
# Published Studies

## OHDSI Community Call
## Sept. 6, 2022 • 11 am ET

# Upcoming OHDSI Community Calls

| Date | Topic |
|---|---|
| Sept. 13 | Clinical Registry Efforts In OHDSI |
| Sept. 20 | 2022 OHDSI Symposium Preview |
| Sept. 27 | HTA Challenge |
| Oct. 4 | OHDSI Debates |
| Oct. 11 | Final OHDSI2021 Logistics |
| Oct. 18 | Welcome To OHDSI |
| Oct. 25 | Future Directions For OHDSI |

# Upcoming OHDSI Community Calls

| Date | Topic |
|---|---|
| Sept. 13 | Clinical Registry Efforts In OHDSI |
| Sept. 20 | 2022 OHDSI Symposium Preview |
| Sept. 27 | HTA Challenge |
| Oct. 4 | OHDSI Debates |
| Oct. 11 | Final OHDSI2021 Logistics |
| Oct. 18 | Welcome To OHDSI |
| Oct. 25 | Future Directions For OHDSI |

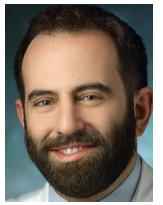
# Sept. 13: Clinical Registry Efforts in OHDSI

## 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

## How to take harmonization to the next step with unit level harmonization; How to create the governance for robust community developed phenotypes

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

# Three Stages of The Journey

## Where Have We Been?

## Where Are We Now?

## Where Are We Going?

# OHDSI Shoutouts!

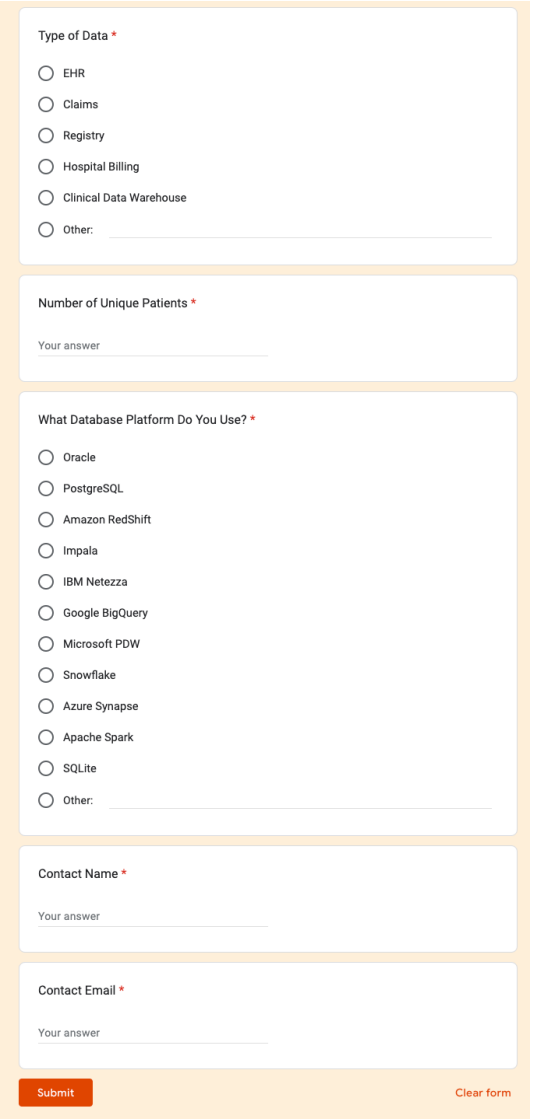
**Congratulations to the 53 individuals or teams who were nominated for a 2022 Titan Award.**

**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 Shoutouts!




**bit.ly/OHDSI-Data-Survey**

# 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 |
|---|---|---|
| Wednesday | 2 am | Patient-Level Prediction/Population-Level Estimation |
| Wednesday | 7 am | Medical Imaging |
| Wednesday | 8 am | Psychiatry |
| Wednesday | 9 am | ATLAS |
| Wednesday | 11 am | Open-Source Community |
| Wednesday | 12 pm | Health Equity |
| Wednesday | 12 pm | FHIR and OMOP Terminologies Subgroup (ZOOM) |
| Wednesday | 4 pm | FHIR and OMOP Data Model Harmonization Subgroup (ZOOM) |
| Thursday | 10 am | Data Quality Dashboard Development |
| Thursday | 12 pm | FHIR and OMOP Oncology Subgroup |
| Friday | 9 am | Education |
| Friday | 9 am | GIS – Geographic Information System General |
| Friday | 9 am | Phenotype Development and Evaluation |
| Friday | 10 pm | China Chapter |
| Monday | 10 am | Healthcare Systems Interest Group |
| Monday | 11 am | Early-Stage Researchers |
| Tuesday | 9 am | OMOP CDM Oncology Genomic Subgroup |

