Support development of effective treatment

**Background**
- Efficacy-effectiveness gap in Randomized Clinical Trials
- Deficient toxicity reporting in phase IV trial setting
- High cost of novel treatments pose challenge to healthcare systems
Through 9 Work Packages, ONCOVALUE will

- Implement OMOP-CDM
- Develop federated analyses platform
- Utilize unstructured Electronic Health record data with AI-based tools
Thank you!

www.oncovalue.org

https://www.linkedin.com/company/oncovalue/
DigiONE: technical challenges and solutions to European cancer OMOP conversions from hospital EHR

Piers Mahon
DIGICORE
DIGICORE’s Mission: reduce the variation in EU cancer outcomes digitally

National 5 year cancer survival, %

% 5 year age-standardized survival 2000-2007

Source: EUROCARE-5 5-year survival by tumor and country based on 2000-2007

1. Improve to top quartile = 100,000 lives a year

Top quartile

2nd quartile

3rd & 4th quartile

Switzerland

Germany

France

Norway

Italy

Portugal

Netherlands

Ireland

Malta

Spain

Denmark

Czech Republic

UK

Slovenia

Croatia

Lithuania

Estonia

Slovakia

Latvia

Poland

Bulgaria

Sweden

Finland

Iceland

Belgium

Austria

Switzerland

France

Norway

Italy

Portugal

Netherlands

Ireland

Malta

Spain

Denmark

Czech Republic

UK

Slovenia

Croatia

Lithuania

Estonia

Slovakia

Latvia

Poland

Bulgaria
DIGICORE includes 37 cancer centres in 18 countries, 21 with Cancer OMOP instances built or funded and US / Asian extensions in planning

1. TAYS Cancer Centre
   Vision Zero Cancer

2. Oslo UH

3. Edinburgh
4. Leeds
5. Trinity St James

6. Maastricht CCC
7. Saint Luc
8. UCC Frankfurt
9. IPO Porto
10 Fund. Jimenez Diaz
Start Madrid, Clínica Universitaria Navarra

Unicancer network (3)
Institut Curie,
Centre Léon Bérard,
Centre Jean Perrin

DigiONE I3 research partner (4)
11. Charite
12. Greiswald
13. Groningen
14. Dresden

US and Asian centres in conversion planning

15. Tartu University Hospital
16. National Cancer Institute
17. Maria Skłodowska Curie
18. Masaryk Memorial
Inst of Oncology Ljubljana
Sestre Milosrdnice

Key:
**Bold** - Cancer OMOP funded

ACC network (10)
IEO, 19. INT Milan,
Gemelli, Reggio Emilia
20 IRST, 21 San Raffaele,
INT Pascale
IRCCS Veneto, San Matteo
Humanitas Mirasole
Problem: Hospital data is challenging data – most will require NLP and primary data capture improvement, so it’s expensive (EHDS take note..)

Data item availability on 40 items agree by consensus “essential for good care management” (%)

- Hospital 1: >90% available & structured
- Hospital 2: >90% available & structured
- Hospital 3: >80% items available, but >30% unstructured (free text rich)
- Hospital 4: >80% items available, but >30% unstructured (free text rich)
- Hospital 5: >80% items available, but >30% unstructured (free text rich)
- Hospital 6: >80% items available, but >30% unstructured (free text rich)
- Hospital 7: >30% items missing (no technology can help)
- Hospital 8: >30% items missing (no technology can help)
- Hospital 9: >30% items missing (no technology can help)
- Hospital 10: >30% items missing (no technology can help)

- 0: Paper, or unknown
- 1: Will need reform to primary data capture
- 2: OCR + NLP on PDFs
- 3: NLP on free text
- 4: NLP on semi-structured data
- 5: ETL from structured data
The outline for DigiONE, our integrated, open standard technology stack was published in Nature Medicine, including a description of target data

1. **Minimal Essential Description Of Cancer (MEDOC)**
   Consensus built care quality minimal data set

2. **Near-real time frontline feedback loops** to improve data

3. **Pan-format Cancer data ingestion to OMOP**. Not just ETL also NLP, OCR

4. **GDPR recital 34 privacy conserving solutions** for next generation sequencing results

5. **Full federation using open-source Vantage6** to allow statistical analysis equivalent to centralised data, but without data pooling and without patient consent

6. **Modular, protocolized implementation plans** to solve for limited data normalisation skills in most hospitals

7. **All in open standards and vendor agnostic**

We use Disease Natural History with care quality assessment to clean up large cohorts of patients, even in hospitals with weak legal basis

