Workgroup Updates

OHDSI Community Call
Oct. 3, 2023 • 11 am ET

Upcoming Community Calls

Date	Topic
Oct. 3	Workgroup Reports, pt 1
Oct. 10	Workgroup Reports, pt 2
Oct. 17	Symposium Week! Final Logistics + Mad Minutes
Oct. 24	Welcome to OHDSI
Oct. 31	TBA
Nov. 7	Meet The Titans
Nov. 14	Collaborator Showcase Honorees

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

👏

Congratulations to the team of **Santiago Frid, Guillem Bracons Cucó, Jessyca Gil Rojas, Antonio López-Rueda, Xavier Pastor Duran, Olga Martínez-Sáez, and Raimundo Lozano-Rubí** on the publication of **Evaluation of OMOP CDM, i2b2 and ICGC ARGO for supporting data harmonization in a breast cancer use case of a multicentric European AI project** in *The Journal of Biomedical Informatics*.

Original Research

Evaluation of OMOP CDM, i2b2 and ICGC ARGO for supporting data harmonization in a breast cancer use case of a multicentric European AI project

Santiago Frid [a,*], Guillem Bracons Cucó [b], Jessyca Gil Rojas [a], Antonio López-Rueda [a,c], Xavier Pastor Duran [a], Olga Martínez-Sáez, MD [d], Raimundo Lozano-Rubí [d]

[a] Clinical Informatics Service, Hospital Clínic de Barcelona, Villarroel 170, 08036 Barcelona, Spain
[b] Fundació de Recerca Clínic Barcelona - Institut d'Investigacions Biomèdiques August Pi i Sunyer, Rosselló 149-153, 08036 Barcelona, Spain
[c] Radiology Service, Hospital Clínic de Barcelona, Villarroel 170, 08036 Barcelona, Spain
[d] Oncology Service, Hospital Clínic de Barcelona, Villarroel 170, 08036 Barcelona, Spain

ABSTRACT

Objective: Observational research in cancer poses great challenges regarding adequate data sharing and consolidation based on a homogeneous data semantic base. Common Data Models (CDMs) can help consolidate health data repositories from different institutions minimizing loss of meaning by organizing data into a standard structure. This study aims to evaluate the performance of the Observational Medical Outcomes Partnership (OMOP) CDM, Informatics for Integrating Biology & the Bedside (i2b2) and International Cancer Genome Consortium, Accelerating Research in Genomic Oncology (ICGC ARGO) for representing non-imaging data in a breast cancer use case of EuCanImage.
Methods: We used ontologies to represent metamodels of OMOP, i2b2, and ICGC ARGO and variables used in a cancer use case of a European AI project. We selected four evaluation criteria for the CDMs adapted from previous research: content coverage, simplicity, integration, implementability.
Results: i2b2 and OMOP exhibited higher element completeness (100% each) than ICGC ARGO (58.1%), while the three achieved 100% domain completeness. ICGC ARGO normalizes only one of our variables with a standard terminology, while i2b2 and OMOP use standardized vocabularies for all of them. In terms of simplicity, ICGC ARGO and i2b2 proved to be simpler both in terms of ontological model (276 and 175 elements, respectively) and in the queries (7 and 20 lines of code, respectively), while OMOP required a much more complex ontological model (615 elements) and queries similar to those of i2b2 (20 lines). Regarding implementability, OMOP had the highest number of mentions in articles in PubMed (130) and Google Scholar (1,810), ICGC ARGO had the highest number of updates to the CDM since 2020 (4), and i2b2 is the model with more tools specifically developed for the CDM (26).
Conclusion: ICGC ARGO proved to be rigid and very limited in the representation of oncologic concepts, while i2b2 and OMOP showed a very good performance. i2b2's lack of a common dictionary hinders its scalability, requiring sites that will share data to explicitly define a conceptual framework, and suggesting that OMOP and its

OHDSI Shoutouts! 👏

Congrats to our 2023 Titan Award Nominees!

Alexander Davydov · Aniek Markus · Anna Ostropolets · Anthony Sena · Asieh Golozar · Asiyah Lin · Atif Adam · Azza Shoaibi · Can Yin · Carlos Diaz · Center for Surgical Science team · Christian Reich · Christie Quarles · Chungsoo Kim · Cindy Cai · Clair Blacketer · Clark Evans · Craig Sachson · Cynthia Sung · Dana Zakrzewski · Danielle Boyce · Davera Gabriel · Debo Wei · Eleanor Davies · Elisse Katzman · Erica Voss · Evan Minty · Frank DeFalco · Geert Byttebier · Georgina Kennedy · Gowtham Rao · Grahame Grieve · Gregory Klebanov · Gyeol Song · Henrik John · Hugo Vernooij · IQVIA OMOP Productized Analytics · Ismail Gogenur · Jack Brewster · James Brash · James Gilbert · Jared Houghtaling · Jasmine Gratton · Jenna Reps · Jiawei Qian · Jiayi (Jessie) Tong · Jing Li · Joel Swerdel · John Gresh · Katherine Duszynski · Katy Sadowski · Kyle Zollo-Venecek · Kyrylo Simonov · LAISDAR Study Team · Lee Evans · Lydia Liu · Manlik Kwong · Marc Suchard · Marc Twagirumukiza · Marcel de Wilde · Masha Khitrun · Marti Catala · Martijn Schuemie · Martin Lavallee · Marty Alvarez · Meghan Pettine · Mengyuan Shang · Michael Matheny · Michelle Hribar · Milou Brand · Montse Camprubi · Nathan Buesgens · Nathan Hall · Nicole Pratt · Nigel Hughes · Nikolai Grewe · OHDSI Vocabulary Team · Oleg Zhuk · Paul Dougall · Paul Nagy · Polina Talapova · Raivo Kolde · Renske Los · Sally Baxter · Sarah Seager · Stephen Town · Tal El-Hay · Thamir Alshammary · Thomas Falconer · Timur Vakhitov · Varvara Savitskaya · Vipina Keloth · Xiaoyu Lin

Winners will be announced during the #OHDSI2023 Closing Talk!

