# Clinical Registry Efforts in OHDSI

## OHDSI Community Call

## Sept. 13, 2022 • 11 am ET

# Upcoming OHDSI Community Calls

| Date | Topic |
|---|---|
| Sept. 20 | 2022 OHDSI Symposium Preview |
| Sept. 27 | HTA Challenge |
| Oct. 4 | OHDSI Debates |
| Oct. 11 | Final OHDSI2022 Logistics |
| Oct. 18 | Welcome To OHDSI |
| Oct. 25 | Future Directions For OHDSI |

# Upcoming OHDSI Community Calls

| Date | Topic |
|---|---|
| Sept. 20 | 2022 OHDSI Symposium Preview |
| Sept. 27 | HTA Challenge |
| Oct. 4 | OHDSI Debates |
| Oct. 11 | Final OHDSI2022 Logistics |
| Oct. 18 | Welcome To OHDSI |
| Oct. 25 | Future Directions For OHDSI |

# Three Stages of The Journey

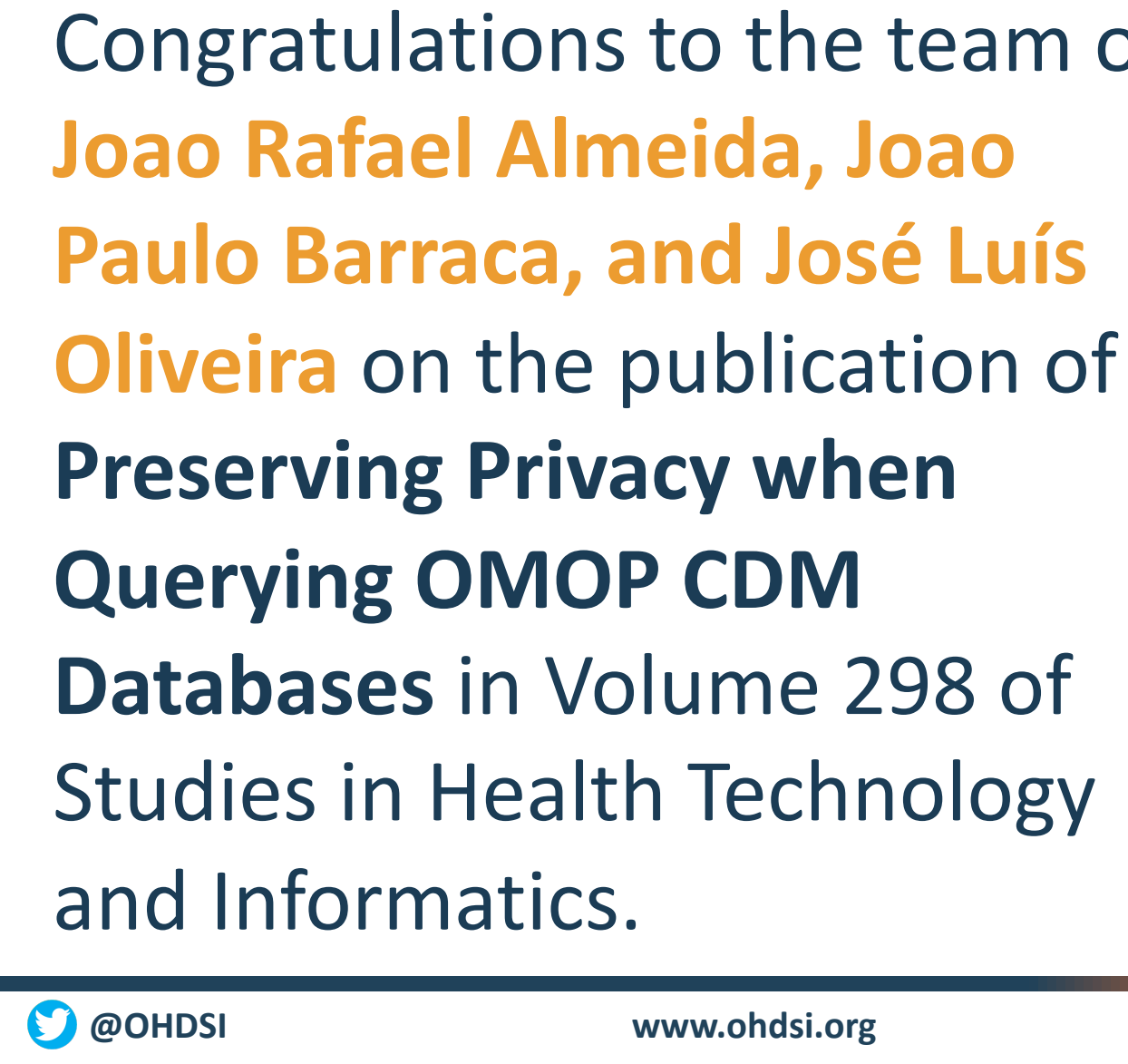
## Where Have We Been?

## Where Are We Now?

## Where Are We Going?

# OHDSI Shoutouts!

Congratulations to the team of **Joao Rafael Almeida, Joao Paulo Barraca, and José Luís Oliveira** on the publication of **Preserving Privacy when Querying OMOP CDM Databases** in Volume 298 of Studies in Health Technology and Informatics.



*Digital Professionalism in Health and Care: Developing the Workforce, Building the Future*
*P. Scott et al. (Eds.)*


*doi:10.3233/SHTI220930*

## Preserving Privacy when Querying OMOP CDM Databases

Joao Rafael ALMEIDA[a,c,1], Joao Paulo BARRACA[b] and José Luís OLIVEIRA[a]
[a]*DETI/IEETA, University of Aveiro, Portugal*
[b]*IT, DETI, University of Aveiro, Portugal*
[c]*Department of Computation, University of A Coruña, Spain*

**Abstract.** Anonymisation is currently one of the biggest challenges when sharing sensitive personal information. Its importance depends largely on the application domain, but when dealing with health information, this becomes a more serious issue. A simpler approach to avoid inadequate disclosure is to ensure that all data that can be associated directly with an individual is removed from the original dataset. However, some studies have shown that simple anonymisation procedures can sometimes be reverted using specific patients' characteristics. In this work, we propose a secure architecture to share information from distributed databases without compromising the subjects' privacy. The anonymiser system was validated using the OMOP CDM data schema, which is widely adopted in observational research studies.