# CBER BEST Seminar Series Returns Tomorrow!

BEST

Webinar Registration

Topic: Quantitative Bias Analysis Methods to Improve Inferences

Speaker: Matthew P. Fox, DSc, MPH
Professor, BUSPH

Description: Observational epidemiologic research around vaccine efficacy and safety can provide important insights into causal relationships, but key sources of bias often impair the inferences we draw from these studies. Uncontrolled confounding, selection bias and information bias are common in epidemiologic research and failure to account for their impacts in a quantitative manner (rather than qualitative assessments in discussion sections of manuscripts after conclusions have been drawn) can led to poor inferences. This talk will give an overview of quantitative bias analysis methods to demonstrate how they can be implemented on both summary and record level data to account for the impact of systematic error on study results and provide some examples using data from the literature on how to apply these methods to vaccines safety data.

## Matthew Fox, DSc, MPH

Professor, Departments of Epidemiology and Global Health @Boston University

Matthew Fox, DSc, MPH, is a Professor in the Departments of Epidemiology and Global Health at Boston University. Dr. Fox joined Boston University in 2001. His research interests include treatment outcomes in HIV-treatment programs, infectious disease epidemiology (with specific interests in HIV and pneumonia), and epidemiologic methods. Dr. Fox works on ways to improve retention in HIV-care programs in South Africa from the time of testing HIV-positive through long-term treatment. As part of this work, he is involved in analyses to assess the impact of changes in South Africa's National Treatment Guidelines for HIV. Dr. Fox also does research on quantitative bias analysis and co-authored a book on these methods, Applying Quantitative Bias Analysis to Epidemiologic Data. He is also the host of a public health journal club podcast called Free Associations designed to help people stay current in the public health literature and think critically about the quality of research studies.

# 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.

# OHDSI Newsletter Is Available

OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Who We Are | OHDSI Updates & News | Standards | Software Tools | OHDSI Studies | Book of OHDSI | Resources | New To OHDSI?

OHDSI Community Calls | Events & Past Collaborations | Workgroups | EHDEN Academy | This Week In OHDSI | Our Journey (PDF)

Publications | Support & Sponsorship | 2022 OHDSI Symposium | 2022 APAC Symposium | Newsletters | Follow OHDSI on Social

- Subscribe
- September 2022
- [illegible] 2022
- July 2022
- June 2022
- May 2022
- April 2022
- March 2022
- February 2022
- 2021 In Review

## Welcome to OHDSI!

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

## 2022 [illegible]sium

The 2022 OH[illegible]eld Oct. 14-16 at the [illegible]otel & Conference [illegible]s opened.

We will hold [illegible]iday, Oct. 14, which wil[illegible]howcase. On Saturday[illegible]-day

