



OHDSI2025 Global Symposium Preview

OHDSI Community Call
Sept. 9, 2025 • 11 am ET



Upcoming Community Calls

Date	Topic
Sept. 9	Global Symposium Preview
Sept. 16	OHDSI/OMOP Research Spotlight
Sept. 23	Educating on OHDSI: Lessons Learned
Sept. 30	OHDSI 2025 Poster Preview Mad Minutes / Symposium Logistics
Oct. 7	No Call – OHDSI Symposium
Oct. 14	Welcome to OHDSI
Oct. 21	Meet the Titans



Sept 16: OHDSI/OMOP Research Spotlight



Jessie Tong

Assistant Professor, Johns Hopkins University

Unlocking efficiency in real-world collaborative studies: a multi-site international study with one-shot lossless GLMM algorithm • *NPJ Digital Medicine*



Kim López Güell

Dphil Candidate, University of Oxford

Clusters of post-acute COVID-19 symptoms: a latent class analysis across 9 databases and 7 countries • *Journal of Clinical Epidemiology*



Jen Wooyeon Park

PhD Student, Johns Hopkins University

Breaking data silos: incorporating the DICOM imaging standard into the OMOP CDM to enable multimodal research • *JAMIA*



Abigail Newbury

PhD Student, Columbia University

Multi-domain rule-based phenotyping algorithms enable improved GWAS signal • *NPJ Digital Medicine*



Benjamin Martin

Postdoctoral Fellow, Johns Hopkins University

Identification of Adult Dermatomyositis Patients Using Real-World Data Sources • *Arthritis Care and Research*



Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?




OHDSI Shoutouts!



Congratulations to the team of **Junqing Xie, Mike Du, Yuchen Guo, Cesar Barboza, James T Brash, Antonella Delmestri, Talita Duarte-Salles, Jasmine Gratton, Romain Griffier, Raivo Kolde, Wai Yi Man, Núria Mercadé-Besora, Marek Oja, Sarah Seager, Katia Verhamme, Dina Vojinovic, Edward Burn, Daniel Prieto-Alhambra, Martí Català, Annika M Jödicke** on the publication of **Trends in prescription opioid use in Europe: A DARWIN EU® multinational cohort study including seven European countries** in *Frontiers in Pharmacology*.

 **frontiers** | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 18 August 2025
DOI 10.3389/fphar.2025.1608051

 Check for updates

OPEN ACCESS

EDITED BY
Ippazio Cosimo Antonazzo,
University of Milan-Bicocca, Italy

REVIEWED BY
Masafumi Yoshimura,
Faculty of Rehabilitation Kansai Medical
University, Japan
Irene Wood,
Centro de Investigaciones Biomédicas
(CEINBIO), Uruguay

*CORRESPONDENCE
Daniel Prieto-Alhambra,
✉ daniel.prietoalhambra@ndorms.ox.ac.uk

[†]These authors have contributed equally to
this work

RECEIVED 08 April 2025
ACCEPTED 15 July 2025
PUBLISHED 18 August 2025

CITATION
Xie J, Du M, Guo Y, Barboza C, Brash JT,
Delmestri A, Duarte-Salles T, Gratton J,
Griffier R, Kolde R, Man WY, Mercadé-Besora N,
Oja M, Seager S, Verhamme K, Vojinovic D,
Burn E, Prieto-Alhambra D, Català M and
Jödicke AM (2025) Trends in prescription opioid
use in Europe: A DARWIN EU® multinational
cohort study including seven
European countries.
Front. Pharmacol. 16:1608051.
doi: 10.3389/fphar.2025.1608051

COPYRIGHT
© 2025 Xie, Du, Guo, Barboza, Brash, Delmestri,
Duarte-Salles, Gratton, Griffier, Kolde, Man,
Mercadé-Besora, Oja, Seager, Verhamme,
Vojinovic, Burn, Prieto-Alhambra, Català and
Jödicke

Trends in prescription opioid use in Europe: A DARWIN EU® multinational cohort study including seven European countries

Junqing Xie^{1†}, Mike Du^{1†}, Yuchen Guo¹, Cesar Barboza²,
James T. Brash³, Antonella Delmestri¹, Talita Duarte-Salles^{2,4},
Jasmine Gratton³, Romain Griffier⁵, Raivo Kolde⁶, Wai Yi Man¹,
Núria Mercadé-Besora⁴, Marek Oja⁶, Sarah Seager³,
Katia Verhamme², Dina Vojinovic⁷, Edward Burn¹,
Daniel Prieto-Alhambra^{1,2*}, Martí Català¹ and Annika M. Jödicke¹

¹Pharmaco- and Device Epidemiology Group, Centre of Statistics in Medicine, NDORMS, University of Oxford, England, United Kingdom, ²Department of Medical Informatics, Erasmus Medical Center, Rotterdam, Netherlands, ³QVIA, Real World Solution, Brighton, United Kingdom, ⁴Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJG), Barcelona, Spain, ⁵Public Health Department, Medical Information Service, Bordeaux University Hospital, Bordeaux, France, ⁶Institute of Computer Science, University of Tartu, Tartu, Estonia, ⁷QVIA, Real World Solutions, Amsterdam, Netherlands

Background: The opioid crisis has been a serious public health challenge in North America for decades, despite numerous efforts to mitigate its devastating consequences. As concerns grow about a similar situation developing in Europe, we evaluated the trends in opioid use and characterized prescribing indications across seven European countries.