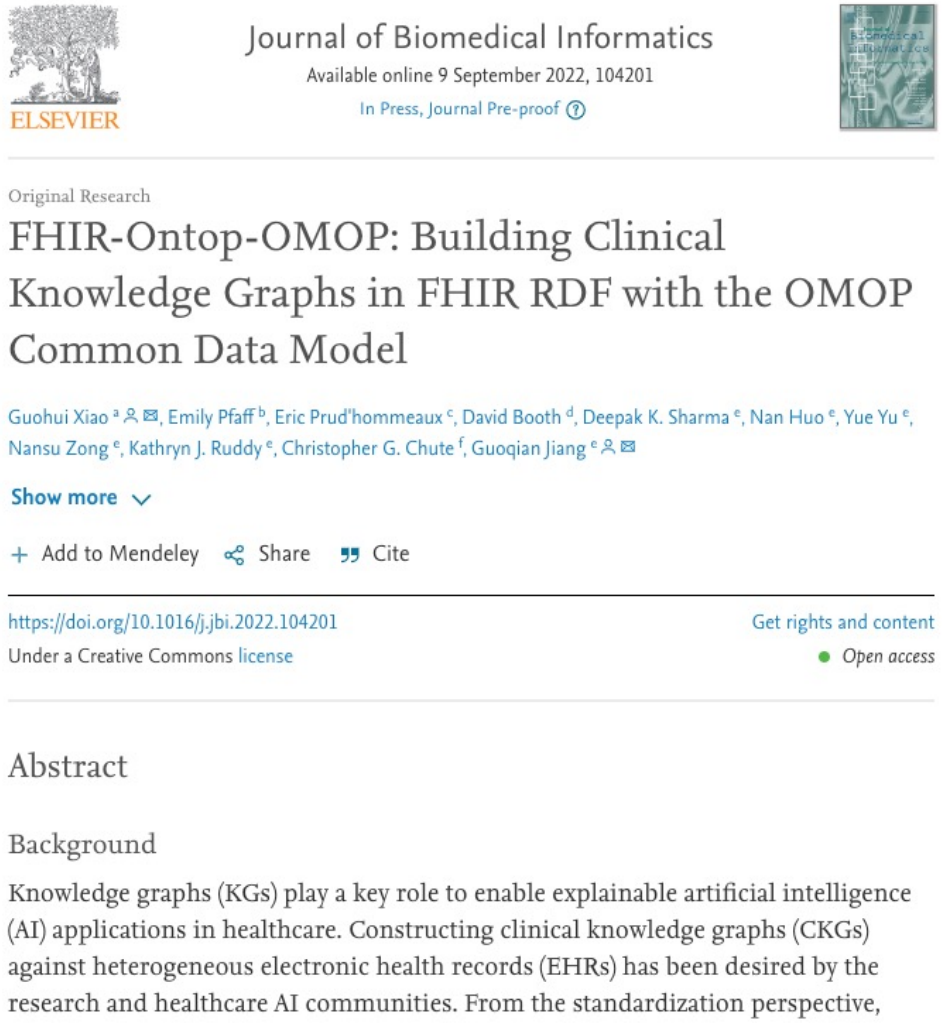
**Keywords.** Privacy preserving, Data anonymisation, k-Anonymity, l-Diversity, OMOP CDM, OHDSI

### 1. Introduction

Current approaches followed when sharing clinical data, simply try to avoid releasing sensitive information when publishing datasets. Their contents are often anonymized through processes that modify the original data, using data transformations that hide or remove subjects' identities, without degrading the data utility. However, anonymization simply based on the users' identity is limited, and there are still relevant challenges related to the privacy preservation of published data, namely on how to ensure the protection of data with smaller datasets, or on how to ensure resilience against future privacy threats [1].

# OHDSI Shoutouts!

Congratulations to the team of **Guohui Xiao, Emily Pfaff, Eric Prud'hommeaux, David Booth, Deepak K Sharma, Nan Huo, Yue Yu, Nansu Zong, Kathryn J Ruddy, Christopher Chute, and Guoqian Jiang** on the publication of **FHIR-Ontop-OMOP: Building Clinical Knowledge Graphs in FHIR RDF with the OMOP Common Data Model** in the Journal of Biomedical Informatics.



ELSEVIER

Journal of Biomedical Informatics

Available online 9 September 2022, 104201

In Press, Journal Pre-proof

Original Research

FHIR-Ontop-OMOP: Building Clinical Knowledge Graphs in FHIR RDF with the OMOP Common Data Model

Guohui Xiao [a], Emily Pfaff [b], Eric Prud'hommeaux [c], David Booth [d], Deepak K. Sharma [e], Nan Huo [e], Yue Yu [e], Nansu Zong [e], Kathryn J. Ruddy [e], Christopher G. Chute [f], Guoqian Jiang [e]

Show more

+ Add to Mendeley  Share  Cite

https://doi.org/10.1016/j.jbi.2022.104201

Get rights and content

Under a Creative Commons license

Open access

Abstract

Background

Knowledge graphs (KGs) play a key role to enable explainable artificial intelligence (AI) applications in healthcare. Constructing clinical knowledge graphs (CKGs) against heterogeneous electronic health records (EHRs) has been desired by the research and healthcare AI communities. From the standardization perspective,

# OHDSI Shoutouts!

Congratulations to the team of **Benoit L Marteau, Yuanda Zhu, Felipe Giuste, Wenqi Shi, Ashley Carpenter, Coleman Hilton, and May D Wang** on the publication of **Accelerating Multi-site Health Informatics with Streamlined Data Infrastructure using OMOP-on-FHIR** in the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).

# OHDSI Shoutouts!

# OHDSI Shoutouts!

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

Have a study published? Please send to sachson@ohdsi.org so we can share during this call and on our social channels.
Let's work together to promote the collaborative work happening in OHDSI!

# Three Stages of The Journey

## Where Have We Been?

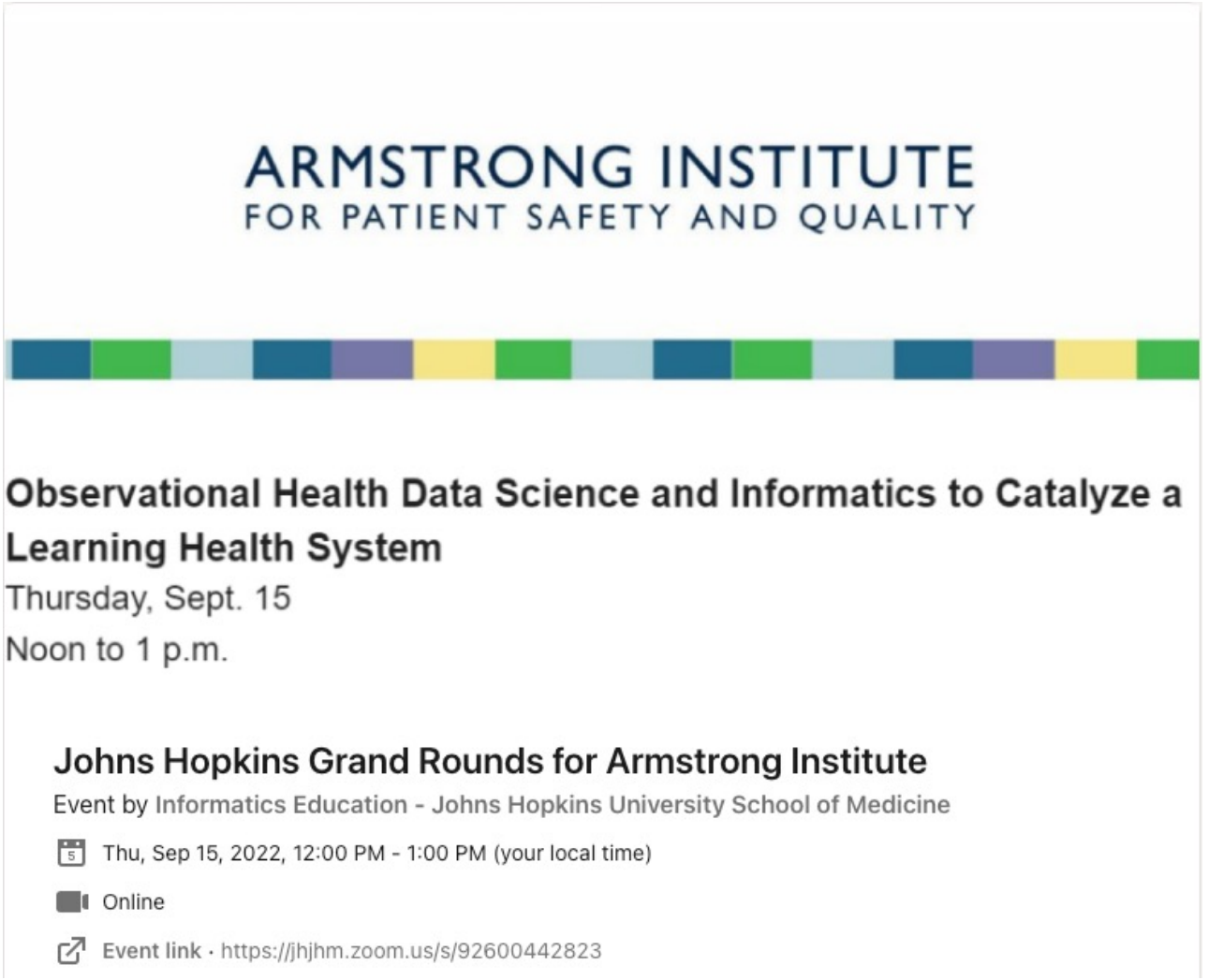
## Where Are We Now?