# #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

---

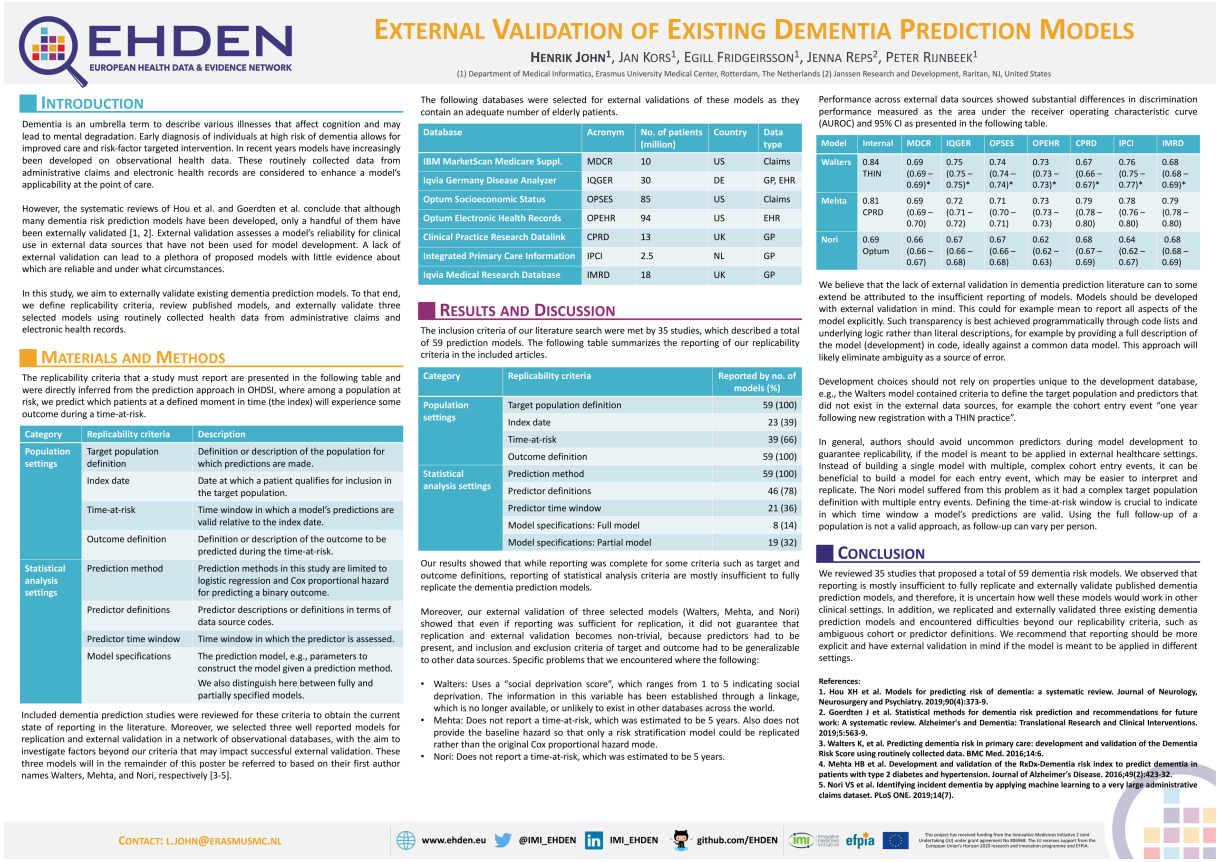
### Hotel Information and Sleeping Room Block

**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**

# #OHDSISocialShowcase This Week



EHDEN
EUROPEAN HEALTH DATA & EVIDENCE NETWORK

## External Validation of Existing Dementia Prediction Models

Henrik John[1], Jan Kors[1], Egill Fridgeirsson[1], Jenna Reps[2], Peter Rijnbeek[1]

(1) Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands (2) Janssen Research and Development, Raritan, NJ, United States

### Introduction

Dementia is an umbrella term to describe various illnesses that affect cognition and may lead to mental degradation. Early diagnosis of individuals at high risk of dementia allows for improved care and risk-factor targeted intervention. In recent years models have increasingly been developed on observational health data. These routinely collected data from administrative claims and electronic health records are considered to enhance a model's applicability at the point of care.

However, the systematic reviews of Hou et al. and Goerdten et al. conclude that although many dementia risk prediction models have been developed, only a handful of them have been externally validated [1, 2]. External validation assesses a model's reliability for clinical use in external data sources that have not been used for model development. A lack of external validation can lead to a plethora of proposed models with little evidence about which are reliable and under what circumstances.

In this study, we aim to externally validate existing dementia prediction models. To that end, we define replicability criteria, review published models, and externally validate three selected models using routinely collected health data from administrative claims and electronic health records.

### Materials and Methods

The replicability criteria that a study must report are presented in the following table and were directly inferred from the prediction approach in OHDSI, where among a population at risk, we predict which patients at a defined moment in time (the index) will experience some outcome during a time-at-risk.

| Category | Replicability criteria | Description |
|---|---|---|
| Population settings | Target population definition | Definition or description of the population for which predictions are made. |
| | Index date | Date at which a patient qualifies for inclusion in the target population. |
| | Time-at-risk | Time window in which a model's predictions are valid relative to the index date. |
| | Outcome definition | Definition or description of the outcome to be predicted during the time-at-risk. |
| Statistical analysis settings | Prediction method | Prediction methods in this study are limited to logistic regression and Cox proportional hazard for predicting a binary outcome. |
| | Predictor definitions | Predictor descriptions or definitions in terms of data source codes. |
| | Predictor time window | Time window in which the predictor is assessed. |
| | Model specifications | The prediction model, e.g., parameters to construct the model given a prediction method. We also distinguish here between fully and partially specified models. |

Included dementia prediction studies were reviewed for these criteria to obtain the current state of reporting in the literature. Moreover, we selected three well reported models for replication and external validation in a network of observational databases, with the aim to investigate factors beyond our criteria that may impact successful external validation. These three models will in the remainder of this poster be referred to based on their first author names Walters, Mehta, and Nori, respectively [3-5].