OHDSI Shoutouts!



Congratulations to the team of **Daniel Neumann, Richard Gebler, Jana Kiederle, Jördis Beck, Fabio Aubele, Alexander Struebing, Florian Schmidt, Matthias Reusche, Helene Koester, Markus Loeffler, Sebastian Staeubert** on the publication of **Development and Implementation of an Open, Modular, and Participatory Toolchain for Distributed IT Development in Healthcare Research – Lessons Learned** in *Volume 331 of Studies in Health Technology and Informatics: German Medical Data Sciences 2025: GMDS Illuminates Health*.

378

German Medical Data Sciences 2025: GMDS Illuminates Health

R. Röhrig et al. (Eds.)

© 2025 The Authors.

This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/SHIT251418

Development and Implementation of an Open, Modular, and Participatory Toolchain for Distributed IT Development in Healthcare Research - Lessons Learned

Daniel NEUMANN^{a,1}, Richard GEBLER^b, Jana KIEDERLE^c, Jördis BECK^d, Fabio AUBELE^e, Alexander STRUEBING^f, Florian SCHMIDT^g, Matthias REUSCHE^h, Helene KOESTERⁱ, Markus LOEFFLER^j, and Sebastian STAEUBERT^a

^a Institute for Medical Informatics, Statistics and Epidemiology (IMISE), University Leipzig, Germany

^b Institute for Medical Informatics and Biometry, Carl Gustav Carus Faculty of Medicine, Technische Universität Dresden, Dresden, Germany

^c Institute of Medical Informatics, University Hospital Heidelberg, Germany

^d University Hospital Giessen and Marburg, Giessen, Germany

^e Medical Data Integration Center, LMU University Hospital, Ludwig-Maximilians-Universität München, Munich, Germany

^f Medical Centre for Information and Communication Technology, Universitätsklinikum Erlangen, Germany

ORCID: Daniel Neumann <https://orcid.org/0000-0002-4639-5189>

Abstract. Introduction: Distributed healthcare research infrastructures face significant challenges when translating routine clinical data into harmonized, research-ready formats using HL7 FHIR standards. **State of the Art:** Existing FHIR-based pipelines such as the SMART/HL7 FHIR Bulk Data Access API, FHIR-to-OMOP mappings, and analytical services like Pathling demonstrate technical feasibility. However, most assume semantically valid FHIR data, operate within single-institution settings, and lack practical guidance for deployment across heterogeneous, regulated environments. **Technical Framework and Deployment:** Within the German Medical Informatics Initiative (MII) and the INTERPOLAR project, we developed an open, modular, and participatory toolchain for decentralized FHIR-based data transformation and export across multiple Data Integration Centers (DICs). The toolchain supports FHIR extraction, profile-based transformation, REDCap integration, and OMOP-compatible export. Deployment required adapting to local infrastructures, regulatory boundaries (e.g., de-identified FHIR stores, restricted network access), and clinical domain needs. Configurable modules, proxy support, and site-specific adaptations were essential for integration into operational hospital workflows. **Lessons Learned:** Key lessons include the necessity of early access to real data, the limitations of synthetic test data, the value of joint workshops for profile interpretation, and the need for adaptable validation tooling. Organizational knowledge gaps, inconsistent FHIR implementations, and performance issues in resource flattening were addressed through co-design and



Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?



Upcoming Workgroup Calls



Date	Time (ET)	Meeting
Tuesday	12 pm	ATLAS/WebAPI
Tuesday	12 pm	Generative AI and Analytics
Tuesday	3 pm	Oncology Outreach/Research Subgroup
Wednesday	9 am	Patient-Level Prediction
Wednesday	2 pm	Natural Language Processing
Wednesday	7 pm	Eyecare and Vision Research
Thursday	7 am	Europe Community Call
Thursday	9:30 am	Network Data Quality
Thursday	10 am	Rare Diseases
Thursday	10:30 am	Evidence Network
Friday	9 am	Phenotype Development & Evaluation
Friday	10 am	GIS – Geographic Information System
Friday	11 am	Clinical Trials
Friday	11:30 am	Steering
Friday	11 pm	China Chapter
Monday	9 am	Vaccine Vocabulary
Monday	10 am	Africa Chapter
Monday	10 am	Getting Started Subgroup
Monday	11 am	Data Bricks User Group
Monday	2 pm	Electronic Animal Health Records
Tuesday	10 am	CDM Survey Subgroup



2025 Europe Community Calls

Date	Topic
Sept. 11	Europe Community Call Introduction / DARWIN EU Update
Oct. 9	TBA
Nov. 13	Patient-Reported Outcome Measures (PROMs)
Dec. 11	Vocabularies in Europe



OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Who We Are > Updates & News > Standards > Software Tools > Network Studies > Community Forums > Education > New To OHDSI? >

Community Calls > Past Events > Workgroups > 2024 'Our Journey' Annual Report > Current Events > Support & Sponsorship >

2025 Africa Symposium > 2025 APAC Symposium > Github > YouTube > Twitter > LinkedIn > Newsletters >

2025 Community Calls

Europe Community Calls

APAC Community Calls

2024 Community Calls

2023 Community Calls

2022 Community Calls

2021 Community Calls

Welcome to OHDSI!