## Where Are We Going?

# Upcoming Workgroup Calls

| Date | Time (ET) | Meeting |
|---|---|---|
| Tuesday | 12 pm | Common Data Model Vocabulary |
| Tuesday | 3 pm | OMOP CDM Oncology Outreach/Research Subgroup |
| Tuesday | 6 pm | Eye Care and Vision Research |
| Wednesday | 10 am | FHIR and OMOP Digital Quality Measurements Subgroup (ZOOM) |
| Wednesday | 2 pm | Natural Language Processing |
| Thursday | 12 pm | HADES |
| Thursday | 12 pm | FHIR and OMOP Oncology Subgroup |
| Thursday | 1 pm | OMOP CDM Oncology Vocabulary/Development Subgroup |
| Thursday | 6 pm | FHIR and OMOP Digital Quality Measurements Subgroup (ZOOM) |
| Friday | 9 am | GIS-Geographic Information System Development |
| Friday | 10:30 am | Clinical Trials |

# Paul Nagy Hosts Grand Rounds Sept. 15



# 2022 Titan Award Nominees

Congratulations to the 53 individuals or teams who were nominated for a 2022 Titan Award.

The Titan Award recipients will be announced during the closing talk at the 2022 OHDSI Symposium.

Thamir Alshammary | Juan Banda | Adam Black | Fan Bu | Montse Camprubi | Yong Chen | Marcel de Wilde | Frank DeFalco | Egill Fridgeirsson | Jamie Gilbert | Jake Gillberg | Jason Hsu | Nigel Hughes | Yu-Chuan Jack Li | Mik Kallfelz | Andy Kanter | Elisse Katzman | Chungsoo Kim | Greg Klebanov | Chris Knoll | Kristin Kostka | Manlik Kwong | Christophe Lambert | Martin Lavallee | Jing Li | Xintong Li | Star Liu | Ajit Londhe | Aniek Markus | Evan Minty | Paul Nagy | Karthik Natarajan | Aki Nishimura | Anna Ostropolets | Melanie Philofsky | Gowtham Rao | Berta Raventos | Craig Sachson | Martijn Schuemie | Azza Shoaibi | Marc Suchard | Cynthia Sung | Joel Swerdel | May Terry | Don Torok | Cynthia Yang | Jacob Zelko | Center for Surgical Science Prediction study team | LEGEND-T2DM | N3C | Thrombosis w Thrombocytopenia phenotype project team | Vaccine Evidence Workgroup

# OHDSI Newsletter Is Available

## The Journey Newsletter (September 2022)

The #OHDSI2022 Symposium is less than seven weeks away, and the weekend agenda is available in this newsletter. There will be presentations, the collaborator showcase, a full-day tutorial, workgroup activities and plenty more. You can also learn more about our new data partner survey, a new collaborator spotlight, 13 recent publications, several presentations from the past month and plenty more in the latest edition of The Journey! **#JoinTheJourney**

## Objective Diagnostics Plenary, OHDSI Support for Regulatory Agencies Highlight Agenda For 2022 Symposium


Simply returning together for the first time since 2019 would be enough reason to look forward to the 2022 OHDSI Symposium, but there will also be several impactful presentations, the largest collaborator showcase in community history and plenty more activities that will make the Oct. 14-16 weekend a special one.

## Publications

Vorisek CN, Lehne M, Klopfenstein SAI, Mayer PJ, Bartschke A, Haese T, Thun S. Fast Healthcare Interoperability Resources (FHIR) for Interoperability in Health Research: Systematic Review. JMIR Med Inform. 2022;10(7):e35724. Epub 20220719. doi: 10.2196/35724. PubMed PMID: 35852842; PubMed Central PMCID: PMCPMC9346559.

Kim S, Bang JI, Boo D, Kim B, Choi IY, Ko S, Yoo IR, Kim K, Kim J, Joo Y, Ryoo HG, Paeng JC, Park JM, Jang W, Kim B, Chung Y, Yang D, Yoo S, Lee HY. Second primary malignancy risk in thyroid cancer and matched patients with and without radioiodine therapy analysis from the observational health data sciences and informatics. Eur J Nucl Med Mol Imaging. 2022;49(10):3547-56. Epub 20220401. doi: 10.1007/s00259-022-05779-9. PubMed PMID: 35362796.

Lin V, Tsouchnika A, Allakhverdiiev E, Rosen AW, Gögenur M, Clausen JSR, Bräuner KB, Walbech JS, Rijnbeek P, Drakos I, Gögenur I. Training prediction models for individual risk assessment of postoperative complications after surgery for colorectal cancer. Tech Coloproctol. 2022;26(8):665-75. Epub 20220520. doi: 10.1007/s10151-022-02624-x. PubMed PMID: 35593971.

Bräuner KB, Rosen AW, Tsouchnika A, Walbech JS, Gögenur M, Lin VA, Clausen JSR, Gögenur I. Developing prediction models for short-term mortality after surgery for colorectal cancer using a Danish national quality assurance database. Int J Colorectal Dis. 2022;37(8):1835-43. Epub 20220718. doi: 10.1007/s00384-022-04207-6. PubMed PMID: 35849195.

Lamer A, Moussa MD, Marcilly R, Logier R, Vallet B, Tavernier B. Development and usage of an anesthesia data warehouse: lessons learnt from a 10-year project. J Clin Monit Comput. 2022. Epub 20220806. doi: 10.1007/s10877-022-00898-y. PubMed PMID: 35933465.

Shah SC, Canakis A, Halvorson AE, Dorn C, Wilson O, Denton J, Hauger R, Hunt C, Suzuki A, Matheny ME, Siew E, Hung A, Greevy RA, Jr., Roumie CL. Associations Between Gastrointestinal Symptoms and COVID-19 Severity Outcomes, Based on a Propensity Score-Weighted Analysis of a Nationwide Cohort. Gastro Hep Adv. 2022. Epub 20220807. doi: 10.1016/j.gastha.2022.06.015. PubMed PMID: 35966642; PubMed Central PMCID: PMCPMC9357443.

## September Update Podcast


In the latest On The Journey video, Patrick Ryan and Craig Sachson discuss the agenda for the main conference at the OHDSI Symposium, including the plenary (Objective Diagnostics: A pathway to provably reliable evidence), a set of presentations around OHDSI support for regulatory agencies, the collaborator showcase, and more. They also discuss the recent data partner survey that went out and the importance of keeping an update global database.