The following databases were selected for external validations of these models as they contain an adequate number of elderly patients.

| Database | Acronym | No. of patients (million) | Country | Data type |
|---|---|---|---|---|
| IBM MarketScan Medicare Suppl. | MDCR | 10 | US | Claims |
| Iqvia Germany Disease Analyzer | IQGER | 30 | DE | GP, EHR |
| Optum Socioeconomic Status | OPSES | 85 | US | Claims |
| Optum Electronic Health Records | OPEHR | 94 | US | EHR |
| Clinical Practice Research Datalink | CPRD | 13 | UK | GP |
| Integrated Primary Care Information | IPCI | 2.5 | NL | GP |
| Iqvia Medical Research Database | IMRD | 18 | UK | GP |

### Results and Discussion

The inclusion criteria of our literature search were met by 35 studies, which described a total of 59 prediction models. The following table summarizes the reporting of our replicability criteria in the included articles.

| Category | Replicability criteria | Reported by no. of models (%) |
|---|---|---|
| Population settings | Target population definition | 59 (100) |
| | Index date | 23 (39) |
| | Time-at-risk | 39 (66) |
| | Outcome definition | 59 (100) |
| Statistical analysis settings | Prediction method | 59 (100) |
| | Predictor definitions | 46 (78) |
| | Predictor time window | 21 (36) |
| | Model specifications: Full model | 8 (14) |
| | Model specifications: Partial model | 19 (32) |

Our results showed that while reporting was complete for some criteria such as target and outcome definitions, reporting of statistical analysis criteria are mostly insufficient to fully replicate the dementia prediction models.

Moreover, our external validation of three selected models (Walters, Mehta, and Nori) showed that even if reporting was sufficient for replication, it did not guarantee that replication and external validation becomes non-trivial, because predictors had to be present, and inclusion and exclusion criteria of target and outcome had to be generalizable to other data sources. Specific problems that we encountered where the following:

- Walters: Uses a "social deprivation score", which ranges from 1 to 5 indicating social deprivation. The information in this variable has been established through a linkage, which is no longer available, or unlikely to exist in other databases across the world.
- Mehta: Does not report a time-at-risk, which was estimated to be 5 years. Also does not provide the baseline hazard so that only a risk stratification model could be replicated rather than the original Cox proportional hazard mode.
- Nori: Does not report a time-at-risk, which was estimated to be 5 years.

Performance across external data sources showed substantial differences in discrimination performance measured as the area under the receiver operating characteristic curve (AUROC) and 95% CI as presented in the following table.

| Model | Internal | MDCR | IQGER | OPSES | OPEHR | CPRD | IPCI | IMRD |
|---|---|---|---|---|---|---|---|---|
| Walters | 0.84 THIN | 0.69 (0.69 – 0.69)* | 0.75 (0.75 – 0.75)* | 0.74 (0.74 – 0.74)* | 0.73 (0.73 – 0.73)* | 0.67 (0.66 – 0.67)* | 0.76 (0.75 – 0.77)* | 0.68 (0.68 – 0.69)* |
| Mehta | 0.81 CPRD | 0.69 (0.69 – 0.70) | 0.72 (0.71 – 0.72) | 0.71 (0.70 – 0.71) | 0.73 (0.73 – 0.73) | 0.79 (0.78 – 0.80) | 0.78 (0.76 – 0.80) | 0.79 (0.78 – 0.80) |
| Nori | 0.69 Optum | 0.66 (0.66 – 0.67) | 0.67 (0.66 – 0.68) | 0.67 (0.66 – 0.68) | 0.62 (0.62 – 0.63) | 0.68 (0.67 – 0.69) | 0.64 (0.62 – 0.67) | 0.68 (0.68 – 0.69) |

We believe that the lack of external validation in dementia prediction literature can to some extend be attributed to the insufficient reporting of models. Models should be developed with external validation in mind. This could for example mean to report all aspects of the model explicitly. Such transparency is best achieved programmatically through code lists and underlying logic rather than literal descriptions, for example by providing a full description of the model (development) in code, ideally against a common data model. This approach will likely eliminate ambiguity as a source of error.

Development choices should not rely on properties unique to the development database, e.g., the Walters model contained criteria to define the target population and predictors that did not exist in the external data sources, for example the cohort entry event "one year following new registration with a THIN practice".

In general, authors should avoid uncommon predictors during model development to guarantee replicability, if the model is meant to be applied in external healthcare settings. Instead of building a single model with multiple, complex cohort entry events, it can be beneficial to build a model for each entry event, which may be easier to interpret and replicate. The Nori model suffered from this problem as it had a complex target population definition with multiple entry events. Defining the time-at-risk window is crucial to indicate in which time window a model's predictions are valid. Using the full follow-up of a population is not a valid approach, as follow-up can vary per person.