<table>
<thead>
<tr>
<th>Example NSCLC Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients diagnosed with mNSCLC between 1st November 2018 – 30th November 2023</td>
</tr>
<tr>
<td>• Subgroups by stage at diagnosis of NSCLC</td>
</tr>
<tr>
<td>- De novo metastases</td>
</tr>
<tr>
<td>- Initially diagnosed with NSCLC at earlier stage</td>
</tr>
<tr>
<td>• Subgroups with metastases at index date in:</td>
</tr>
<tr>
<td>- Each of these single sites: Brain, liver, adrenal gland, bone, other lung, other single sites</td>
</tr>
<tr>
<td>- Multiple sites: Including brain, excluding brain</td>
</tr>
<tr>
<td>• Subgroup prescribed immunotherapies as 1st LoT for mNSCLC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Research objectives summarised</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Describe demographic and clinical characteristics, <strong>genetic phenotype, re-biopsy rates</strong>, and Tx received for NSCLC prior to index</td>
</tr>
<tr>
<td>2. Describe <strong>SACT patterns by 1st and 2nd LoT, radiotherapy and surgery</strong> for mNSCLC</td>
</tr>
<tr>
<td>3. Assess <strong>OS and TtNT by 1st and 2nd LoT</strong> incl. adjustment for prognostic characteristics</td>
</tr>
<tr>
<td>4. Describe <strong>duration of Tx, starting dose, and dose intensity</strong> by age and gender in patients prescribed <strong>1st LoT immunotherapies</strong> for mNSCLC</td>
</tr>
<tr>
<td>5. <strong>Benchmark care quality</strong> against ESMO guideline recommendations</td>
</tr>
</tbody>
</table>

**DINASTY** = clinical audit = legitimate use = legal to process anywhere

ESMO: European Society for Medical Oncology; LoT: Line of Therapy; mNSCLC: metastatic Non Small Cell Lung Cancer; OS: Overall Survival; SACT: Systemic Anti Cancer Treatment; TtNT: Time to Next Treatment; Tx: Treatment
DINASTY studies are just the first step on a longer journey to transform the impact of cancer real world evidence globally.

THE WORLD TODAY

Traditional eCRF observational study

- Manual retype post-consent into central eCRF
- Statistical analysis of centralised pooled data

DINASTY STUDIES (large scale data preparation)

Automate a cancer outcomes study

- Network wide Cancer OMOP study implementation
- Data “ready” before protocol in research environment
- Federation to get pooled data equivalent insights without pooling data

THE FUTURE...

(the exciting narrow cohort stuff)

- Health system research
- Comparative effectiveness
- Patient finding / trial planning
- Pragmatic trial
- Safety ascertainment
- Discovery – omics
- Biomarker validation
- Care quality management
- AI driven Decision Support

Key: Research users

- Pharma RWE/Med affairs
- Pharma clinical dev/R&D
- Diagnostics companies
- National governments

Join us!
The PHEMS Project: New Strategies in Health Data Sharing

Sofia Bazakou
The Hyve
New Strategies in Health Data Sharing

Sofia Bazakou, sofia@thehyve.nl
Pediatric Hospitals as European drivers for Multi-party computation and Synthetic data generation capabilities across clinical specialties and data types

3 years

October 2023 - September 2026

7 M€

Funded through Horizon Europe and UK Research & Innovation

Coordinated by HUS (Helsinki University Hospital)

The clinical cases use pediatric data stored in the OMOP Common Data Model (CDM)
PHEMS transcends boundaries, supporting access and use of health data between children’s hospitals and across borders

**Objectives**

- Increase access to health data while protecting patient privacy
- Advance federated health data analysis through predictive modelling and machine learning
- Enable on-demand generation of shareable synthetic and anonymized datasets
- Demonstrate the value of the data ecosystem using three clinical use cases in four countries

The project will create a decentralized and open health data ecosystem consisting of technical components and governance frameworks, empowering institutions to collaborate without relinquishing control over their data.