## Community Updates

**Where Have We Been?**

- Paul Nagy led an informative session on what it takes to build organizational support for adopting the OMOP CDM and OHDSI tools, as well as building organizational capacity for conducting observational research. Video from this panel, which included Keran Moll, Greg Klebanov and Ajit Londhe, is available here.
- The third workgroup leadership summit of 2022 was held in August and focused on building connections between workgroups and maximizing time together at the OHDSI symposium. There are several activities planned during the symposium weekend to strengthen bonds both within and between the workgroups.
- The first session of the Early-Stage Researchers Speaker Series, Asieh Golozar discussed her career path, how OHDSI has influenced her journey, and shares some tips for moving forward in health data sciences. You can watch it here, and be on the lookout for similar sessions in the future.

**Where Are We Now?**

- The agenda for the OHDSI 2022 Symposium was recently announced. This document lists the full agenda for the main conference, including all planned presentations, as well as the lightning talks and software demos for the Collaborator Showcase. A later version will include the 100+ posters that will also be presented at the showcase. If you haven't already registered for any of the events, please visit our symposium homepage and assure your spot at our biggest event of the year.

# #OHDSI2022 Agenda

OHDSI
Observational Health Data Sciences and Informatics

OHDSI 2022 Symposium
Oct. 14-16, 2022
Bethesda North Marriott Hotel & Conference Center

## Main Conference Agenda • Oct. 14

| | |
|---|---|
| 7:30 am - 8:30 am<br>*Ballroom AE Foyer* | Registration and Lite Breakfast |
| 9:00 am - 10:00 am<br>*Ballroom DE* | **State of the Community**<br>George Hripcsak, Columbia University<br>• *presentation of 2020, 2021 Titan Awards* |
| 10:00 am - 10:45 am<br>*Ballroom DE* | **Workgroup and Chapter connections**<br>• *workgroup/chapter leads will be distributed across the venue and available for networking to share activities and progress and connect for future collaborations*<br>**OHDSI Speed Dating** |
| 10:45 am - 12:15 pm<br>*Ballroom DE* | **Plenary: Objective Diagnostics: A pathway to provably reliable evidence**<br>Martijn Schuemie, Johnson & Johnson |
| 12:15 pm - 1:00 pm<br>*Ballroom Foyer* | **Buffet Lunch**<br>• *buffet in exhibitor space* |
| 1:00 pm - 2:00 pm<br>*Ballroom DE* | **Presentations: OHDSI support for regulatory authorities**<br>moderator: Jody-Ann McLeggon, Columbia University<br>• *"US FDA/CBER: Performance of vaccine safety surveillance methods" Fan Bu, UCLA*<br>• *"Korea Ministry of Food and Drug Safety: Replication of clinical trials in electronic health records" Seng Chan You, Yonsei University*<br>• *"European Medicines Agency: DARWIN-EU" Peter Rijnbeek, Erasmus MC* |
| 2:00 pm - 3:00 pm<br>*Ballroom ABC* | **Collaborator Showcase, Round 1**<br>• *Poster presentations with poster walks*<br>• *Software demonstrations*<br>• *Exhibitors* |
| 3:00 pm - 4:00 pm<br>*Ballroom DE* | **Collaborator Showcase Lightning Talks**<br>moderator: Kristin Kostka, Roux Institute at Northeastern University<br>• *"Disambiguation of ICPC codes using free-text and active learning to improve concept mappings" Tom Seinen, Erasmus MC*<br>• *"OHDSI Phenotype Phebruary: lessons learned" Azza Shoaibi, Johnson & Johnson* |

OHDSI
Observational Health Data Sciences and Informatics

OHDSI 2022 Symposium
Oct. 14-16, 2022
Bethesda North Marriott Hotel & Conference Center

## Main Conference Agenda • Oct. 14

| | |
|---|---|
| 3:00 pm - 4:00 pm<br>*Ballroom DE*<br>*(continued)* | • *"Reduce, Reuse, & Recycle: Going Green with Atlas Reusables" Ajit Londhe, Amgen*<br>• *"Best practices for prognostic model development using observational health data: a scoping review" Cynthia Yang, Erasmus MC*<br>• *"Machine Learning for Predicting Patients at Risk of Prolonged Opioid Use Following Surgery" Behzad Naderalvojoud, Stanford University*<br>• *"When does statistical equality meet health equity: developing analytical pipelines to compare associational and causal fairness in their application to EHR data" Linying Zhang, Columbia University*<br>• *"Analyzing the Effect of Hypertension on Retinal Thickness Using Radiology Common Data Model (R-CDM)" Chul Hyoung Park, Ajou University*<br>• *"Multinational Patterns of Second-line Anti-hyperglycemic Drug Initiation: A LEGEND-T2DM Study" Lovedeep Dhingra, Yale University* |
| 4:00 pm - 5:00 pm<br>*Ballroom ABC* | **Collaborator Showcase, Round 2**<br>• *Poster presentations with poster walks*<br>• *Software demonstrations*<br>• *Exhibitors* |
| 5:00 pm - 6:00 pm<br>*Ballroom DE* | **Closing Talk: Building A Healthier World Together**<br>Patrick Ryan, Johnson & Johnson, Columbia University<br>• *2022 Titan Awards*<br>• *Group photo at conclusion* |
| 6:00 pm - 7:00 pm<br>*Ballroom ABC* | **Networking Reception** |

**Register Here:**
**ohdsi.org/ohdsi2022symposium/**

**#JoinTheJourney**

**bit.ly/OHDSI2022-Agenda**

# #OHDSI2022 Agenda

OHDSI
Observational Health Data Sciences and Informatics

OHDSI 2022 Symposium
Oct. 14-16, 2022
Bethesda North Marriott Hotel &
Conference Center

## Full-Day Tutorial • Oct. 15

### An Introductory Journey From Data To Evidence

In this tutorial, we will introduce participants to steps along the journey from data to evidence using the OMOP Common Data Model, OHDSI tools and scientific best practices.In each 50-minute segment, the class will learn the conceptual framing of the problem and approach to the solution.Then, the class will have the opportunity to have hands-on exposure to design and implementation of analyses and interpretation of results.The course will be motivated by a real use case: using observational data to generate evidence about the relationship between an exposure and outcome, and will highlight how the suite of OHDSI tools and practices can enable such learning.

This class is designed for newcomers to the OHDSI community who are looking for a high-level summary across a wide range of topics covered within the OHDSI community. It's also designed for those in the OHDSI community who may be focused in one particular area of the journey who want exposure to the other areas, so they can better understand how their work contributes to be 'big picture,' and advances the mission to improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care

The tutorial will be held in White Oak A.

| Time | Title | Faculty |
|---|---|---|
| 7:30 am - 8:30 am | Registration/Lite Breakfast (White Oak Foyer) | |
| 8:30 am - 9:00 am | Overview of the OHDSI Journey: where are we going? | Patrick Ryan |
| 9:00 am - 9:50 am | OMOP Common Data Model and vocabulary | Clair Blacketer |
| 9:50 am - 10:00 am | Energy Break | |
| 10:00 am - 10:50 am | ETL a source database into OMOP CDM | Melanie Philofsky |
| 10:50 am - 11:00 am | Energy Break | |
| 11:00 am - 11:50 am | Creating Cohort Definitions | Asieh Golozar |
| 11:50 am - 12:30 pm | Buffet Lunch | |
| 12:30 pm - 1:20 pm | Phenotype Evaluation | Gowtham Rao |
| 1:20 pm - 1:30 pm | Energy Break | |
| 1:30 pm - 2:20 pm | Characterization | Kristin Kostka |
| 2:20 pm - 2:30 pm | Energy Break | |
| 2:30 pm - 3:20 pm | Estimation | Martijn Schuemie |
| 3:20 pm - 3:30 pm | Energy Break | |
| 3:30 pm - 4:20 pm | Prediction | Jenna Reps |
| 4:20 pm - 5:00 pm | Recap of the OHDSI Journey: Where do we go from here? | George Hripcsak |

OHDSI
Observational Health Data Sciences and Informatics

OHDSI 2022 Symposium
Oct. 14-16, 2022
Bethesda North Marriott Hotel &
Conference Center

## Collaborator Showcase

### Poster Presentations

The 2022 OHDSI Symposium will host two sessions featuring a total of over 100 posters that highlight the breadth of global research happening within our community. Closer to the symposium weekend, we will announce all of the posters and presenters.