### Conclusion

We reviewed 35 studies that proposed a total of 59 dementia risk models. We observed that reporting is mostly insufficient to fully replicate and externally validate published dementia prediction models, and therefore, it is uncertain how well these models would work in other clinical settings. In addition, we replicated and externally validated three existing dementia prediction models and encountered difficulties beyond our replicability criteria, such as ambiguous cohort or predictor definitions. We recommend that reporting should be more explicit and have external validation in mind if the model is meant to be applied in different settings.

**References:**

**1. Hou XH et al. Models for predicting risk of dementia: a systematic review. Journal of Neurology, Neurosurgery and Psychiatry. 2019;90(4):373-9.**
**2. Goerdten J et al. Statistical methods for dementia risk prediction and recommendations for future work: A systematic review. Alzheimer's and Dementia: Translational Research and Clinical Interventions. 2019;5:563-9.**
**3. Walters K, et al. Predicting dementia risk in primary care: development and validation of the Dementia Risk Score using routinely collected data. BMC Med. 2016;14:6.**
**4. Mehta HB et al. Development and validation of the RxDx-Dementia risk index to predict dementia in patients with type 2 diabetes and hypertension. Journal of Alzheimer's Disease. 2016;49(2):423-32.**
**5. Nori VS et al. Identifying incident dementia by applying machine learning to a very large administrative claims dataset. PLoS ONE. 2019;14(7).**

CONTACT: L.JOHN@ERASMUSMC.NL

www.ehden.eu @IMI_EHDEN IMI_EHDEN github.com/EHDEN

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 806968. The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA.

**MONDAY**

## External validation of existing dementia prediction models on observational data

**Lead: Henrik John**

# #OHDSISocialShowcase This Week

***Characteristics and outcomes of inflammatory bowel disease patients: an open, multinational OHDSI network study***


PRESENTER: **Chen** Yanover
KI Research Institute


**INTRO**

- Crohn's disease (CD) and ulcerative colitis (UC) are chronic inflammatory bowel diseases with consistently increasing incidence rates
- These conditions significantly impact the quality of life of patients and families

**METHODS**

**Study design.** A multinational cohort study using routinely collected healthcare data.

**Data sources.**

- IQVIA Medical Research Data, primary care electronic health records (EHRs) from the United Kingdom (**IMRD-UK,** version: 2019-03); over 12.5 million patients, approximately 5% of the UK population
- The Ajou University School of Medicine (**AUSOM**) data, an EHR database from Ajou university hospital in South Korea, containing >2.8 million patients, from 1995 to 2022

**Study population.** IBD cohorts include individuals with at least two diagnoses of IBD or with an IBD diagnosis and a prescription for an IBD medication; CD and UC cohorts further require at least one diagnosis of the corresponding disease and none of the other

**Characteristics and outcomes.**

- Baseline characteristics: during patients' entire history, 1Y, 1M before index date
- Outcomes and treatments: during the 1M, 1Y, 3Y, 5Y, 10Y, following index date and full follow-up time window
- Attributes: OHDSI predefined features (demographics, condition groups, drug era groups), +100 IBD-specific features

## A unified, global characterization of INFLAMMATORY BOWEL DISEASE patient cohorts


Setting the stage for IBD-related predictive and estimation OHDSI network studies

Scan QR to

... download analysis package

... read study protocol

... explore full results

| | Crohn's disease | | Ulcerative colitis | |
|---|---|---|---|---|
| | AUSOM | IMRD-UK | AUSOM | IMRD-UK |
| **n** | 402 | 6,936 | 635 | 13,924 |
| **Females** | 36.6% | 54.4% | 36.2% | 49.1% |
| **Age** | | | | |
| <20 years | 25.9% | 16.3% | 7.9% | 5.8% |
| 20-65 years | 67.3% | 69.8% | 83.2% | 71.3% |
| >65 years | 4.7% | 13.8% | 8.0% | 22.8% |
| **BMI category**[a] | | | | |
| Underweight | 4.0% | 5.6% | 1.9% | 2.7% |
| Normal | 12.7% | 31.5% | 15.7% | 29.3% |
| Overweight | 4.0% | 25.8% | 5.0% | 30.2% |
| Obese | - | 17.4% | - | 18.5% |
| **Medications (index date – 30 days)** | | | | |
| 5-ASA | 65.9% | 43.2% | 84.4% | 67.0% |
| Immunomodulator | 15.4% | 8.2% | 4.7% | 2.0% |
| Corticosteroids[b] | 7.0% | 6.2% | 5.0% | 4.2% |
| Antibiotics | 23.1% | 1.1% | 9.3% | 0.6% |
| **Outcomes (index date – 3 years)** | | | | |
| Anxiety | 10.4% | 6.7% | 13.2% | 6.7% |
| Depression | 4.7% | 9.3% | 5.8% | 8.2% |
| Anemia | 5.5% | 7.6% | 1.7% | 6.6% |
| **Procedures (index date – 3 years)** | | | | |
| S bowel resection | 4.7% | 10.5% | 1.1% | 4.7% |