#OHDSISocialShowcase

ohdsi.org/europe2023-showcase

#OHDSISocialShowcase

MONDAY

Privacy-preserving using k-anonymity and l-diversity in OMOP CDM databases

(João Rafael Almeida, José Luís Oliveira)

Privacy-preserving using k-anonymity and l-diversity in OMOP CDM databases

👤 João Almeida

INTRO:

- The official OMOP CDM guidelines can prevent the re-identification of patients stored in the databases.

- However, literature has shown that procedures that omitted key identifiers are not robust anonymisation procedures.

METHODS

- *k*-Anonymity limits the information released, based on generalisation and suppression of data concepts, as well as the number of repetitive elements.

- *l*-Diversity technique was proposed aiming to fill some gaps of the *k*-anonymity model.

- These techniques can be applied to the OMOP CDM schema requiring the characterisation of each field.

- When sharing a sample of the database, these techniques increase the anonymization levels.

RESULTS

- We only considered some tables of the OMOP CDM.

- Tables in the groups "Standardized vocabularies", "Standardized health economics", "Standardized derived elements" and "Standardized metadata" were omitted.

- Key attributes, quasi-identifies and sensitive attributes were mapped to apply the privacy-preserving techniques.

Sharing aggregate data?

Robust anonymization through privacy-preserving techniques.

The key attributes were defined as the unique identifiers directly associated with a patient. We considered all fields in the format "*_source_value" as sensitive attributes. All of the remaining fields were considered quasi-identifiers because these represent the equivalence classes in the view.

Scan **QR** to download the **full paper**

👤 João R. Almeida and José L. Oliveira

university of aveiro
theoria poiesis praxis   IEETA   OHDSI

#OHDSISocialShowcase

TUESDAY

Capture and consolidation of renal specific concepts into a cohesive OMOP dataset

(**Jared Houghtaling**, Jose Antonio Ramírez García, Clémence Le Cornec, Lore Vermeylen, Nir Assaraf, Lars Halvorsen)

Pre-coordinating mapping conventions for these data will enable efficient and effective collaboration on studies that have **potential to improve outcomes** for those individuals with rare kidney diseases

Title: Capture and consolidation of renal-specific concepts into a cohesive OMOP dataset

BACKGROUND: The European Rare Kidney Disease Reference Network (ERKNet) was formed by the European Union in 2017 and is one of 24 European Reference Networks (ERNs). ERKNet -- made up of 32 pediatric and 20 adult nephrology centres -- aims to improve the quality of patient management by: (1) educating healthcare professionals, (2) establishing best practices, (3) enabling virtual expert consultations for unusual cases, and (4) promoting clinical research activities. Patient registries are a critical component of these aims: beginning in 2017, ERKNet centres were active in more than 60 disease-specific registries with regional or national patient coverage or voluntary participation. At that time, however, no single registry was used across all ERKNet centres and few provided monitoring of relevant disease or treatment specific performance and outcome measures. In an attempt to improve the quantity and quality of this valuable registry information, ERKNet created the European Rare Kidney Disease Registry (ERKReg) [1]. ERKNet is now an active data partner in the European Health Data and Evidence Network (EHDEN) and is working together with edenceHealth to transform their rich registry dataset, ERKReg, into the Observational Medical Outcomes Partnership (OMOP) common data model (CDM). ERKReg is focused on rare kidney diseases, and, despite the low prevalence of these diseases, it represents more than 17.000 patients. This increased patient coverage in comparison with other registries is a direct benefit of being part of the ERKNet collaborative effort. Finally, ERKReg is characterised by its continuous longitudinal patient follow-up, which distinguishes it as unique and valuable resource for the study and treatment of rare kidney diseases.

Figure 1: Overview of three-machine installation, with one managed instance hosting the OMOP CDM database and ETL processes, one instance hosting the OHDSI tooling (via Docker) required to interact with the OMOP data, and a user workstation capable of connecting to both instances for updating/launching ETL processes and viewing web applications.
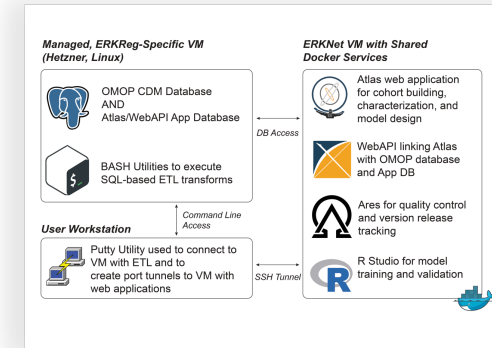
METHODOLOGY: We designed the Extract-Transform-Load (ETL) process in a way that is both flexible to accommodate different input files and datatypes, and simple to update. The ETL itself is executed via bash script and references an array of SQL queries that perform the necessary lookups and transformations in a PostgreSQL database. Critical to these processes is the semantic mapping file, which holds both the structural and conceptual logic necessary to convert source variables to associated domain-specific events with proper date references. In total, ERKReg mapped approximately 1500 unique source variables to standard concepts in OMOP CDM, capturing both events as well as the absence thereof. Mapping was done collaboratively by multiple team members within a shared and versioned spreadsheet, and coverage was monitored between versions using the OHDSI Ares application. The solution addressed several constraints related to data security and privacy, as well as resources available: (1) the OMOP CDM data needed to remain in Germany and (ideally) within a managed VM, (2) the managed VM selected could not support native Docker processes, and the proposed alternative uDocker was not sufficient to host OHDSI tooling, and (3) ERKNet had already provisioned a VM for other Docker processes, and that server had space available for OHDSI tooling.