### Software Demos

The 2022 OHDSI Symposium will feature 17 software demonstrations during the Collaborator Showcase sessions, listed below:

**A demonstration of the EnsemblePatientLevelPredition package** (Jenna M. Reps, Jenna Wong, and Ross Williams)

**CohortIncidence: A Software Demonstration** (Christopher Knoll)

**Criteria2Query 2.0: Combining Human and Machine Intelligence for Cohort Identification** (Yilu Fang, Betina Idnay, Yingcheng Sun, Hao Liu, Zhehuan Chen, Karen Marder, Hua Xu, Rebecca Schnall, Chunhua Weng)

**Data Network Feasibility Tool - Software Demonstration** (Frank DeFalco, Clair Blacketer)

**Data Quality Dashboard v2.0** (Clair Blacketer, Frank DeFalco, Anthony Molinaro, Dmitry Ilyn,Luis Alaniz, Maxim Moinat)

**Einstein-ATLAS: Leveraging OHDSI/ATLAS and Open-Source Development to Support Translational Research, Data Science, and Regulatory Compliance** (Parsa Mirhaji, Selvin Soby, Erin Henninger, Chandra Nelapatla, Manuel Wahle, Boudewijn Aasman, Eran Belin)

**ohdsitargets - An R package for building OHDSI study pipelines using targets** (Adam Black, Martin Lavallee, Asieh Golozar, Gregory Klebanov)

**OmopPopEpi: An R package to compute population-level incidence and prevalence using the OMOP common data model** (Marta Catala, Berta Raventas, Mike Du, Yuchen Guo, Xintong Li, Ross Williams, Talita Duarte Salles, Daniel Prieto Alhambra, Edward Burn)

**PHOEBE 2.0: selecting the right concept sets for the right patients using lexical, semantic, and data-driven recommendations** (Anna Ostropolets, George Hripcsak, Patrick Ryan)

**REal World Assessment and Research of Drugs (REWARD): presenting an open-source package for Population-level effect estimation at the scale of all outcomes by all exposures** (James Gilbert)

**Simple and practical EMR to OMOP CDM ETL tool** (Pieter-Jan Lammertyn, Stijn Dupulthys, Louise Berteloot, Peter De Jaeger, Kim Denturck, Nathalie Mertens)

**Standardizing Knowledge of Drug Effects: An Application of PheKnowLator for Drug Safety** (Tiffany J. Callahan, Patrick B. Ryan, George Hripcsak)

**Strategus: Marching towards transparent, reproducible research** (Anthony G. Sena, Christopher Knoll, James Gilbert, Jenna Reps, Frank DeFalco, Clair Blacketer, Anthony Molinaro, Joshua Ide, Patrick Ryan, Martijn Schuemie)

**The OHDSI Community Dashboard: Tracking the Health and Impact of the Open Science Observational Health Data Sciences and Informatics Community** (Star Liu, Asieh Golozar, Jody-Ann McLeggon, Adam Black, Paul Nagy)

**Understanding circe-be logic through Capr for generating complex cohort definitions** (Martin Lavallee, Adam Black and Asieh Golozar)

**Using dbt - a free and open-source software - to transform data into OMOP CDM in the ETL process** (Thanapat Pitchayarat, Gun Pinyo, Watcharaporn Tanchotsrinon, Somkid Khamsrimuang, Chalita Issarasittiphap, Chaiyanun Bootnumpech, Noppon Siangchin, Kanphitcha Promma, Nattachai Bovornmongkolsak, Prapat Suriyaphol, Natthawut Adulyanukosol)

**Vocabulary Versioning: Tracking Concepts over Time Software Demonstration** (Tom Seinen, Peter Rijnbeek)

**bit.ly/OHDSI2022-Agenda**

# 2022 OHDSI Symposium

## Registration is OPEN for #OHDSI2022!





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

**www.ohdsi.org/ohdsi2022symposium**

# 2022 OHDSI Symposium

OHDSI Community Calls | Events & Past Collaborations | Learn About & Join OHDSI Workgroups | This Week In OHDSI | EHDEN Academy
Annual Report: Our Journey | Publications | Support & Sponsorship | OHDSI2022 Symposium | Newsletters | Follow OHDSI on Social

## 2022 OHDSI Symposium

### Oct. 14-16 · Bethesda North Marriott Hotel & Conference Center

We are thrilled to announce that registration for the 2022 OHDSI Symposium, which will be held Oct. 14-16 at the Bethesda North Marriott Hotel & Conference Center, is now open!

It is so exciting to bring our community back together this fall. Our collaborator showcase will return; please click the link to see how you can take part in our poster presentations, software demos and lightning talks. The full agenda for our conference is still being developed, so please continue to check the OHDSI website (www.ohdsi.org) and our social platforms for updates as we plan for the 2022 Symposium.

The main conference will be held Friday, Oct. 14. A full-day tutorial will be held Saturday, Oct. 15, while other community activities will be held both Oct. 15 and Oct. 16.

### Symposium Registration Details

#### Friday, Oct. 14 – Main Conference

**Registration Fee:** $500*
** this is an open and inclusive event; if the registration fee represents a burden to you, please contact symposium@ohdsi.org.*

Should you need to make changes or cancel your registration ticket, please follow the instructions you will receive on your Eventbrite confirmation upon registration completion. Please note that tickets can be refunded up until 7 days prior to the event; Eventbrite fees are not refundable.

Register For The Main Conference · Friday, Oct. 14

#### Saturday, Oct. 15 – Full-Day Tutorial: An Introductory Journey From Data To Evidence

**Registration Fee:** $300*
** this is an open and inclusive event; if the registration fee represents a burden to you, please contact symposium@ohdsi.org.*

Should you need to make changes or cancel your registration ticket, please follow the instructions you will receive on your Eventbrite confirmation upon registration completion. Please note that tickets can be refunded up until 7 days prior to the event; Eventbrite fees are not refundable.

Register For The Full-Day Tutorial · Saturday, Oct. 15

What Will Be Taught At This Tutorial?

#### Saturday, Oct. 15 and Sunday, Oct. 16 — Community Activities

A highlight of the OHDSI Symposium will be a full weekend of workgroup activities and meetings within the Bethesda North Marriott Hotel & Conference Center. You are now able to register for any workgroup sessions as long as there is no overlap between any two sessions; registration is free, but please do so early as this will be first-come, first-served due to room capacity.