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

OHDSI has established an international network of researchers and observational health databases with a central coordinating center housed at Columbia University.

Read more [about us](#), about [our goals](#), and how you can [help support the OHDSI community](#).

2025 Global Symposium

Please join us at the 2025 Global Symposium, which will be held at the Hyatt Regency Hotel in New Brunswick, N.J., on Oct. 7-9. There will be tutorials Oct. 7, the main conference Oct. 8, and workshop activities Oct. 9.

[Global Symposium Homepage](#)

[Register Me for the Symposium](#)



Science Summit 2025

alongside the United Nations General Assembly (UNGA80)

9 – 26 September 2025



Science for a Sustainable Future: Showcasing Science Collaboration

The role and contribution of **science in attaining the United Nations Sustainable Development Goals (SDGs)** will be the central theme of the Science Summit. The objective is to enable science collaborations to demonstrate how science supports the attainment of the UN SDGs and Agenda 2030.

The Summit will examine what **enabling policy, regulatory and financial environments** are needed to implement and sustain the science mechanisms required to support genuinely global scientific collaborations across continents, nations and themes.

Scientific discovery through the analysis of massive data sets is at hand. This data-enabled approach to science, research and development will be necessary if the SDGs are to be achieved.

[SCIENCE FOR GLOBAL CHALLENGES →](#)

Full programme is [here](#)

<https://sciencesummitnyc.org/>



📅 Sep, Thu 18 | ⌚ 08:30 AM - 10:30 AM | 🖥️ Virtual | 📄 Public

Standardizing Health Data and Analytics to Accelerate Clinical Impact and Global Reach: Part 1

🔊 Theme: Digital / AI



Observational Health Data Science and Informatics (OHDSI) is a global community that uses data harmonized to the OMOP Common Data Model, standardized vocabulary, data quality checks and validated analytics to produce rigorous evaluation of big data from existing health databases. Through sharing of computer codes and summary statistics instead of patient-level data, OHDSI preserves privacy while enabling collaboration across institutions, countries, and continents. Large-scale, real-world studies through OHDSI network collaborations have revealed valuable insights into clinical care and public health.

Speakers:



Agnes Kiragga
Global Health Leader,...

Organization: African Population
Health and Research Centre



Chan Seng You
Assistant Professor

Organization: Yonsei University College
of Medicine



Nicole Pratt
Professor, Biostatistic...

Organization: University of South
Australia



George Hripcsak
Professor, Biomedical...

Organization: Columbia University

Register

Session details



📅 Sep, Thu 18 | ⌚ 11:00 AM - 12:45 PM | 🖥️ Virtual | 🟢 Public

Standardizing Health Data and Analytics to Accelerate Clinical Impact and Global Reach: Part 2

🔊 Theme: Digital / AI



Observational Health Data Science and Informatics (OHDSI) is a global community that uses data harmonized to the OMOP Common Data Model, standardized vocabulary, data quality checks and validated analytics to produce large-scale evaluation of real world data. Through sharing of computer codes and summary statistics instead of patient-level data, OHDSI preserves privacy while enabling collaboration across institutions, countries, and continents. Large-scale, real-world studies by OHDSI members have revealed valuable insights into clinical care and public health.

Speakers:



Cynthia Sung
Adjunct Associate...

Organization: Duke-NUS Medical School Centre of Regulatory Excellence



Patrick Ryan
VP Janssen...

Organization: OHDSI Observational Health Data Science and Informatics



Katia Verhamme
Associate Professor of...

Organization: Erasmus University Medical Center



Peter Rijnbeek
Professor, Medical...

Organization: Erasmus University Medical Center



Julio Oliveira
CEO

Organization: Precision Data



Registration links

Part 1 Sep 18, 8:30-10:30 EDT: <https://event.sciencesummitnyc.org/list-of-sessions/detail/131>

Part 2: Sep 18, 11:00-13:00 EDT <https://event.sciencesummitnyc.org/list-of-sessions/detail/130>

Full programme here: <https://event.sciencesummitnyc.org/list-of-sessions>

Part 1 (8:30 am ET)

1. *Observational Health Data Science and Informatics (OHDSI): Inclusive and Collaborative Science.* George Hripcsak
2. *Promoting Data Harmonization and Data Science in Africa.* Agnes Kiragga
3. *Rapid Response to the Covid-19 Pandemic Using a National Scale Database.* Chan Seng You
4. *OHDSI in Asia and the Pacific Rim.* Nicole Pratt
5. *Q&A Session*

Part 2 (11 am ET)

1. *Enabling Reliable Evidence Generation from Real-world Data in Europe.* Peter Rijnbeek
2. *DARWIN-EU® – Delivering Real World Evidence to Support Regulatory Decision-making by the European Medicines Agency.* Katia Verhamme
3. *OHDSI Adoption and Current Implementation Landscape in Latin America.* Julio Cesar Barbour Oliveira
4. *Learning Opportunities for OHDSI Skills Development.* Cynthia Sung
5. *Clinical and Public Health Impact of OHDSI.* Patrick Ryan
6. *Q&A Session*



Titan Award Nominations Are Due TONIGHT!