LIMITATIONS AND DISCUSSION: ERKNet has transformed more than 10'000 unique patients with approximately 500'000 total records (300'000 Observations, 150'000 Measurements, 35'000 Conditions, 15'000 Procedures) in OMOP format. We have mapped more than 90% of the available source variables to OMOP standards, relying on the Participant Provided Information (PPI) vocabulary for many concepts related to family histories and prior medical records. We also took advantage of the recently released Orphanet-to-SNOMEDCT mappings to capture various elements related to rare kidney diseases [2]. One particular (and yet unmet) challenge we have faced with this data is proper linking of patients through familial relationships. Many of the diseases captured within the ERKReg source data have a genetic component, and we are interested in cataloguing any familial connections between patients within these registries. We are currently using Fact Relationship table, though due to its limited integration with OHDSI tooling we are also in the preliminary phases of designing a rare disease extension to the standard OMOP tables that could handle this type of information, along with other useful data such as patient reported outcomes (PROs) and detailed clinical assessments of disease progression, both of which contain valuable information on the overall quality of life, disease severity, response to treatments, and other relevant factors in the rare disease research environment.

Jared Houghtalinga, Jose Antonio Ramírez Garcíab, Clémence Le Cornecb,
Lore Vermeylena, Nir Assarafa and Lars Halvorsena

AFFILIATIONS
a edenceHealth NV (Kontich, Belgium)
b European Rare Kidney Disease Reference Network (ERKNet)

REFERENCES
[1] Bassanese, Giulia, et al. "The European rare Kidney disease registry (ERKReg): Objectives, design and initial results." Orphanet journal of rare diseases 16.1 (2021): 1-19.
[2] Health Terminology Standards Development Organisation, I. (n.d.). SNOMED CT to Orphanet Map Package Production Release Notes - July 2022. SNOMED Confluence. Retrieved April 6, 2023, from https://confluence.ihtsdotools.org/display/RMT/SNOMED+CT+to+Orphanet+Map+package+Production+Release+Notes+-+July+2022

#OHDSISocialShowcase

WEDNESDAY

Lessons learned from four population-based cancer registries: Mapping of ICD-O-3 codes to standard concepts

(Peter Prinsen, Maaike van Swieten, Chiara Attanasio, Espen Enerly, Siri Larønningen, Elisabetta Rapiti, David Marcic, Evelyne Fournier, Pierre Künzli, Michael Schnell, Sophie Couffignal, Claudine Backes)

The OMOP **ICDO3 vocabulary** is **incomplete.**
**Give us your missing codes!**

Mapping of ICD-O-3 codes to standard concepts:
Lessons learned from four population-based cancer registries

Background: ICD-O-3 contains lists of histology and topography codes. A complete list of conditions, i.e., valid combinations of a histology and a topography, does not exist. The OMOP ICDO3 vocabulary consists of the most commonly occurring conditions in a limited set of data sources: several conditions are missing from the vocabulary.

Result 1: A significant fraction of ICD-O-3 codes are missing from the OMOP ICDO3 vocabulary, but they represent only a small fraction of tumors in each registry.

Result 2: For certain cancers, this fraction is significant, potentially affecting studies on those cancers.

Result 3: Most missing codes occur in only one of the four registries. Therefore, to create a list of missing codes that is as complete as possible, we need missing codes from as many sources as possible.
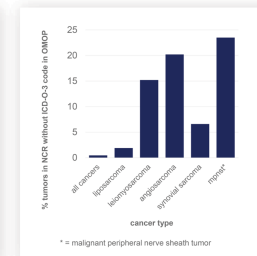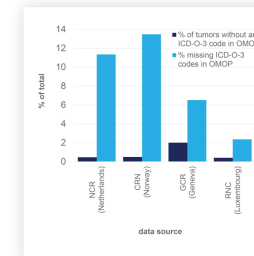
Methods

We compared the ICD-O-3 codes for invasive cancers in the OMOP ICDO3 vocabulary with codes used in four population-based cancer registries from Geneva (Geneva Cancer Registry (GCR)), Luxembourg (Registre National du Cancer (RNC)), The Netherlands (Netherlands Cancer Registry (NCR)) and Norway (Cancer Registry of Norway (CRN)).

Limitation: We believe that with enough data sources we can capture most of the missing ICD-O-3 codes and make the OMOP ICDO3 vocabulary sufficiently complete. How many data sources are needed must be determined.

@OHDSI   www.ohdsi.org   #JoinTheJourney   ohdsi

#OHDSISocialShowcase

THURSDAY

Automated Retrospective Data Extraction from Electronic Health Records Using Natural Language Processing to Identify Transthyretin Amyloid Cardiomyopathy in a Real-World Heart Failure Population
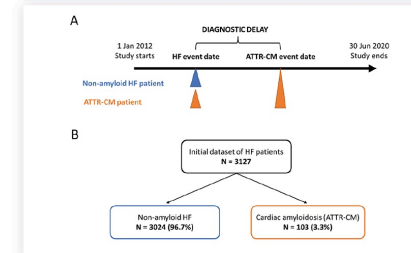
(**Dries Hens**, Ana Moya, Clara L. Oeste, Monika Beles, Sofie Verstreken, Riet Dierckx, Ward Heggermont, Jozef Bartunek, Eline Bogaerts, Imke Masuy, Marc Vanderheyden)

Detection of aTTR-CM in a real-world heart failure population by NLP-guided automated data extraction from EHRs

Automated Retrospective Data Extraction from Electronic Health Records Using Natural Language Processing to Identify Transthyretin Amyloid Cardiomyopathy in a Real-World Heart Failure Population

Background: ATTR-CM is a fatal and progressive illness that frequently has delays in diagnosis and treatment. The objective of this study is to analyze individuals with ATTR-CM in a real-world population of heart failure (HF) patients and pinpoint single and combined ATTR-CM predictive factors in this European population.