Funded by the European Union
**Project Overview**

- **Dissemination and exploitation including business models**

- **Developing the governance framework and playbook**

- **Validating the results**
  - Clinical validation of synthetic data vs. raw data
    - Clinical Use case 1: Cardiology patients operation management
    - Clinical Use case 2: Pediatric Intensive Care Unit (PICU) Sepsis
    - Clinical Use case 3: Hematology; hemophilia

**Feedback**

**Data Controller**

- HUS
- Erasmus MC
- SJD

**PHEMS data and analytics federated services**

- The Hyve
- HUS
- SJD

**PHEMS federation analytics network**

- On demand data anonymization synthetization

- **ECHO**
- **GENESIS Biomed**
- **Veil.ai**
- **tietoevry**
Find out more

Clinical Use Cases

PHEMS: validating novel federated ecosystems for analytics and synthetic data generation methods through real-world data investigations, particularly in the context of pediatric healthcare

New strategies in Health Data Sharing – Clinical Use Cases

Pediatric Hospital in Europe (PHEMS) is an innovative approach to data sharing and analytics that leverages federated ecosystems to enable real-world data investigations in pediatric healthcare. PHEMS is a network of hospitals that collaborate to share and analyze data securely, allowing for the development of analytics and synthetic data generation methods that can be used to inform and improve patient care across multiple hospitals.

VEILAI's Next-Generation Anonymization enables cutting-edge research for children's diseases

VEILAI’s Next-Generation Anonymization is beyond-state-of-the-art, enabling secondary use of pediatric health data

VEILAI’s Next-Generation Anonymization enables cutting-edge research for children's diseases

Enhancing Pediatric Care Data Collaboration through Privacy-Enhanced Federated Learning and/or Anonymization

Background: Health data is a critical resource for improving patient care and understanding disease. However, the use of health data is often limited due to privacy concerns. VEILAI is developing innovative data anonymization and privacy protection techniques to enable the secondary use of health data without compromising patient privacy.

VEILAI’s Next-Generation Anonymization enables cutting-edge research for children's diseases

Tuomo Pentikäinen

VEILAI

Funded by the European Union

UK Research and Innovation
Ecraid: European Clinical Research Alliance on Infectious Diseases

Ankur Krishnan
Heidelberg University Hospital
ECRAID-Base

Building a European clinical research alliance on infectious diseases

Presented by: Ankur Krishnan
On behalf of the Ecraid consortium
Ecraid: European Clinical Research Alliance on Infectious Diseases

- **ECRAID-Base** is a ‘warm-base’ pan-European clinical research network that facilitates faster, easier and cost-effective infectious diseases (ID) and antimicrobial resistance (AMR) research to reduce their impact on individual and population health in Europe.

- ECRAID-Base addresses **ID outbreak preparedness and response** through efficiently and effectively generating rigorous evidence to **improve the diagnosis, prevention and treatment of infections**

- This is facilitated by a **European multidisciplinary clinical research network** and innovative research approaches. The network has the capacity and capability to:
  - directly enrol patients with infectious diseases to generate evidence to support the testing and development of new diagnostic, preventive and/or therapeutic strategies and therapies
  - conduct a broad range of rigorous clinical studies efficiently and rapidly
  - function as a platform for a rapid research response in the face of serious infectious disease outbreaks
Consortium

19 organisations based in ten countries: Belgium, Croatia, Czech Republic, Italy, Germany, France, the Netherlands, Spain, Switzerland, and the United Kingdom.
Network of experts in epidemiology of infectious diseases and in the design of clinical studies, statistical methods and analyses

- **CLIN-Net**: Network of clinics and hospitals capable of quickly and reliably recruiting, treating, monitoring and reporting data for multinational, multicenter studies.
- **LAB-Net**: Network of microbiology laboratories delivering high-quality and standardized information on microbial strains and antibiotic resistance.
- **STAT-Net**: Network of statistics and clinical study design experts in infectious diseases (ID) and antimicrobial resistance (AMR) research.
- **EPI-Net**: Network of ID/AMR epidemiology and surveillance experts.
- **Penta ID**: Network devoted to advancing research on optimising the prevention, diagnosis and treatment of infectious diseases in children and in pregnancy.

- **900 labs - 42 countries**
- **300 primary care sites - 18 countries**
- **90 sites - 18 countries**
- **1251 hospitals**
- **27 Ecraid Clinical Liaisons**
- **45 sites - 11 countries**
<table>
<thead>
<tr>
<th>Study</th>
<th>Start Date</th>
<th>Participants Enrolled</th>
<th>Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>POS-VAP (Ventilator-Associated Pneumonia in ICUs)</td>
<td>August 2022</td>
<td>2967</td>
<td>40</td>
</tr>
<tr>
<td>POS-cUTI (Complicated Urinary Tract Infections in Hospitals)</td>
<td>October 2022</td>
<td>2359</td>
<td>43</td>
</tr>
<tr>
<td>POS-ARI-ER (Acute Respiratory Infections in Emergency Rooms)</td>
<td>June 2023</td>
<td>796</td>
<td>42</td>
</tr>
<tr>
<td>POS-ARI-PC (Acute Respiratory Infections in Primary Care)</td>
<td>February 2024</td>
<td>62</td>
<td>11</td>
</tr>
<tr>
<td>POS-Disease X (Unexplained febrile illness with unusual epidemiology and/or clinical presentation in Emergency Rooms)</td>
<td>December 2023</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>REMAP-CAP (Europe) (Community-Acquired Pneumonia in Hospitals)</td>
<td>March 2018</td>
<td>7928</td>
<td>131</td>
</tr>
</tbody>
</table>