[a] Underweight: BMI <18.5 kg/m², normal weight: 18.5-25 kg/m², overweight: 25-30 kg/m², obese: >30 kg/m².
[b] Systemic (prednisone, methylprednisolone) and intestinal locally acting (prednisolone, hydrocortisone, prednisone, betamethasone, tixocortol, budesonide, beclometasone) steroids.

**INTERESTED? Join us!**


COLLABORATORS

Ajou University — Jimyung Park, Rae Woong Park, Kwang Jae Lee, Sung Jae Shin
Ben-Gurion University — Nadav Rappoport
Rambam Medical Center — Ramit Magen-Rimon, Yehuda Chowers, Roni Weisshof
KI Research Institute — Tal El-Hay, Maytal Bivas-Benita, Pinchas Akiva


OHDSI

## TUESDAY

## Characteristics and outcomes of inflammatory bowel disease patients: an open, multinational OHDSI network study

**Lead: Chen Yanover**

# #OHDSISocialShowcase This Week

***Mapping PROMs data from the Dutch PROFILES registry to the OMOP CDM - experiences and challenges***

PRESENTER: **Peter** Prinsen

**INTRO**
- Data from the **Netherlands Cancer Registry (NCR)** is being converted to the OMOP-CDM.
- **PROFILES** contains patient-reported outcome measures (**PROMs**) data that is linked to patients in the NCR.
- **How do we add this PROMs data to the OMOP-CDM** to make the data set even more interesting? OHDSI is not very clear on that.

**METHODS**
Use cases we foresee:
- **Specific studies** on PROMs data.
- **General studies** where PROMs data provides additional information about the patient.

**A possible solution** for adding PROMs data to the OMOP-CDM **is outlined in Fig. 1**.

**RESULTS**
We mapped the **EORTC QLQ-C30** and the **Hospital Anxiety and Depression Scale** (HADS, see Fig. 2), and added those to our OMOP-CDM.

They both contain **question/answer (Q/A) pairs** and **scores**. The latter are calculated from subsets of Q/A pairs.

**Examples of mappings** are shown in **Figs. 3 & 4**.

**DISCUSSION**
There are several **open questions, to be answered by the OHDSI community**:

- Should there be an **overarching "PROMs questionnaire filled out" concept** that Q/A pairs and scores can be linked to?
**Yes**, that is a good way to group data from a single filled out questionnaire.
**No**, irrelevant for studies.

- Should **Q/A pairs** be **added to the OMOP-CDM**? Should **scores**?
**Yes**, they are relevant for specific studies.
**No**, only clinical facts and events are relevant.

- Should **Q/A pairs** be **mapped** to clinical concepts?
**Yes**, always.
**Yes**, but only if there are no scores (questions are just a tool to determine scores, the latter are the only relevant outcomes).

## Guideline for adding PROMs to OMOP-CDM nonexistent

## PROMs vocabulary incomplete

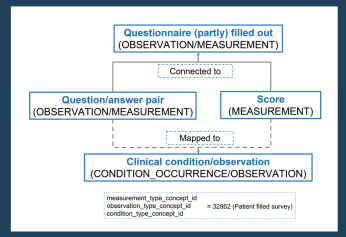



Figure 1. A solution for adding PROMs data to the OMOP-CDM.

| Score | Value | Concept ID | Name | Vocabulary |
|---|---|---|---|---|
| Anxiety score | 0-7 (Normal) | do not map | | |
| | 8-10 (Borderline abnormal) | 4322025 | Mild anxiety | SNOMED |
| | 11-21 (Abnormal) | 441542 | Anxiety | SNOMED |
| Depression score | 0-7 (Normal) | do not map | | |
| | 8-10 (Borderline abnormal) | 40546087 | Depressed mood | SNOMED |
| | 11-21 (Abnormal) | 40546087 | Depressed mood | SNOMED |

Figure 3. Example of score mapping: HADS scores.