Result 1: Out of 3127 HF patients, there were 103 HF patients diagnosed with ATTR-CM. This data was analyzed to detect phenotypes of patients at risk for ATTR-CM. The mean (±SD) diagnostic delay between HF and ATTR-CM was 1.8 (±1.6) years.

Result 2: The strongest cardiac predictor was atrial fibrillation, whereas the strongest non-cardiac predictor was carpal tunnel syndrome.
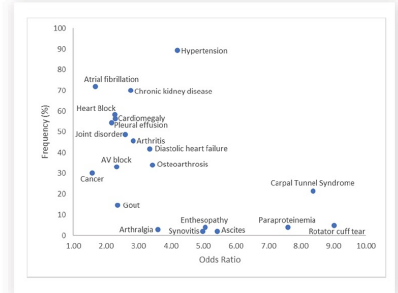
Methods
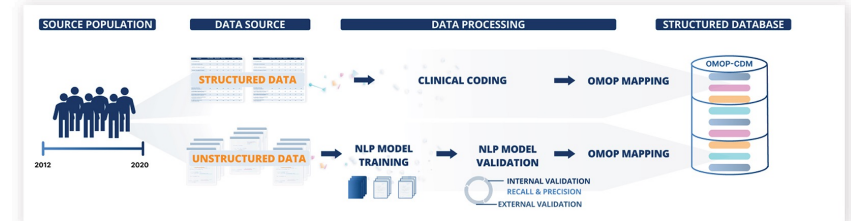
Conclusion: With the growing adoption of EHRs, vast databases containing both structured and unstructured clinical data are being produced. They offer a sizable possibility to create algorithms for automatically interpreting data and assisting medical professionals in the diagnosis and treatment of patients.

Dries Hens, Ana Moya, Clara L. Oeste, Monika Beles, Sofie Verstreken, Riet Dierckx, Ward Heggermont, Jozef Bartunek, Eline Bogaerts, Imke Masuy, Marc Vanderheyden

Hartcentrum OLV Aalst   LYNXCARE

#OHDSISocialShowcase

FRIDAY

FinnGen ETL using BigQuery SQL

(Shanmukha Sampath Padmanabhuni, Javier Gracia-Tabuenca, Mary Pat Reeve)

FinnGen ETL using BigQuery SQL

Background: In addition to 500,000 samples of genotype data, FinnGen continuously expanding its register base phenotype data. We developed a process to transform current and incoming registers to the OMOP-CDM.

Methods
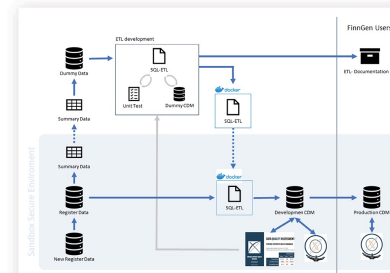
1. Custom R scripts to summarize the source data to be download from the safe environment.
2. Summary data is used to generate dummy data.
3. ETL is develop on the dummy data and unit tested.
4. Up to date documentation is generated and published.
5. SQL-ETL and support scripts are dockerized and uploaded into the safe environment.
6. SQL-ETL is applied to the source data to generate a development CDM instance.
7. Development CDM instance is tested on DQD and new features tested in Atlas.
8. When ready the development CDM is copied into production to be utilized by the users.

Results

- Six out of 17 national registries converted.
- 73 unit tests passed covering 10 CDM tables.
- Used stem and mapping tables.
- ETL ran within an hour.
- Utilized DQD to fix most of Plausibility and Conformance problems.
- Scan the QR code to explore the ETL documentation.
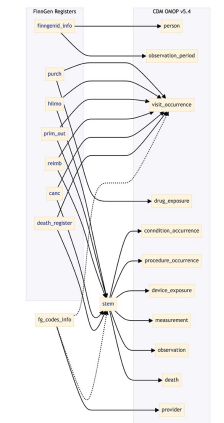
ETL: source and custom stem tables

Limitation: There are quite good number of unmapped codes in the vocabularies. Dummy data generator does not consider sex variable while assigning codes.

Shanmukha Sampath Padmanabhuni
Javier Gracia-Tabuenca, Mary Pat Reeve

OHDSI Shoutouts!

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

Do you have anything you want to share? Please send to sachson@ohdsi.org so we can highlight during this call and on our social channels.
Let's work together to promote the collaborative work happening in OHDSI!

Three Stages of The Journey

Where Have We Been?
Where Are We Now?
Where Are We Going?

the journey

Upcoming Workgroup Calls

Date	Time (ET)	Meeting
Tuesday	12 pm	Common Data Model Vocabulary Subgroup
Wednesday	2 am	Methods Research
Wednesday	7 am	Medical Imaging
Wednesday	8 am	Psychiatry
Thursday	9:30 am	Themis
Thursday	12 pm	Methods Research
Thursday	1 pm	OMOP CDM Oncology Vocabulary Development Subgroup
Thursday	7 pm	Dentistry
Friday	9 am	GIS – Geographic Information System General
Friday	11 am	Clinical Trials
Monday	10 am	Healthcare Systems Interest Group
Monday	11 am	Early-Stage Researchers
Monday	4 pm	Eyecare & Vision Research
Monday	6 pm	OMOP & FHIR
Tuesday	9 am	OMOP CDM Oncology Genomic Subgroup

Spotlight: Jody-Ann McLeggon

Jody-Ann McLeggon is a program manager at the Columbia University Irving Medical Center. She focuses specifically on the collaboration between OHDSI and the U.S. Food & Drug Administration CBER Biologics Effectiveness and Safety (BEST) Community Engagement and Development Initiative. She also took a leadership role in leading the global community on the development and execution of four network studies in the 2023 Save Our Sisyphus Challenge.

Jody-Ann spent more than five years at Northwell Health as a research supervisor and coordinator. She managed project portfolios for multisite research projects and oversaw research activities and development within the department. She also managed the implementation of NIH studies, clinical trials, and investigator-initiated studies.