The Titan Awards have been handed out annually since 2018 to recognize OHDSI collaborators (or collaborating institutions) for their contributions towards OHDSI's mission.

Nominations for the 2025 Titan Awards are now open. **Please complete your nominations by our Sept. 9 (8 pm ET) deadline!**

ohdsi.org/titan-awards





Upcoming Presentation (Sept. 29, 1-2 pm ET)

Delivery, Not Hype: How to Harmonise FHIR × openEHR × OMOP in practice

A Health Futures Collective Series - powered by Evidentli, Fuzzy, openEHR, Professional Development and Summer Programmes, Institute of Extended Learning at Imperial College

Confirmed panellists

Grahame Grieve — Inventor of FHIR

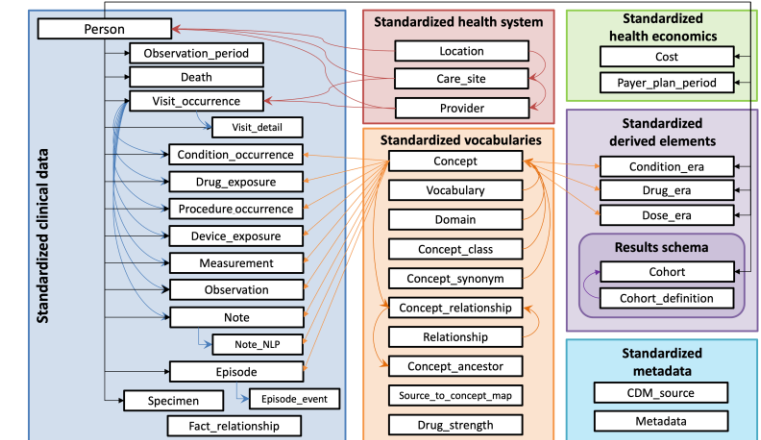
Rachel Dunscombe — CEO, openEHR Foundation

Davera Gabriel — Lead Author, FHIR to OMOP IG

Guy Tsafnat — Expert in real-world data & AI

Moderator

Ram Rajaraman — Healthcare & Life Sciences Lead, Quantexa





Jamie Weaver Scholarship at University of Oxford

< PhDs

♥ **Improving the quality of real world evidence by measuring and minimising outcome misclassification using the OMOP common data model and large multinational health data (Botnar-2025-8)**

University of Oxford > Botnar Research Centre

👤 Prof Dani Prieto-Alhambra 📅 Tuesday, December 02, 2025

📁 Funded PhD Project (Students Worldwide)

About the Project

This scholarship and work has been proposed to continue and expand work started by the late James (Jamie) Weaver. Jamie was a talented and bright data scientist and DPhil student working with us on the use of methods to minimise the impact of outcome misclassification in real world evidence (RWE). Funding has been secured, from the Medical Sciences Division, Brasenose College, and NDORMS, for this project to continue his important work on this extremely relevant topic; the successful candidate will be assigned to Brasenose College.

Real world evidence (RWE) is generated by leveraging and processing large routinely collected health data. Despite difficulties in the analysis of such information for causal inference purposes, RWE has recently been shown as a reliable source of data when used using adequate methods for trial emulation [1, 2]. We participate in multiple European and international networks to generate reliable information to inform, amongst others, regulatory decision making and health technology assessments.

Through ongoing collaborations, we leverage multiple international datasets mapped to the Observational Medical Outcomes Partnership (OMOP) Common Data Model in a federated manner. Previous work led by our student Jamie Weaver uncovered the impact of outcome misclassification on the estimation of background rates of adverse events, and proposed new methods to account for this in future studies [3].

Through this 3-year PhD funded studentship, we aim to investigate how novel methods can be applied to measure and account for outcome misclassification in RWE studies, by researching:

1. The use and application of artificial intelligence (and specifically large language models) for the generation and validation of computable phenotypes
2. The impact of outcome misclassification in different data assets
3. The performance of existing and novel methods to account for outcome misclassification in international RWE studies



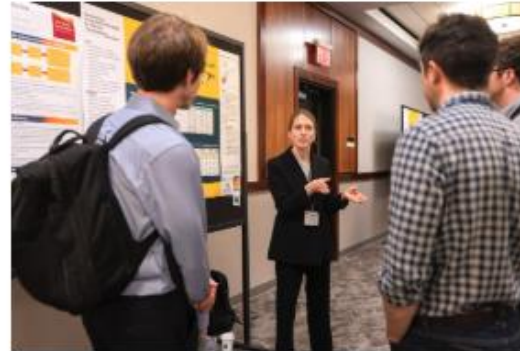
Global Symposium: Oct. 7-9



OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

- Who We Are ▾
- Updates & News ▾
- Standards
- Software Tools ▾
- Network Studies ▾
- Community Forums ▾
- Education ▾
- New To OHDSI? ▾
- Community Calls ▾
- Past Events ▾
- Workgroups ▾
- 2024 'Our Journey' Annual Report
- Current Events ▾
- Support & Sponsorship
- 2025 Global Symposium ▾
- 2025 Africa Symposium
- 2025 APAC Symposium
- Github
- YouTube
- Twitter
- LinkedIn
- Newsletters ▾