**Perpetual Observational Studies (POS)**

The first studies to benefit from Ecrad’s infrastructure are -
- Five Perpetual Observational Studies (POS)
- European arm of the REMAP-CAP adaptive platform trial

A POS is a prospective, multicentre, observational clinical study that perpetually enrolls patients.

They address key clinical research gaps, including variations in clinical practices, incidence of ID syndromes, and associated risk factors.

They create a clinical research backbone, ready to concurrently or sequentially embed studies (observational, experimental, investigator-initiated, or commercial).
Our OHDSI journey till now...

Joined through the **EHDEN 5th Open Call**

Partnered with **edenceHealth NV** to transform POS-VAP data to OMOP-CDM

Established a **scalable, sustainable and secure architecture to deploy and maintain ETL pipelines and OHDSI tooling for ECRAID-Base** at UMC Utrecht (please visit our poster #13 for more details)

**Leveraged the Oncology extension (OMOP-CDM v5.4)** to map hierarchical and inter-event dependent patient-level data on microbiological identification and cure, AMR profiles, VAP clinical criteria based onset and clinical cure.

Successfully set up all infrastructure requirements for ETL validation (ACHILLES, DQD, ARES, ATLAS).

Published the data source on the **EHDEN Database Catalogue** (‘ECRAID-Base POS VAP’)

Submitted an Expression of Interest to become a **DARWIN EU data partner**

ECRAID-Base metadata published on **EMA RWD catalogues** (Data source ID – 1111205)

Executed **MoU between Ecraid and EHDEN foundations**

Links between each patient’s microbiological identification and microbiological cure episodes, as well as their VAP onset and clinical cure episodes, can provide significant clinical insights into disease diagnoses, presentation, progression, treatment, and prognosis.
Next steps and future directions

Continue transformation of the **remaining POSs to OMOP-CDM** through CoMeCT project

**Train and educate the researchers and clinicians in our consortium** on OMOP-CDM and OHDSI suite of tools and packages

**Participate in OHDSI working groups** (Common Data Model, CDM vocabulary subgroup, Registry, Natural Language Processing, Clinical Trials, Patient-level Prediction, Phenotype Development and Evaluation)

Leverage ECRAID-Base study data in OMOP-CDM, established infrastructure and partnerships with EHDEN, OHDSI and DARWIN EU communities to **participate in and initiate projects and studies on** -

- Diagnosis, prevention and treatment of infections
- Infectious diseases epidemiology and antimicrobial resistance surveillance
- Pandemic preparedness and response
- Antimicrobial Consumption (AMC) and Antimicrobial Stewardship (AMS)
- And, many more...

Initiate an **ID-AMR working group** to develop mapping guidelines and best practice recommendations

**Identify a core AMR data structure** that harmonizes the various potential sources of AMR data (EHR, LIMS, Cohorts, etc.)

**Collaborate to develop a ‘Microbiology extension’**, which leverages and builds upon the hierarchical structure in ‘Oncology extension’ and incorporates the nuances and complexities of ID/AMR research data
Thank you for your attention!

Got a question?
Please get in touch with me during the symposium
Visit our website - https://www.ecraid.eu/
Or, please reach out to me at - Ankur.Krishnan@uni-heidelberg.de
PHederation – the Federated Network of Pulmonary Hypertension Registries

Eva-Maria Didden, Actelion, a Johnson & Johnson Company
The federated network of Pulmonary Hypertension registries

Presenter: Eva-Maria Didden
Co-authors: Valerie van Baalen, Michel van Speybroeck, Monika Brand
Date: June 3, 2024
Location: European OHDSI Symposium, SS Rotterdam, the Netherlands
Background

- Pulmonary Arterial Hypertension (PAH) is a rare subgroup of Pulmonary Hypertension (PH).
- Real-World Evidence (RWE) generation in rare diseases is often restricted due to the small patient numbers, geographic distribution, and limited data access.
- In disease-agnostic EHR or insurance claims databases, it is challenging to accurately identify PH or PAH patients.
- PHederation is a disease-specific Federated Data Network (FDN) that brings together multiple fit-for-purpose PH data sources.
Introduction

PHederation is a public-private partnership connecting* harmonized** disease-specific clinical data sources and subject matter experts, for enhanced research in Pulmonary Hypertension.