She discusses the collaboration between OHDSI and the FDA, how OHDSI can assist regulator agencies in the active surveillance of vaccines and drugs, lessons she learned about network studies, and more in the latest edition of the Collaborator Spotlight.

What drew you to a career in research and what have been a couple exciting projects you have worked on?

My journey in research started out as an undergraduate student working in a cell biology lab at the beginning of my sophomore year investigating the function of Lark protein in Drosophila melanogaster (fruit flies). Since then, my interest in research has only evolved, and I've transferred my knowledge and skills from my wet lab experience to public health. Currently, I've been able to participate in many OHDSI initiatives. The most recent one was the Save Our Sisyphus (SOS) Challenge, where we have led the community through all aspects of conducting research study, from initial idea through design and dissemination. I also support our FDA initiative on methods development research for vaccine safety monitoring. I'm also a member of the LEGEND initiative led by Marc Suchard working on LEGEND T2DM.

ohdsi.org/spotlight-jody-ann-mcleggon

October Newsletter Is Available

The Journey Newsletter (October 2023)

It is officially Symposium Month! The 2023 OHDSI Global Symposium will be held Oct. 20-22 at the Hilton East Brunswick Hotel & Executive Meeting Center in East Brunswick, N.J., and there is still time to register. This newsletter takes a closer look at the symposium weekend, including the three-day collaborator showcase, while also reflecting on the growth of the DARWIN EU® Initiative. We also celebrate the 90+ Titan Award nominees! #JoinTheJourney

OHDSI Videocast: Symposium Showcase, DARWIN EU®, Titans & more

In the latest On The Journey video, Patrick Ryan and Craig Sachson take a deeper look at the expanded Collaborator Showcase for the 2023 OHDSI Global Symposium main conference. They discuss the 90+ Titan Award nominees who were nominated by members of the community, and they reflect on two European initiatives that were highlighted during September community calls, DARWIN EU® and the 11th Revision of the ENCePP Guide on Methodological Standards in Pharmacoepidemiology. (if video does not appear, click 'View this email in your browser')

Community Updates

Where Have We Been?
• The Sept. 5 OHDSI Community Call provided a look at the DARWIN EU® Progress and Roadmap. DARWIN EU® is an initiative that works to deliver real-world evidence from across Europe on diseases, populations and the uses and performance of medicines. Members of the leadership team provided an update on both its progress and goals.
• ENCePP editor Catherine Cohet and seven authors of the 11th Revision of the ENCePP Guide on Methodological Standards in Pharmacoepidemiology led a discussion on the purpose and audience of the guide, some of the specific focuses within this revision, and how it can benefit the research mission of the OHDSI community.

Where Are We Now?
• After months of preparation, the OHDSI Global Symposium is less than three weeks away. The Global Symposium will be held Oct. 20-22 at the Hilton East Brunswick Hotel & Executive Meeting Center in East Brunswick, N.J. The full weekend agenda is now available, and it includes the Friday conference agenda, descriptions on the various weekend activities, and all 137 posters and 24 software demos that will be presented during the collaborator showcase.
• Congratulations to the 93 individuals/teams who were nominated for a 2023 Titan Award. Each year, Titan Awards are presented to members of the community who have made valuable contributions to OHDSI's mission, vision and values. The recipients of the seven Titan Awards will be announced during the closing talk on Friday, Oct. 20 at the Global Symposium.

Where Are We Going?
• We are going to East Brunswick, of course, and we hope to see you there. We are still accepting registrations for the Global Symposium, but don't wait too long to secure your spot at the highlight event of the OHDSI year! Use this link to register for #OHDSI2023!
• The first two weeks of OHDSI community calls this month (Tuesdays, 11 am ET) will feature updates from our various workgroups, and they will help set the foundation for our symposium weekend activities. Community calls are open to everybody, and they are a great way to stay updated on everything happening within the community. The meeting link is available within our Community Calls page.

#OHDSI2023 Month is Here!
Check Out The Full Weekend Agenda For The Oct. 20-22 Global Symposium

The agenda for the 2023 Global Symposium main conference is now available and it highlights the most diverse agenda in our event history. It lists the full schedule for all three days, descriptions on the various weekend activities, and all 137 posters and 24 software demos that will be presented during the collaborator showcase.

- **Register for the 2023 Global Symposium**
- **2023 Global Symposium Weekend Agenda**
- **Collaborator Showcase Posters**
- **Collaborator Showcase Software Demos**
- **Collaborator Showcase Lightning Talks**

September Publications

Przysucha M, Hüsers J, Liberman D, Kersten O, Schlüter A, Fraas S, Busch D, Moelleken M, Erfurt-Berge C, Dissemond J, Hübner U. Design and Implementation of an ETL-Process to Transfer Wound-Related Data into a Standardized Common Data Model. Stud Health Technol Inform. 2023 Sep 12;307:258-266. doi: 10.3233/SHTI230723. PMID: 37697861.

de Groot R, Püttmann DP, Fleuren LM, Thoral PJ, Elbers PWG, de Keizer NF, Cornet R; Dutch ICU Data Sharing Against COVID-19 Collaborators. Determining and assessing characteristics of data element names impacting the performance of annotation using Usagi. Int J Med Inform. 2023 Aug 29;178:105200. doi: 10.1016/j.ijmedinf.2023.105200. Epub ahead of print. PMID: 37703800.

Lee S, Shin H, Choe S, Kang MG, Kim SH, Kang DY, Kim JH. MetaLAB-HOI: Template standardization of health outcomes enable massive and accurate detection of adverse drug reactions from electronic health records. Pharmacoepidemiol Drug Saf. 2023 Sep 14. doi: 10.1002/pds.5694. Epub ahead of print. PMID: 37710363.