- 2025 Global Symposium Homepage
- Register for OHDSI2025
- OHDSI 2025 Agenda
- OHDSI 2025 Collaborator Showcase
- OHDSI 2025 Tutorials



2025 OHDSI Global Symposium

ohdsi.org/ohdsi2025







Africa Symposium: Nov. 10-12

[2025 Global Symposium](#) [2025 Africa Symposium](#) [2025 APAC Symposium](#) [Github](#) [YouTube](#) [Twitter](#) [LinkedIn](#) [Newsletters](#)

Join Us At The Inaugural OHDSI Africa Symposium

Nov. 10-12, 2025 • Joint Clinical Research Centre (JCRC) & Mestil Hotel Kampala



The inaugural OHDSI Africa Symposium will be held in Kampala at the Joint Clinical Research Centre (JCRC) and Mestil Hotel. Our community is delighted to introduce a new face-to-face opportunity in Africa, where OHDSI is growing at an exciting pace. We hope you will join us for this historical moment.

The first OHDSI Africa symposium will be hosted by JCRC and will begin with a dedicated one-day training course at JCRC, followed by a two-day main conference at Mestil hotel. Below are some important dates for you to save to your calendar:

Collaborator Showcase

- Submissions deadline: September 10
- Submissions review: September 11-30
- Notification of acceptance: October 5

Symposium

- Tutorial: November 10 at JCRC
- Main conference: November 11-12 at Mestil Hotel

Mestil Hotel Accommodations

Booking Code: JCRC
Booking Link: https://direct-book.com/properties/MestilDIRECT?promotion_code=JCRC25

Register Me for the 2025 OHDSI Africa Symposium!

2025 OHDSI Africa Symposium Full Agenda

ohdsi.org/africa2025



APAC Symposium: Dec. 6-7

The 2025 OHDSI APAC Symposium will be held Dec. 6-7 in Shanghai, China at the Shanghai Jiao Tong University. It will feature a 1-day tutorial and a 1-day main conference. Here are some important dates for you to save to your calendar:

Collaborator Showcase

- Submissions deadline: September 7
- Submissions review: September 8 – October 9
- Notification of acceptance: October 17



ohdsi.org/apac2025



#OHDSISocialShowcase This Week

Monday

ECRAID-Base OMOP-CDM ETL Architecture for Scalable Health Data Integration

(Panagiotis Gialernios, Shirah Cashriel, Marc Padros Goossens, Frank Leus, Freija Descamps, Lauren Maxwell, Ankur Krishnan)

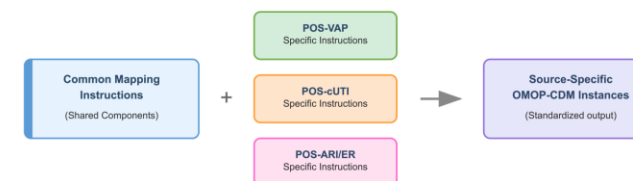
A Modular ETL Framework for Transforming Multi-Source Medical Data to OMOP Common Data Model

Title: Architecture for Scalable Health Data Integration

Background: ECRAID-Base is an EU-funded project aiming to efficiently generate rigorous evidence to improve the diagnosis, prevention, and treatment of infections, while responding to emerging infectious diseases and antimicrobial resistance threats. At its core are 5 perpetual observational studies (POS), with POS-VAP (Ventilator Associated Pneumonia) harmonized to OMOP-CDM through the EHDEN (European Health Data and Evidence Network) project. To optimize the harmonization process, common medical reference data such as pathogen and virus lists are mapped once as shared components and reused across all POS studies ensuring consistency while reducing redundant mapping efforts.

Methods: The ECRAID-Base data is collected in a question/answer format, with each pair mapped to OMOP-CDM using expert-defined instructions that specify target tables, fields, and concept IDs.

Modular ETL Architecture Overview



Modular approach enables easy addition of new data sources while maintaining consistency
Each data source maintains its specific mappings while leveraging common instructions

Key Updates to ECRAID-Base ETL Codebase



Framework Benefits

- ✓ Isolated components ensure source-specific changes don't impact other data sources
- ✓ New data sources are easier to add
- ✓ Streamlined integration while maintaining scalability

Conclusion: Our modular ETL architecture significantly enhances scalability, maintainability, and adaptability for integrating heterogeneous medical data into OMOP CDM. The framework's success with harmonized eCRFs (electronic Case Report Forms) under ECRAID-Base demonstrates the importance of broader harmonization across medical data sources, streamlining management and ensuring adaptability to evolving health data challenges.



Panagiotis Gialernios*, Shirah Cashriel*, Marc Padros Goossens*, Frank Leus*, Freija Descamps*, Lauren Maxwell*, Ankur Krishnan*
*EvidenceHealth NV
*European Clinical Research Alliance on Infectious Diseases (ECRAID)



ecraid



#OHDSISocialShowcase This Week

Tuesday

Multiple myeloma lines of treatment: from drug_exposure to proper Episode table

(Dmytro Dymshyts, Rupa Makadia, Laura L. Hester)

Multiple myeloma lines of treatment: from drug_era to proper OMOP Episode table

How to extract Multiple Myeloma (MM) treatment regimen from drug_era into Episode table

Background: The drug data can be abstracted on the 4 principal levels: drug exposure (single drug administration or prescription), drug era (continuous drug administration or prescription), treatment regimen (drugs used in combination with fixed schedule), line of therapy (several regimens used consecutively united by one clinical intent). The first two levels exist in our OMOP common data model (CDM) datasets, and we need an effective way of capturing and storing lines and regimen.