See The Schedule & Agenda For Workgroup Activities · Weekend of Oct. 15-16

Register For Workgroup Activities · Weekend of Oct. 15-16

---

### Hotel Information and Sleeping Room Block

**Hotel:** Bethesda North Marriott Hotel & Conference Center
**Address:** 5701 Marinelli Road, Rockville, Maryland, 20852
**Hotel Main Number:** (301) 822-9200
**Reservations Toll Free:** (877) 212-5752
**Reservations Local Phone:** (301) 822-9200

This year, OHDSI is holding a sleeping room block for the nights of Oct. 13 and 14 with a special room rate of $179 plus taxes. Please note that all sleeping rooms are on a first-come, first-served basis. To help us in the planning process, we ask that you do not cancel your hotel room ordered through the OHDSI Room Block. If you must cancel, please let us know prior to Thursday, Sept. 1, so that we can offer the room to others who may need one. Once the room block is full, or if specific nights are sold out, you may make additional room reservations on the hotel's website or by calling the hotel phone number above. Please note that OHDSI is not holding any sleeping rooms on Saturday, Oct. 15. Therefore, please call the hotel phone number or make this reservation online should you plan to stay Saturday night.

**ohdsi.org/ohdsi2022symposium**

# Job Openings

**Talita Duarte-Salles**, a 2020 OHDSI Titan Award honoree, recently announced an opening for a post-doc to join her Real World Epidemiology group at the The Institut d'Investigació en Atenció Primària Jordi Gol (IDIAPJGol).

The application deadline is Sept. 18, 2022. More information and an application link will be available on the community calls page.

IDIAP JORDI GOL

**Job Description Post-Doctoral Researcher**

**"Institut d'Investigació en Atenció Primària Jordi Gol"**

The Institut d'Investigació en Atenció Primària Jordi Gol i Gurina is looking for a postdoctoral researcher to join the Real-World Epidemiology (RWEpi) research group led by Dr. Talita Duarte-Salles. The RWEpi research group is a multidisciplinary team including epidemiologists, statisticians, clinicians, data scientists, and pre- and post-doctoral researchers with extensive experience in analysing and interpreting large, real-world data from different countries. Our main goal is to generate reliable scientific evidence by using routinely collected data from different settings of the healthcare system. We are currently involved in a number of national and international studies focused on the understanding of the natural history and determinants of multiple diseases/conditions (e.g. cancer, COVID-19, obesity, mental health, etc), the safety and effectiveness of new medicines and vaccines, and the implementation of methodological approaches to analyse and evaluate the quality of real-world data. We are also particularly interested in population subgroups such as the pediatric and pregnant women populations, being recently able to link a large cohort of mother and child electronic healthcare records from primary care in Catalonia; as well as understanding health inequalities, and the role of lifestyle and urban environmental exposures on health.

# #OHDSISocialShowcase This Week

*Common data environment for source vocabularies mapping*

PRESENTER: **Irina Zherko**

**BACKGROUND**

The problem of mapping similar codes from different vocabularies that have similar source concepts (e.g., ICD10, ICD10CM) with every new vocabulary and on every refresh called for a more structured semi-automated approach. Commercial datasets, while being specifically designed for each customer, often contain duplicate or similar data that can potentially be combined, mapped together or reused.

**METHODS**

The CDE should contain the most complete set of source data, structured and organized into groups by code, source_code_description, as well as the frequency of occurrence of each concept and the group as a whole.

**RESULTS**

The algorithm for CDE creating includes:

- Gathering data. For each source_code/source_code_description combination, a flag with the data source (customer) name and/or the name of the dataset in case of combining datasets in one environment as well as respective count of the records for every customer/dataset and overall count must be stored in a separate fields.
- Preparing concept names or descriptions for sorting.
- Additional information per domain adding. In accordance with the requirements of custom mapping, additional fields can be filled for certain domains.
- Defining the groups using 2 fields at once (source_code and cleaned source_code_description) with the help of recursive joins.
- Sorting by group count and concept count for mapping prioritization.
- Mapping to standard concepts.

**A common data environment** approach can streamline source vocabulary processing for **repetitive processes, harmonize and enrich mapping,** speed up the process of **data reuse.**

The **common data environment** should contain the most complete set of source data, organized into groups by **code**, **code_description**, frequency of occurrence of **each concept and the whole group.**

Scan QR to link to **script with recursive joins** for grouping using several fields at once

Table 1. Differences in the same concept codes mapping in ICD10 and ICD10CM.

| ICD10/ ICD10CM code | ICD10/ ICD10CM name | target_concept_id for ICD10 | target_name for ICD10 | target_concept_id for ICD10CM | target_name for ICD10CM |
|---|---|---|---|---|---|
| A06.2 | Amoeboma of intestine | 37116456 | Amebioma of large intestine caused by Entamoeba histolytica | 4051447 | Amebioma |
| A32.1 | Listerial meningitis and meningoencephalitis | 4190307 | Inflammatory disease of the central nervous system | 4136364 | Listeria meningitis |
| A32.1 | Listerial meningitis and meningoencephalitis | 4190709 | Listeriosis | | |
| E30.6 | Type 1 diabetes mellitus, With other complications | 435216 | Disorder due to type 1 diabetes mellitus | 201254 | Type 1 diabetes mellitus |
| | | | | 4417781 | Complication due to diabetes mellitus |


Figure 1. The theoretical scheme of workflow with CDE application

Table 2. Examples of source_code_description transformation

| Original source_code_description | Cleaned source_code_description |
|---|---|
| Photographic diagnosis by simple radiography (a) | photographic diagnosis by simple radiography |
| Steady-state Plasma Glucose 0 min | steadystate glucose min |
| Above the knee deep vein thrombosis age in days | abovekneedthrombosisage |
| BOTH LUNGS: DIFFERENCE IN NUMBER VOXELS ABOVE OR EQUAL TO -200 PERCENT | differencenumbervoxelsabovezeroequalto percent |

Table 3. The part of the CDE of two custom datasets

There are groups of concepts with identical meaning. Mappings, which need improvements are marked red. Improvements can be easily done using mappings from the rows nearby.

| source_code_description | count dataset1 | count dataset2 | cumulative count | flag | target_concept_code | target_concept_name |
|---|---|---|---|---|---|---|
| stereotactic radiation therapy with a gamma knife | 85595 | | 85595 | dataset1 | 4401560008 | Stereotactic destruction of lesion using gamma radiation |
| stereotactic radiotherapy with gamma knife | | 11823 | 11823 | dataset2 | 2000153003 | Radionuclide therapy |
| | | | | | | |
| subconjunctival injection procedure | 83201 | | 83201 | dataset1 | 744132004 | Subconjunctival injection |
| subconjunctival injection | 25238 | | 25238 | dataset2 | No matching concept | No matching concept |
| | | | | | | |
| transcutaneous nephrostomy/pyelostomy percutaneous nephrostomy | 84428 | | 84428 | dataset1 | 19810009 | Nephrostomy |
| | 7553 | | 7553 | dataset2 | No matching concept | No matching concept |
| | | | | | | |
| bun | 633454697 | | 633454697 | dataset1 | 105011006 | Blood urea nitrogen measurement |
| bun (urine) | 5193674 | | 5193674 | dataset1 | 88263 | Urea nitrogen, urine |
| urea nitrogen (bun) | | 1127199967 | 1127199967 | dataset2 | 105011006 | Blood urea nitrogen measurement |
| urea nitrogen (urine) | | 869412 | 869412 | dataset2 | 88263 | Urea nitrogen, urine |


Irina Zherko, MD,
Mikhail Nerovnya, MD,
Michael Kallfelz, MD,
Alexander Davydov, MD