Raventós B, Fernández-Bertolín S, Aragón M, Voss EA, Blacketer C, Méndez-Boo L, Recalde M, Roel E, Pistillo A, Reyes C, van Sandijk S, Halvorsen L, Rijnbeek PR, Burn E, Duarte-Salles T. Transforming the Information System for Research in Primary Care (SIDIAP) in Catalonia to the OMOP Common Data Model and Its Use for COVID-19 Research. Clin Epidemiol. 2023 Sep 13;15:969-986. doi: 10.2147/CLEP.S419481. PMID: 37724311; PMCID: PMC10505380.

Moya A, Oeste CL, Beles M, Verstreken S, Dierckx R, Heggermont W, Bartunek J, Bogaerts E, Masuy I, Hens D, Bertolone D, Vanderheyden M. Detection of transthyretin amyloid cardiomyopathy by automated data extraction from electronic health records. ESC Heart Fail. 2023 Sep 19. doi: 10.1002/ehf2.14517. Epub ahead of print. PMID: 37726928.

Matcham F, Simblett SK, Leightley D, Dalby M, Siddi S, Haro JM, Lamers F, Penninx BWHJ, Bruce S, Nica R, Zormpas S, Gilpin G, White KM, Oetzmann C, Annas P, Brasen JC, Narayan VA, Hotopf M, Wykes T; RADAR-CNS consortium. The association between persistent cognitive difficulties and depression and functional outcomes in people with major depressive disorder. Psychol Med. 2023 Oct;53(13):6334-6344. doi: 10.1017/S0033291722003671.

mailchi.mp/ohdsi/october2023

October Newsletter Is Available

mailchi.mp/ohdsi/october2023

Global Symposium

Oct. 20-22 • East Brunswick, NJ, USA
Hilton East Brunswick Hotel & Executive Meeting Center

bit.ly/OHDSI2023Registration

@OHDSI   www.ohdsi.org   #JoinTheJourney   ohdsi

Global Symposium Conference Agenda

Time	Topic
7:30 - 8:30 am East Brunswick Room + Grand Ballroom Foyer	**Symposium Registration, Lite Breakfast Buffet, All-Day Exhibits**
8:30 - 9:30 am Grand Ballroom	**State of the Community** OHDSI: Where have we been? Where are we going? George Hripcsak, Columbia Univ. **Community Highlights:** • OMOP CDM users and the OHDSI data network Clair Blacketer, Johnson & Johnson • OHDSI standardized vocabularies Alexander Davydov, Odysseus Data Services • OHDSI's open-source community Katy Sadowski, Boehringer Ingelheim • OHDSI Europe 2024 Peter Rijnbeek, Erasmus MC • OHDSI Asia-Pacific 2024 Mengling Feng, National Univ. of Singapore
9:30 - 10:30 am Grand Ballroom	**OHDSI Community Networking** **Moderators:** • Faaizah Arshad, Univ. of California-Los Angeles • Cynthia Sung, Duke-NUS Medical School
10:30 am - 12:00 pm Grand Ballroom	**Plenary: Improving the reliability and scale of case validation** **Presenters:** • Patrick Ryan, Johnson & Johnson, Columbia Univ. • Anna Ostropolets, Odysseus Data Services • Martijn Schuemie, Johnson & Johnson, Univ. of California-Los Angeles
12:00 pm - 1:00 pm Grand Ballroom Foyer	**Buffet Lunch**

All events take place at the Grand Ballroom Level • Exhibits will be available throughout the day

Time	Topic
1:00 pm - 2:00 pm Grand Ballroom	**Panel: Lessons learned from OHDSI network studies** **Presenters:** • Insights from LEGEND-T2DM Marc Suchard, Univ. of California-Los Angeles • Intravitreal anti-VEGF and risk of kidney failure: A Sisyphus Challenge Study Cindy X Cai, Johns Hopkins Univ. • Fluoroquinolones and the risk of aortic aneurysm: A Sisyphus Challenge study Seng Chan You, Yonsei Univ. • Lessons learned applying the Strategus framework across the OHDSI network Anthony Sena, Johnson & Johnson **Moderator:** Sarah Seager, IQVIA
2:00 pm - 2:45 pm Grand Ballroom	**Collaborator Showcase, Lightning Talk Session #1:** **Data Standards and Methods Research** • Mapping of Critical Care EHR Flowsheet data to the OMOP CDM via SSSOM Polina Talapova, SciForce • Paving the way to estimate daily dose in OMOP CDM for Drug Utilisation Studies in DARWIN EU® Theresa Burkard, Univ. of Oxford • Generating Synthetic Electronic Health Records in OMOP using GPT Chao Pang, Columbia Univ. • Comparing concepts extracted from clinical Dutch text to conditions in the structured data Tom Seinen, Erasmus MC • Finding a constrained number of predictor phenotypes for multiple outcome prediction Jenna Reps, Johnson & Johnson **Moderator:** Davera Gabriel, Johns Hopkins University
2:45 pm - 3:30 pm Grand Ballroom	**Collaborator Showcase, Poster / Demo Session #1** **Poster walk leads:** • Data standards: Mui Van Zandt, IQVIA • Methods research: Christophe Lambert, Univ. of New Mexico • Open-source development: Paul Nagy, Johns Hopkins Univ. • Clinical applications: Kristin Kostka, Northeastern University

All events take place at the Grand Ballroom Level • Exhibits will be available throughout the day