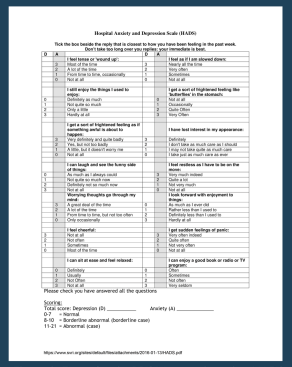



Figure 2. Example of a PROMs questionnaire: HADS.

| Question | Concept ID* | Answer | Concept ID** | Mapped question/answer | Concept ID*** |
|---|---|---|---|---|---|
| Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase? | 701340 | not at all | 45883172 | Capable of heavy physical activity | 763659 |
| | | a little | 45876949 | ? | ? |
| | | quite a bit | 45884456 | ? | ? |
| | | very much | 45885256 | ? | ? |
| Do you have any trouble taking a long walk? | 701341 | not at all | 45883172 | Able to walk | 4086870 |
| | | a little | 45876949 | Unable to walk long distances | 44792042 |
| | | quite a bit | 45884456 | Unable to walk long distances | 44792042 |
| | | very much | 45885256 | Unable to walk long distances | 44792042 |

* OMOP Extension vocabulary, ** LOINC vocabulary, *** SNOMED vocabulary

Figure 4. Example of question and answer concepts, and mapping of Q/A pairs: the first two questions of QLQ-C30.

- Should **scores** be **mapped** to clinical concepts?
**Yes**, that is the clinically most important part.
**No**?

- Should we add **negative concepts** (no pain)?
**Yes**, they are relevant observations.
**No**, this information is not useful in studies.

- Can we use ...type_concept to indicate **patient-reported** (vs **physician-reported**)?
**Yes**, we do not want to create a whole bunch of new concepts.
**No**, nobody uses the ...type_concept field in analyses.

- What **observation or event date** do we associate with the PROMs data ("Did you experience pain in the past month?")?

**CONCLUSIONS**
- A **guideline for adding PROMs to the OMOP-CDM does not exist**: the OHDSI community should develop conventions for capturing PROMS data in OMOP.
- The **vocabulary is very incomplete** when it comes to representing PROMS data.

**JOIN US!**
Do you want to **collaborate on harmonizing PROMS data** in OMOP? Then contact Sebastiaan and join the **EHDEN PROMS WG** (sebastiaan.van.sandijk@odysseusinc.com)


Peter Prinsen[1]*,
Chiara Attanasio[1],
Corina van den Hurk[1],
Nicole Horevoorts[1],
Sebastiaan van Sandijk[2]

[1]IKNL,
[2]Odysseus Data Services,
*p.prinsen@iknl.nl


IKNL integraal kankercentrum Nederland · profiles · EHDEN · ODYSSEUS DATA SERVICES INC · OHDSI

# WEDNESDAY

## Mapping PROMs data from the Dutch PROFILES registry to the OMOP CDM - experiences and challenges

## Lead: Peter Prinsen

# #OHDSISocialShowcase This Week

***OHDSI Germany***
*A recap after one year*
PRESENTER: Michele Zoch

**INTRO**
- Establishing OHDSI Germany as a multi-stakeholder interest group in Spring 2021
- Goals:
  1. Establishing a German research network of hospitals on OMOP for OHDSI
  2. Creating and sharing best practices as well as deliverables
  3. Involving German stakeholder and providing assistance with getting started in the OHDSI community
  4. Collaborating for administrative work in terms of German data security laws and ethics

**METHODS**
- Monthly community meetings
- Presentations and discussions
- Sharing best practices in plenary

**RESULTS**
- Formation of an interdisciplinary interest group
- Identifying common topics and synergy effects
- Offering workshops and tutorials

**OUTLOOK**
- Onboarding of further participants in Germany
- Exploiting further data sources and terminologies
- Extending of ETL processes
- Providing a technology stack (esp. German OMOP Stack, ETL jobs)
- Participating in DARWIN
- Collaborating in newly applied projects of the German Medical Informatics Initiative
- Working together to overcome administrative hurdles
- Intensify joint work

**OHDSI Germany** is one of the first contact points for **sharing best practices** and **experiences** with a focus on **German patient data.**

Networking Sharing Adapting Collaborating

**PREVIOUS TOPICS**
- Development of a German specific OMOP Stack and ETL jobs
- Strategies for using OMOP for rare diseases
- Integration of federally-mandated Medical Information Objects
- Effects of pre- and post-coordination on mapping