Methods

Main principles applied:

1. First regimen starts as any MM-specific drug exposure after MM diagnosis
2. All MM-specific drugs within 30 days of regimen start considered a part of regimen (applicable to first and subsequent regimen)
3. Regimen ends if either:
 - a. New drug is added
 - b. Drug is removed (not used for more than 90 days)
4. New regimen starts as either
 - a. Next day of the previous regimen end
 - b. Start of the new drug era, if there was gap between previous regimen
5. Maintenance therapy is a monotherapy duration >60 days that follows a combination therapy that lasts more than 30 days
6. Regimens are grouped into line of therapy when they are:
 - a. Stem cell transplant and related (conditioning/lymphodepletion) therapy
 - b. Apheresis, anti-plasma cell treatment, CAR-T
 - c. Regime and its corresponding maintenance therapy
 - d. Addition of immunomodulatory or proteasome inhibitor drugs within 90 days of the previous regimen start
7. The regimen are mapped to HemOnc concepts by matching ingredients and populate EPISODE. episode_object_concept_id
8. Custom concepts were created to support the line of treatment.
9. The line of treatment becomes a parent episode of treatment regimen

Results

Figure 1. Abstraction of drug era into Regimen and line of therapy in individual patient

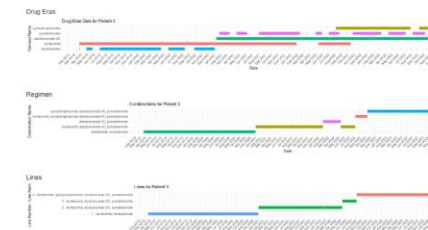


Table 1. Characterization of treatment regimen

	Flatiron MM	Optum EHR	Optum SES	HealthVerity
total number of patients	17433	39311	44696	71005
median count of regimen	2	2	2	2
median count of lines	1	1	1	2
median length of regimen, days	66	56	89	85
median length of line, days	126	85	129	132
most frequent line of therapy	lenalidomide, dexamethasone (with bortezomib, lenalidomide, dexamethasone as the next most common)		bortezomib, lenalidomide, dexamethasone	bortezomib, lenalidomide, dexamethasone

Regimen and lines are consistent across the databases

Conclusion: The algorithm creates two additional levels of abstraction of multiple myeloma treatment: Regimen and Lines of therapy, - which are useful in oncological observational research. It shows consistent and plausible results in US databases, with our next steps to abstract such regimen for our in-house European databases as well as sharing the approach with OHDSI data partners across Europe.



Dmytro Dymshyts, Rupa Makadia, Laura L. Hester

Johnson & Johnson



#OHDSISocialShowcase This Week

Wednesday

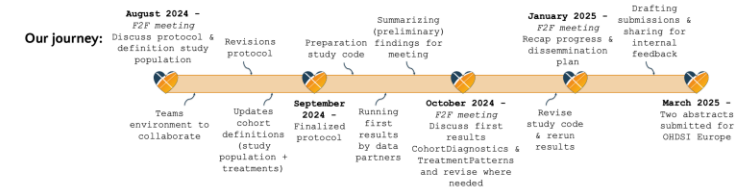
Lessons learned from the OHDSI-NL study-a-thon on breast cancer

(Aniek F. Markus, Sofia Bazakou, Renske Los, Julia Kurps, Jelle Evers on behalf of OHDSI NL)

The OHDSI NL study-a-thon has been a fun and productive way to learn together, but we need to expand our network of data partners in the Netherlands!

Lessons learned from the OHDSI NL study-a-thon on breast cancer

Background: To foster collaborations on both national and global level, OHDSI national nodes have been set up in several countries. The Dutch node (OHDSI NL) initiated a network study to assess the extent to which we are ready to execute network studies and answer clinical questions at a national level in the Netherlands. In this work, we reflect on our findings and share the key lessons learned throughout the process.



Lessons learned:

Start with clearly defining the research question. We were striving for too many goals at first, which made the effort less focused.

Perform a feasibility check (leading to a go / no go):

- Assess whether we have the clinical expertise (domain experts) and right databases (feasibility).
- Identify for each database which part of the study can be contributed to.
- Assess whether OMOP CDM fits the use case, benefits and drawbacks?

Define meeting habits. We started with face-to-face meetings but later added intermittent online meetings. We found regular short check-ups helped everyone to prioritize work and keep making progress.

Clear timelines are essential. Given that most participants are contributing voluntarily and there are no external deadlines, it is important to align on the time span to move forward.

Organization is key

Clearly assign tasks and let more and less experienced people work together/share responsibilities. This way older can share knowledge and newbies can learn.