Time	Topic
3:30 - 4:15 pm Grand Ballroom	**Collaborator Showcase, Lightning Talk Session #2:** **Methods Research and Clinical Applications** • Synthesizing Evidence for Rare Events: a Novel Zero-Inflated Bivariate Model to Integrate Studies with Double-Zero Outcomes Lu Li, Univ. of Pennsylvania • Active Safety Surveillance Using Real-world Evidence (ASSURE): An application of the Strategus package Kevin Haynes, Johnson & Johnson • Patient's outcomes after endoscopic retrograde cholangiopancreatography (ERCP) using reprocessed duodenoscope: a descriptive study using real-world data Jessica Maruyama, Precision Data • Quantification of Racial Differences in Post-acute Sequelae of SARS-CoV-2 Infection (PASC) in Children: an EHR-Based Cohort from the RECOVER Program Bingyu Zhang, Univ. of Pennsylvania • Eye Care and Vision Research Workgroup: First Year Update Michelle Hribar, National Institutes of Health – National Eye Institute **Moderator:** Atif Adam, IQVIA
4:15 - 5:00 pm Grand Ballroom	**Collaborator Showcase, Poster / Demo Session #2** **Poster walk leads:** • Data standards: Melanie Philofsky, Odysseus Data Services • Methods research: Andrew Williams, Tufts Univ. • Open-source development: Nsikak Akpakpan, Accenture • Clinical applications: Hanieh Razzaghi, Childrens Hospital of Pennsylvania
5:00 pm - 6:00 pm Grand Ballroom	**Closing session: Scaling community, scaling collaboration** • Titan Awards • Group Photo **Presenter** Patrick Ryan, Johnson & Johnson, Columbia Univ.
6:00 pm - 7:00 pm East Brunswick Room Grand Ballroom Foyer	**Networking Reception and Exhibits**
7:00 pm - 8:00 pm Grand Ballroom	**OHDSI Got Talent!**

bit.ly/OHDSI2023-Agenda

Register →

Global Symposium Conference Agenda

Agenda • Saturday, Oct. 21

Time	Topic
7:00 - 8:00 am Grand Ballroom Foyer	Lite Breakfast Buffet, All-Day Exhibits
8:00 am - 12:00 pm Various rooms	Introduction to OHDSI Tutorial Common Data Model/Network Data Quality WG Meeting Health Analytics Data-to-Evidence Suite (HADES) Hackathon Health EquityWG Meeting Medical Imaging WG Meeting Natural Language Processing WG Meeting OHDSI Industry WG Kickoff Meeting Oncology WG Meeting Phenotype Development & Evaluation WG Meeting Pregnancy and Reproductive Health Group (PRHeG) WG Meeting
12:00 - 1:00 pm Ballroom Foyer/ Ballroom	Lunch Buffet, Collaborator Showcase, All-Day Exhibits
1:00 pm - 5:00 pm Grand Ballroom	HowOften Large-Scale Characterization Workshop
5:00 pm	Free Time

Agenda • Sunday, Oct. 21

Time	Topic
7:00 - 8:00 am Grand Ballroom Foyer	Lite Breakfast Buffet, All-Day Exhibits
8:00 am - 12:00 pm Grand Ballroom	HowOften Large-Scale Characterization Workshop
12:00 - 1:00 pm Ballroom Foyer/ Ballroom	Lunch Buffet, Collaborator Showcase, All-Day Exhibits
1:00 pm - 5:00 pm Various Rooms	Africa Chapter Workshop Eye Care & Vision Research WG Meeting Health Analytics Data-to-Evidence Suite (HADES) Hackathon Healthcare Systems Interest Group (HSIG) WG Meeting HL7 FHIR-OMOP Connectathon ISPE RWE for Pharmacovigilance Medical Devices WG Meeting Psychiatry WG Meeting Vocabulary WG Meeting
5:00 pm	Symposium Closing

bit.ly/OHDSI2023-Agenda

Register ➡

@OHDSI   www.ohdsi.org   #JoinTheJourney   ohdsi

Welcome, 1ˢᵗ-Time Attendees!

All OHDSI first-time attendees are welcome to attend an orientation on Friday at 7:45 am within the Woodbridge/ Piscataway room. **Paul Nagy**, a 2022 Titan honoree for community leadership, will lead this session.

OHDSI

1ˢᵗ Time Attendees

AFFIX BLUE STICKER

Register →

ohdsi.org/ohdsi2023

Openings at Boehringer Ingelheim

Opening: Postdoctoral Associate/Data Analyst

Job Announcement: Postdoctoral Associate/Data Analyst - LEGEND Hypertension Project

Position: Postdoctoral Associate/Data Analyst

Organization: Yale University, School of Medicine

Location: 195 Church Street, 5th floor, New Haven, CT, 06510

Application Deadline: Rolling basis

Job Description:

We are seeking a talented and dedicated Postdoctoral Associate/Data Analyst to join our dynamic team. In this role, you will play a pivotal part in advancing our mission of improving health outcomes through data-driven research. You will have the opportunity to work with diverse healthcare datasets, develop innovative analytical methods, and collaborate with experts in the field.

The Postdoctoral Associate/Data Analyst should possess significant experience in R and Rstudio, with specific expertise in database management using PostgreSQL—critical requirements within the OHDSI network. Your responsibilities will include assisting the Principal Investigator (Dr. Yuan Lu from Yale University) and Co-Investigator (Drs. Marc Suchard from UCLA) in creating the analytic tool stack and performing related analyses.

Key Responsibilities:

- Collaborate with multidisciplinary teams to design and execute data analysis projects.
- Develop and implement statistical and machine learning models for healthcare data.
- Perform data extraction and preprocessing tasks to prepare datasets for analysis.
- Conduct exploratory data analysis and visualization to extract insights from healthcare data.
- Assist in the development and maintenance of OHDSI's open-source tools and resources.
- Communicate findings and insights through reports, presentations, and publications.
- Stay up-to-date with the latest advancements in data science and healthcare informatics.

Email: y.lu@yale.edu

Where Are We Going?