ODYSSEUS DATA SERVICES INC

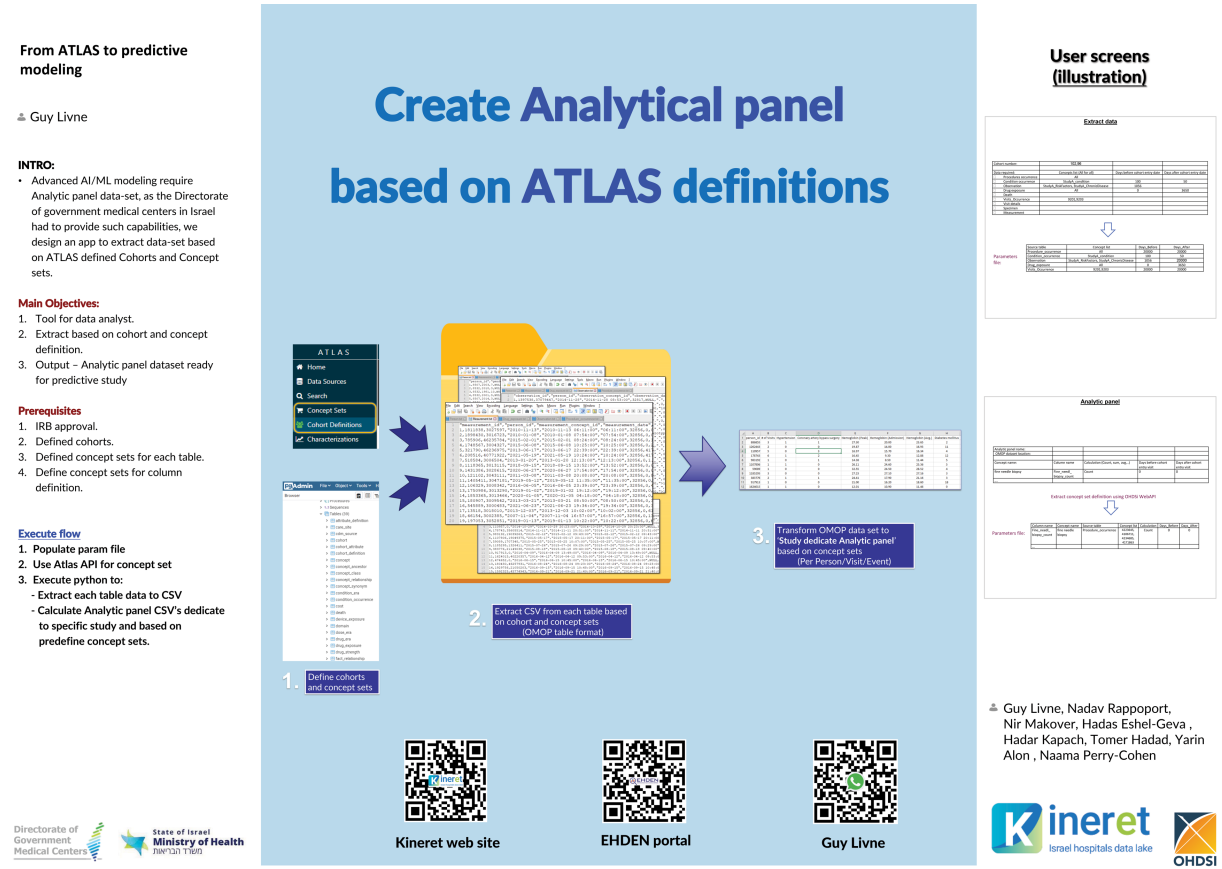
OHDSI

## MONDAY

## Common data environment for source vocabularies mapping

**Lead: Irina Zherko**

# #OHDSISocialShowcase This Week




**TUESDAY**

## From ATLAS to predictive modeling CDM data extracting & Study preparations

**Lead: Guy Livne**

# #OHDSISocialShowcase This Week

***Utilising real-world evidence for health technology assessment***
*Development of a cancer survival use case*

PRESENTER: **Ravinder Claire**

**INTRODUCTION**

- In cancer health technology assessments (HTA), extrapolation techniques are used to estimate overall survival in people who receive a treatment, beyond observed trial data.
- This is important to inform economic evaluations used for HTA, which assess cost effectiveness over a lifetime.
- However, this is a key source of decision uncertainty because it involves forecasting the future based on shorter-term observed data.
- Real-world evidence can help address this uncertainty. For example, country-specific, real-world survival data for patients receiving the current standard of care could be used to validate the trial-based survival extrapolations typically used in HTA processes.
- Here, we describe an EHDEN use case that is in development to demonstrate how real-world data from the OMOP-CDM could help address this priority issue for HTA.

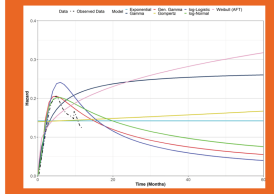
**AIMS**

Using CPRD data in the first instance, we aim to:

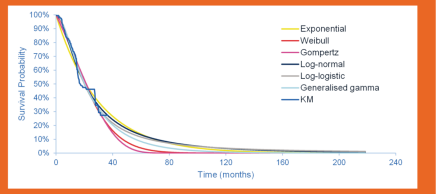
1. Develop and assess data quality of phenotypes for the identification of breast, colorectal, head and neck, lung, liver, prostate, and stomach cancer.
2. Estimate overall survival of the studied cancers, stratifying by key demographic variables and comorbidity.
3. Fit standard parametric survival functions to the data to extrapolate long-term natural history of the studied cancers, including visual and statistical goodness of fit.
4. Include these outputs in a user-friendly, interactive 'EHDEN Cancer Survival Data Dashboard'.

We aim to develop an **EHDEN Cancer Survival Dashboard**, allowing users to quickly examine **survival data** and explore **long-term projections**.

**Smoothed hazard plots**



**Extrapolated survival curves**




**Goodness of fit**

| Parametric Model | Akaike Information Criterion | Bayesian Information Criterion | 3-year survival estimate | 5-year survival estimate |
|---|---|---|---|---|
| **Exponential** | 346.4 | 349.1 | 31.4% | 14.6% |
| **Generalised Gamma** | 342.5 | 350.6 | 25.5% | 9.3% |
| **Gompertz** | 345.0 | 350.3 | 21.9% | 1.5% |
| **Log-logistic** | 340.3 | 345.6 | 25.8% | 11.9% |
| **Log-normal** | 341.3 | 346.7 | 28.4% | 13.6% |
| **Weibull** | 341.6 | 347.0 | 22.0% | 4.0% |

Potential survival outputs from the EHDEN Cancer Data Dashboard. Please note, these figures are samples and have been intentionally de-labelled for illustration purposes. Data from NICE technology appraisal 722: Pemigatinib for treating relapsed or refractory advanced cholangiocarcinoma with FGFR2 alterations.

**RESULTS**

- We engaged with potential end users of the Dashboard—academic experts, HTA decision makers, industry experts and potential EHDEN data partners—to understand what information would be most useful to them.
- Some of the potential outputs identified from this consultation can be found in Figure 1.

**DISCUSSION**

- The development of this survival analysis use case will demonstrate a potential benefit of EHDEN and OHDSI tools to address priority areas of HTA agencies and industry stakeholders.
- The main strengths of this use case are likely to be a large sample size and the observational nature of the data, meaning its outputs will be representative of real-world clinical practice and outcomes.
- The development of the EHDEN Cancer Survival Data Dashboard will encourage clear and transparent reporting.
- From an HTA perspective, the dashboard will allow alternative parametric functions for survival extrapolation to be overlaid on observed Kaplan-Meier curves.
- This could have real benefits to healthcare reimbursement decisions based on HTA—for example, by informing comparator survival estimates where there is only a single-arm study, or by validating survival estimates from a clinical trial in the most relevant, real-world population.
- Other identified priority areas for HTA will be considered in future EHDEN HTA use cases in other disease areas.