Start by setting up a collaboration space. Immediately create a central place:

- To store documents (e.g. Teams channel). In the beginning we circulated documents via e-mail, however, this made version tracking difficult.
- To store the study code (e.g. GitHub repository). Given that individuals are working in different organizations there is no single space to use for the project by default.

Because of our intention to do this as a team, we had difficulties breaking the protocol. Not everyone being present at all meetings meant that we didn't always have the right expertise in the room. We ended up revising a lot which caused delays with starting approvals.

We do really learn by doing

No project as driving force, purely volunteer based. Some data partners were expected to participate but didn't manage because of other priorities.

The study-a-thon led to insightful discussions about topics important to OHDSI NL members, e.g.:

- Awareness that database differences can impact analysis. For example, we found the (implementation of) observation periods can be different across databases which is very important for estimating incidence. This can go wrong if study packages are not fully understood by data partners.
- It is difficult to make cohort definitions that work in multiple datasets especially if these contain different information. As it is important to understand the data, it is beneficial to actively involve the data partners in this step.

Use a progress tracker to monitor task completion. We first assigned tasks in meeting notes, but didn't have an overview of progress and responsibilities. The task list in Teams worked.

Invest in preparation

Finally, understand the data approval procedure of all datasets to proactively work on a protocol that matches these needs. In our study we encountered approval procedures that were judged very differently, leading to setbacks in the timelines we set.



Sofia Bazakou, Julia Kurps, Jelle Evers, Renske Los, Aniek Markus on behalf of OHDSI NL





#OHDSISocialShowcase This Week

Thursday

Catalysing biobank diversity analysis through a common data model

(**Karyn Mégy**, Ben Hollis, Celia Burgos, Katherine R Smith, Prasad Gunasekaran, Sean O'Dell, Sebastian Wasilewski, Quanli Wang, Slavé Petrovski)

Catalysing biobank diversity analysis through a common data model.

Karyn Mégy¹, Ben Hollis¹, Celia Burgos Sequeros¹, Katherine R. Smith¹, Prasad Gunasekaran¹, Sean O'Dell¹, Sebastian Wasilewski¹, Quanli Wang², Slavé Petrovski¹.

¹ Centre for Genomics Research, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Cambridge, UK. ² Centre for Genomics Research, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Waltham, MA, USA

AstraZeneca's Centre for Genomics Research (CGR) is building one of the world's most comprehensive and ethnically diverse genetic resources by integrating genetic and phenotypic data from multiple biobanks. We are expanding the ancestral diversity of our datasets, ensuring the data used in our research reflects global populations, so our science is designed and able to benefit a broad range of people. However, to fully leverage the power of this ethnic diversity, it is essential to standardise the data into a common data model.

1. Different data sets, several data types, multiple standards

	Source of health data available in our largest cohorts					And also....
	Hospital data	Primary care	Cancer data	Questionnaires	Free text	
UK Biobank	WHO ICD9 & 10	Read2 & 3	yes	formatted	-	Lab. tests, procedures, multi-omics, drugs
US cohort #1	CM ICD9 & 10, SNOMED	CM ICD9 & 10, SNOMED	CM ICD9 & 10	-	-	Lab. tests, procedures, drugs
Mexico cohort #1	-	-	WHO ICD10	WHO ICD10, formatted	yes	Lab. Tests, drugs (in Spanish), Metabolomic
Pakistan cohort #1	-	-	-	Ad hoc data dictionary	yes	-

WHO: International Classification of Diseases, versions 8, 9 and 10. In the WHO or ICD system
SNOMED: Systematized Nomenclature of Medicine Clinical Terms

2. Tailored vs. large biobanks

Smaller tailored biobanks are focused on specific diseases or populations.

Clinical data:

- Questionnaires, not derived from electronic health records
- Collected at the point of recruitment
- Missingness and error-prone (typos)
- => non-standard OMOP concepts, or no concepts

Format of the data, or even of the questions, can change from one release to the next (e.g. 'Gender: male / female' vs. 'Gender male: yes / no').

Unique phenotypic data not typically gathered across other biobank (e.g. types of smoking/chewing tobacco) prompting considerations on the extent of original data that should be converted to OMOP.

Large biobanks are disease-agnostic, with a broad spectrum of diagnoses, medications, and measurements.

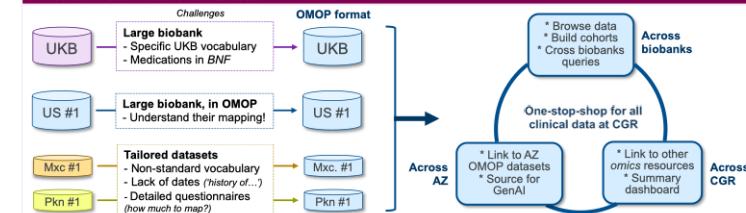
Clinical data:

- Sourced from electronic health records and optionally additional targeted questionnaires
- Well-structured, minimal missing values, systematic and automatic collection

Format remains largely consistent between releases.

However, harmonising **medications** or **measurements** poses challenges due to inconsistencies in data collection, such as varying units within the same measurement concept, or the lack of an international standard for drugs (e.g. *intest in British National Formulary –BNF- vs. active substance in Anatomical Therapeutic Chemical classifications –ATC-*)

3. A common data model for all resources



4. Take home messages

- **Harmonising data format and content** is essential for maximising the value of multiple biobanks.
- Effective transformation strategies depend on a **deep understanding of each dataset**.
- **Lack of standards** in the scientific community (e.g. measurements, medications) can limit the mapping.
- Taking a **holistic view** and considering the **broader context** is key when standardising multiple datasets.