➔ PHederation will contribute to an increase in transparency and reproducibility of RWE in PH***.

---


** A) Standardizing PH registry data to the OMOP Common Data Model: https://pubmed.ncbi.nlm.nih.gov/34727871/
B) Handbook for PH registries to OMOP CDM conversion: https://github.com/OHDSI/ETL--PulmonaryHypertensionRegistries

*** Increase transparency and reproducibility of RWE in rare diseases through disease-specific FDNs: https://pubmed.ncbi.nlm.nih.gov/38556812/
**PHederation Portal**

### DATABASE

<table>
<thead>
<tr>
<th>Database</th>
<th>Description</th>
<th>Observation period</th>
<th># PH and #PAH</th>
<th>Regions</th>
<th>Source data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canadian PH Registry (CPHR)</td>
<td>Prospective PH patient registry</td>
<td>2017 - ongoing</td>
<td>PH: 1,995, PAH: 1,076</td>
<td>Canada</td>
<td>PAHTool</td>
</tr>
<tr>
<td>EXPOSURE (EUPAS19085)</td>
<td>Registry of PAH patients newly treated with Uptravi or another PAH-specific therapy</td>
<td>2017 - ongoing</td>
<td>PAH: 2,354</td>
<td>Europe, Canada</td>
<td>CDISC SDTM</td>
</tr>
<tr>
<td>OPUS (NCT02126943)</td>
<td>Opsumit drug registry</td>
<td>2014 - 2018</td>
<td>PH: 2674, PAH: 1,076</td>
<td>USA</td>
<td>CDISC SDTM</td>
</tr>
<tr>
<td>ORPHEUS (NCT03197688)</td>
<td>Opsumit user medical chart review to supplement OPUS</td>
<td>2013 - 2017</td>
<td>PAH: 2,206, PAH: 2,031</td>
<td>USA</td>
<td>CDISC SDTM</td>
</tr>
<tr>
<td>Porto center of Portuguese PH network</td>
<td>Northern Region Portuguese PH registry</td>
<td>2001 - ongoing</td>
<td>PAH: 2,410, PAH: 216</td>
<td>Portugal - Northern Region</td>
<td>PAHTool</td>
</tr>
<tr>
<td>SPHERE (NCT03278002)</td>
<td>Selexipag drug registry</td>
<td>2016 - 2020</td>
<td>PH: 829, PAH: 759</td>
<td>USA</td>
<td>Registry-specific</td>
</tr>
<tr>
<td>Stanford clinical PH database</td>
<td>PH Registry</td>
<td>2004 - ongoing</td>
<td>PH: 1,189, PAH: 987</td>
<td>USA - Western Region</td>
<td>Registry-specific</td>
</tr>
</tbody>
</table>

### Additional Information

- **Modules**
  - Data Catalogue
  - Study Catalogue
  - Task Manager

- **Development**
  - R Studio Connect

- **Analysis**
  - Atlas
  - PHederation Dashboard

- **Knowledge**
  - User Documentation
  - Forum

---

**J&J Innovative Medicine**
1. Select databases using *Data Catalog* and *Dashboards*; perform fit-for-purpose evaluations.

2. Create study team to develop protocol and analysis plan.

3. Translate analysis plan into standard queries.

4. Distribute queries, execute analysis, and collect aggregate results; potentially conduct meta-analysis.

5. Interpret results, write study report, publish.
Conclusion & Outlook

PHederaion established a network of databases of diverse purpose and origin:

- with the goal of advancing scientific knowledge in PH through distributed data sources and analytics, harmonization, and automation.
- to expand the breadth and depth of individual PH databases, to increase diversity and geographic coverage, and to accelerate and enhance RWE generation.

Visit us at PHederaion.org!

First PHederaion network study (ongoing): Drug utilization of endothelin receptor antagonists and phosphodiesterase-5 inhibitors in newly-diagnosed PAH patients. This study will complement Darwin EU's EUPAS106052* with evidence from a disease-specific FDN.

* EUPAS106052: https://catalogues.ema.europa.eu/node/3797/administrative-details
Improve the care for patients with Pulmonary Hypertension with the power of real-world data.
Coffee break!

Next session starts here @ 11:30!