Ravinder Claire[1], Jamie Elvidge[1], Dalia Dawoud[1], Danielle Newby[2], Ed Burn[2]
[1]National Institute for Health and Care Excellence, UK.
[2]University of Oxford, UK

NICE National Institute for Health and Care Excellence
EHDEN EUROPEAN HEALTH DATA & EVIDENCE NETWORK
OHDSI

## WEDNESDAY

## Utilising real-world evidence for health technology assessment: development of a cancer survival use case

**Lead: Ravinder Claire**

# #OHDSISocialShowcase This Week

*Applying 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**

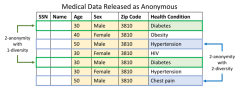
- We only considered some tables of the OMOP CDM.
- Tables in the groups "Standardized vocabularies", "Standardized health economics", "Standardized derived elements" and "Standardized metadata" were omitted.
- Key attributes, quasi-identifies and sensitive attributes were mapped to apply the privacy-preserving techniques.

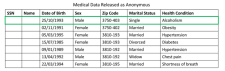
**Sharing aggregate data?**

**Robust anonymization through privacy-preserving techniques.**

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

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.




João Rafael Almeida and José Luís Oliveira

university of aveiro theoria poiesis praxis · IEETA · OHDSI

## THURSDAY

## Applying k-anonymity and l-diversity in OMOP CDM databases

**Lead: João Almeida**

# #OHDSISocialShowcase This Week

***Trial feasibility assessments in federated hospital Electronic Health Record networks, based on OMOP CDM***
*An objective of the IMI2 EU-PEARL Consortium*

PRESENTER: Eva-Maria Didden

**INTRODUCTION**

- Hospital Electronic Health Record (EHR) systems can inform the design of clinical trial protocols and optimize recruitment.
- EU-PEARL aims to assess the potential to use the OHDSI tooling and the OMOP CDM to evaluate protocol feasibility in hospital EHR networks.
- Here, we present results of a trial feasibility study in Neurofibromatosis Type 1 (NF1) with Optical Pathway Glioma (OPG) in the Erasmus MC hospital EHR system, using Atlas.

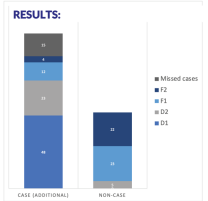
**METHODS**

Reusable NF1-OPG phenotype algorithms, two only based on diagnosis codes (D) and two also based on follow-up visits (F) :

**D1:** NF1 diagnosis & OPG diagnosis
**D2:** OPG diagnosis
**F1:** NF1 diagnosis & brain MRI & 4 encounters with an ophthalmologist within 365 days
**F2:** NF1 diagnosis & brain MRI & 3 encounters with an ophthalmologist within 365 days

- To identify missing cases, selected patients were compared with a list of known OPG patients.
- Patients additionally selected via the Atlas phenotype algorithms were classified as cases or non-cases via clinical chart review.

**RESULTS:**




Number of cases and non-cases included by each Atlas phenotype algorithm. D1 initially included 48 cases. Using D2, an additional 23 cases were included. With F1 and F2, 12 and 4 cases were identified. Each step also included more non-cases. 15 cases were not included in any of the 4 definitions.

OMOP CDM/ATLAS allows **identifying Neurofibromatosis Type1 patients** with an **Optical Pathway Glioma** in the **Electronic Health Record** database of a **clinical site.**

Leveraging this approach to a **federated site network** provides the potential to identify patients matching **trial eligibility** criteria on **large scale** and to **refine** criteria as appropriate.

EU-PEARL has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 853966.
The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and CHILDREN'S TUMOR FOUNDATION, GLOBAL ALLIANCE FOR TB DRUG DEVELOPMENT NON PROFIT ORGANISATION, SPRINGWORKS THERAPEUTICS INC.


**True cases:** 104 (61 from a pre-existing list; 43 additional ones from chart review of Atlas cohorts)

Patients selected by phenotype algorithms:
**D1:** 48; **D2:** 76; **F1:** 62; **F2:** 90

| Atlas cohort | N selected in cohort | N cases compared to original list (n=61) | N cases additionally reviewed (n=43) | Sensitivity (original cases only) | PPV (original and reviewed) |
|---|---|---|---|---|---|
| D1 | 48 | 36 | 12 | 59.0% (36/61) | 100.0% (48/48) |
| D2 | 76 | 40 | 31 | 65.6% (40/61) | 93.4% (71/76) |
| F1 | 62 | 29 | 10 | 47.5% (29/61) | 62.9% (39/62) |
| F2 | 90 | 32 | 13 | 52.5% (32/61) | 50.0% (45/90) |

**CONCLUSION**

**Summary:**

- NF1 patients with an OPG could be identified in the EHR database of a clinical site.
- Clinically meaningful variations between phenotype algorithms could be evaluated.

**Note:**

- For NF1, being a rare condition, a sensitive phenotype algorithm may be preferrable.
- For more common conditions, one may tend to use more specific algorithms.

**Next steps:**

1. Share the NF1-OPG phenotype algorithms with other sites of the federated EHR network
2. Obtain large-scale aggregate query results, including patient counts and characteristics
3. Overall and for each individual site, evaluate patient counts against expected prevalence
4. Refine phenotype algorithms as appropriate
5. Re-run the query, identify potential patients, and conduct in depth chart review to confirm eligibility to specific study protocol recruitment criteria
6. Evaluate site recruitment potential and study eligibility criteria

➔ **This promising approach will be replicated in 1-2 other diseases, and a general description of the methodology will be made available through EU-PEARL.**

AUTHORS: Eva-Maria Didden, Maxim Moinat, Britt Dhaenens, Esther Arévalo de Andres, Camille Couvert, Susana Kalko, Andreas Kremer, Martine Lewi, Cécile Spiertz, Eng Hooi Tan, Courtney Worrell, Nadir Ammour, Peter Rijnbeek, Rianne Oostenbrink, Dipak Kalra

EU-PEARL
imi innovative medicines initiative
OHDSI

## FRIDAY

## Trial feasibility assessments in federated hospital EHR networks, based on OMOP CDM: An objective of the IMI EU-PEARL Consortium

**Lead: Eva-Maria Didden**

# 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?

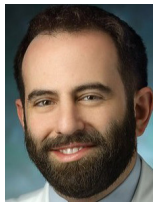
# Sept. 13: Clinical Registry Efforts in OHDSI

## How clinical registries and OHDSI can benefit from each other

**Presenter:** Paul Nagy • Program Director for Graduate Training in Biomedical Informatics and Data Science, Deputy Director of the Johns Hopkins Medicine Technology Innovation Center

## How to adapt a manual clinical registry to OMOP

**Presenter:** Matt Robinson • Assistant Professor, The Johns Hopkins University School of Medicine



## How to lower the ETL barrier going to OMOP with Perseus (Demonstration)

**Presenter:** Zachary Wang • Graduate Student, Johns Hopkins (2022 Kheiron Cohort member)

## Lowering the deployment burden with the cloud

**Presenter:** Lee Evans • Owner, LTS Computing LLC

## Distributed Machine Learning Using OMOP

**Presenter:** Emily Pfaff • Research Assistant Professor, University of North Carolina at Chapel Hill