Acknowledgements - We thank the participants and investigators in the UK Biobank and all other biobanks who made these resources available. We acknowledge members of the AstraZeneca Genomics Initiative for their contribution.

AstraZeneca



#OHDSISocialShowcase This Week

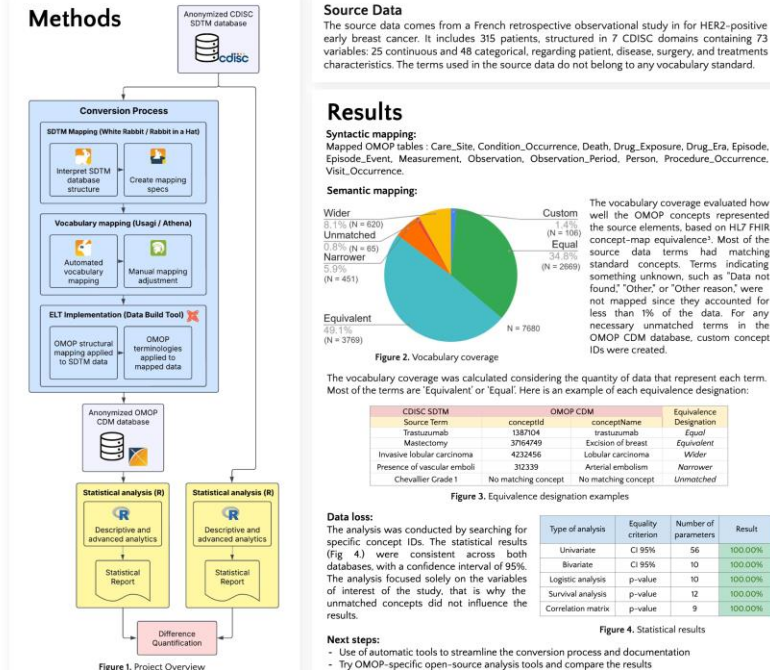
Friday

Key learnings after transformation of a CDISC SDTM database into OMOP CDM

(Amélie Lambert, Claire Castagne, Jacek Chmiel, Eric Boernert, François Margraff, David Pau, Camille Bachot, Lukasz Kaczmarek, Dimitar Toshev, Thomas Stone)

Key learnings after transformation of a CDISC SDTM database into OMOP CDM

Background: Interoperability between databases is becoming increasingly important to facilitate analyses from multiple sources. The OMOP Common Data Model (CDM) is widely used to standardize the structure and content of databases, enabling large-scale observational studies. Converting existing databases to the OMOP format is essential for leveraging its benefits. The aim is to assess the statistical and scientific usefulness^{1,2} of a transformation from CDISC Study Data Tabulation Model (SDTM) to OMOP CDM.



Key learnings: Our study showed that most of our database could be mapped to the OMOP CDM. The model is proficient at representing positive information such as what a patient has or experiences. However, representing negative information (what a patient does not have) requires additional effort and consideration, as the OMOP CDM is not designed for this purpose. Information loss during conversion varies based on the original database's level of detail and the mapping approach. Successful adaptation to the CDM necessitates modifications in both data generation and statistical analysis scripts. This involves ensuring that the terminology used in data collection matches the standardized vocabularies defined by OMOP, such as SNOMED, LOINC, and RxNorm. By doing so, it facilitates smoother data integration, reduces information loss, and enhances the accuracy of subsequent analyses.

Pau D, Bachot C, Monnet C, Vivet L, Boucher M, Sella N, Jégou R. Comparison of anonymisation techniques regarding statistical reproducibility. *PLoS Digital Health*. 2025 Feb 3;4(2):e0000735. 'Jégou R, Bachot C, Monnet C, Boernert E, Chmiel J, Boucher M, Pau D. Capability and accuracy of usual statistical analyses in a real-world setting using a federated approach. *PLoS one*. 2024 Nov 14;19(11):e0302007. 'HL7 FHIR. Resource concept map. <https://www.hl7.org/fhir/terminology.html>. Accessed June 5, 2025.



Amélie Lambert¹, Claire Castagne¹, David Pau¹, Jacek Chmiel¹, Alberto Labarga¹, Eric Boernert², François Margraff², Camille Bachot¹, Lukasz Kaczmarek², Dimitar Toshev², Judith Martinez Gonzalez², Thomas Stone¹

¹Roche SAS, Boulogne-Billancourt, France. ²F. Hoffmann-La Roche AG, Basel, Switzerland. ³Roche Polska Sp. z o.o., Warsaw, Poland. ⁴Roche Products Limited, Welwyn Garden City, United Kingdom. ⁵Aventa, Cologne, Germany. ⁶ALG, Pamplona, Spain. ⁷Billerium S.A., Warsaw, Poland.





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?



**The weekly OHDSI community call is held
every Tuesday at 11 am ET.**

Everybody is invited!

Links are sent out weekly and available at:
ohdsi.org/community-calls-2025