Any other announcements of upcoming work, events, deadlines, etc?

the journey

OHDSI Workgroup Objectives and Key Results (OKR)

NLP Workgroup

Hua Xu

NLP 2023 Objectives and Key Results

Objective 1: Knowledge dissemination – Contribute a chapter on NLP in the Book of OHDSI

Key results

1. Deliver the initial draft of the chapter; Timeline: 1Q2024

Objective 2: ETL for textual data representation and normalization

Key results

1. Revised Note_NLP proposal; Timeline: 1-2Q2023

2. Proof-of-concept study (Poster accepted); Timeline: 3-4Q2023

Objective 2: Conduct multi-site clinical studies that utilize both structured and textual data

Key results

1. T2DM-SDoH-NLP study; Timeline: 2-4Q2023, 1-4Q2024

2. Oncology NLP study; Timeline: 2-4Q2023, 1-4Q2024

3. Psychiatry NLP study; Timeline: 2-4Q2023, 1-4Q2024

Latin America Workgroup Objectives and Key Results (OKR)

Jose Posada

Steering Workgroup
co-leads: Patrick Ryan, George Hripcsak

Purpose: Steering WG exists to support the community and its leaders in collaboratively generating the evidence that promotes better health decisions and better care, by identifying, organizing, and guiding collaborative activities, facilitating communications across the community, providing input to operations of the OHDSI Central Coordinating Center, and building consensus on the vision for where the OHDSI community should go together.

Oct2023 update:

Objective 1: Empower workgroups to contribute to collaboratively generating the evidence that promotes better health decisions and better care

> Key accomplishments:
> 1. Workgroup Leader Summits convened to ensure appropriate communication across workgroups

Objective 2: Create collaboration activities that encourage collaborative generation and dissemination of the evidence that promotes better health decisions and better care

> Key accomplishments:
> 1. OHDSI2023 Global Symposium!
> 2. Community activities: 1- Phenotype Phebruary; 2- Sisyphus Challenge

Objective 3: Draft a strategy for a general purpose OHDSI asset library

2023 OHDSI Health Equity WG: Symposium Session Highlights

Date: 11th October 2023
Time: 8am -12pm

In-depth Research Showcases: Delve into rigorous studies as our WG members unveil critical insights on global health equity challenges.

Expert Panel Discussion: "From Policy to Practice" – Join thought leaders as they elucidate the practicalities of integrating SDoH factors into real-world scenarios.

Hands-on Exploration: Engage directly with the innovative Health Equity Explorer toolkit and witness its capabilities first-hand.

Project Prospectus: Anticipate the future as ambitious WG members pitch their prospective projects for 2024!

OHDSI Industry Work Group

Foster a stronger collaboration between the life science, pharma, and biotech industries, and the OHDSI community.

Identify and develop strategies to encourage the active participation of these industries in OHDSI studies and initiatives.

Facilitate knowledge transfer, sharing industry expertise and learnings with the broader OHDSI community.

Identify opportunities for mutual support, leveraging industry resources and capabilities to advance OHDSI's goals.

Increase the visibility and understanding of OHDSI's initiatives within these industries, promoting active involvement and commitment.

Industry WG Symposium Meeting

Meeting Agenda:

- Introductions
- Industry Working Group Objectives
- Roundtable on canvassing those interests
- Formalise structure of the working group moving forwards

Who should attend?

This is an open meeting with a focus on those members of the OHDSI community who have ties and affiliations with the Pharma and Biotech industries and would like to work together to represent those interests more broadly within OHDSI.

Pregnancy and Reproductive Health Group (PRHeG)
Objectives and Key Results (OKR)

Alison Callahan

Registry Workgroup
Objectives and Key Results (OKR)

Tina Parciak

Eye Care and Vision Research Workgroup

- Accomplishments
  - SOS Challenge Anti-VEGF and kidney injury study
  - Phenotypes submitted to How Often (Eye disease, kidney injury, uveitis)
  - Submitted missing terms to SNOMED and LOINC
  - Active subgroups in Retina, Glaucoma, Uveitis, Pediatrics
  - Publication about mapping, more to come

- Upcoming
  - Starting new subgroups supporting imaging integration & ETL development
  - Pilot eye exam concepts (visual acuity, IOP, refraction) at 4-5 sites
    - Pilot imaging, too?
  - Increase visibility of work

Join us in person at the OHDSI Symposium 10/22 1-5 pm as we plan imaging integration work!

Common Data Model
Objectives and Key Results (OKR)

Clair Blacketer

Network Data Quality
Objectives and Key Results (OKR)

Clair Blacketer

OMOP+ FHIR 2023 WG Progress

OKR1: Consolidate Prior Work

- Consolidated WG Meetings
  - Mondays Weekly 6pm EASTERN (June 12 - October 1)
- Reviewed Prior Subgroups Work
- Obtained FHIR IG GitHub from HL7 FHIR Management Group
- *Next Steps*
  - **Review results from October Connectathon**
  - **Populate FHIR IG GitHub with draft content**

OMOP+ FHIR 2023 WG Progress

OKR2: OHDSI Vocabularies FHIR Terminology Server Specification

- Design for OMOP CodeSystem & ConceptMap FHIR resources: Pass vote
  - HL7 Terminology Services Management Group
  - HL7 Terminology Infrastructure Working Group

- HL7 Terminology Authority Record - Pass vote

- Identified FHIR API Host for OMOP Vocabularies

- ***Next Steps***
  - **OHDSI Vocab Relationships ---> FHIR Terminology Server Operations**
  - **"Using OHDSI Vocabularies on FHIR" page for terminology.HL7.org (THO)**

OMOP+ FHIR 2023 WG Progress

OKR3: Hands-on Activity Supporting Specification Generation

- Sunday Session at Global Symposium
- Crowdsourced transformation experience
- Persistent Sample / Transform Records
- FHIR overview & demonstration

- **Next Steps**
  - **Gap & Overlap Analysis**
  - **Generate Draft Specification Content**
  - **Follow-up Connectathon Event (2024)**

OHDSI 2023 Global Symposium
OMOP + FHIR Connectathon Workflow