**RELATED LINKS**

Michele ZOCH[1], Elisa HENKE[1], Yuan PENG[2], Najia AHMADI[1], Joshua WIEDEKOPF[2], Mareike PRZYSUCHA[3], Josef SCHEPERS[4], Martin SEDLMAYR[1], Ines REINECKE[1]
[1] Institute for Medical Informatics and Biometry at Carl Gustav Carus Faculty of Medicine, Technische Universität Dresden, Germany
[2] IT Center for Clinical Research, University of Lübeck, Germany
[3] Health Informatics Research Group, University AS Osnabrück, Germany
[4] Berlin Institute of Health, Charité – Universitätsmedizin Berlin, Germany

TECHNISCHE UNIVERSITÄT DRESDEN OHDSI

## THURSDAY

## OHDSI Germany: A recap after one year

**Lead: Michele Zoch**

# #OHDSISocialShowcase This Week

***RCTrep: An R package for the validation of methods for treatment effect estimation using real-world data***

PRESENTER: Lingjie Shen[1*]

**INTRO:**

- **Who cares?** – policy makers; regulators; real world evidence (RWE) evaluators.
- **Why?** There is an increasing attention for the leverage of large real-world data (RWD) in treatment effect estimation to drive fast and precise decision making.
- **Challenge:** Since we do not observe the true treatment effect for each individual- which is the fundamental problem of causal inference - validation of treatment effect estimation methods using RWD is challenging.
- **Aim:** In the absence of a ground truth, how can we validate different methods using RWD to select the most reasonable method for the data at hand, driving fast regulatory and clinical decision making?

**METHODS:**

- We identify under which conditions the estimate from randomized control trial (RCT) can be regarded as the ground truth for methods validation using RWD. We illustrate differences between RCT and RWD in Figure 1. We assume the RWD and RCT data are two random samples from a, potentially different, population, and hence allow for a fair comparison of estimates of treatment effect between two samples after population composition is controlled for.
- We consider a set of candidate treatment effect estimators $\mathcal{F} = \{f_1, \ldots, f_m\}$, where $f(\boldsymbol{x})$: $\mathcal{X} \mapsto \mathrm{E}[Y(1) - Y(0) \mid \boldsymbol{X} = \boldsymbol{x}]$, $f(\boldsymbol{x})$ is an estimator of conditional average treatment effect of population with characteristics $\boldsymbol{X} = \boldsymbol{x}$. We select the best one using the following evaluation metric:

$$f^* = \mathrm{argmin}_{f \in \mathcal{F}} \mathrm{L}(\hat{\tau}; f) = \mathrm{argmin}_{f \in \mathcal{F}} \left( \hat{\tau} - \sum_x w(\boldsymbol{x}) f(\boldsymbol{x}) \right)^2, \quad s.t.\ p(\boldsymbol{x}) = q(\boldsymbol{x}) w(\boldsymbol{x})$$

where $\hat{\tau}$ is an unbiased estimate of average treatment effect of a population that a RCT represents, $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$ are the empirical density of $\boldsymbol{x}$ in RCT data and RWD, $w(\boldsymbol{x})$ is a weight for individuals in RWD with characteristics $\boldsymbol{X} = \boldsymbol{x}$.

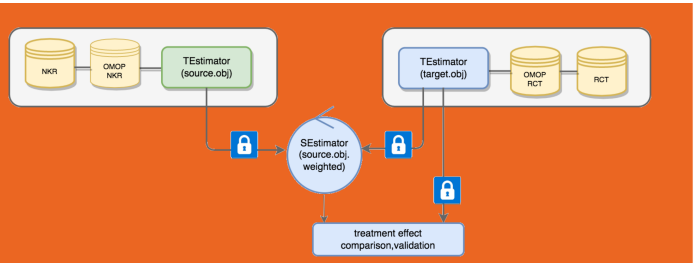
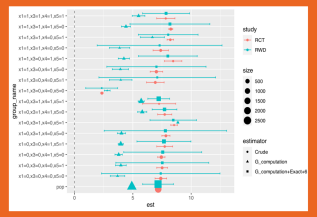



Figure 2: Diagram of RCTrep basic structure

- **TEstimator:** R6 class TEstimator is responsible for **estimating population- and subpopulation-level treatment effects**, and diagnosing assumptions.
- **SEstimator:** R6 class SEstimator is responsible for computing **weights**, so that the weighted covariates in **source.obj** and covariates in **target.obj** are **balanced**. **The two objects communicate within** the object of the class SEstimator, sharing either **unit-level data** or **aggregated data** for computing the weights.



Paper Package

Figure 3: A working example of RCTrep. We use the **G-computation method** to adjust the **treatment assignment mechanism,** and use **exact matching** to adjust the **sampling mechanism.** Results show that **only** correcting for **both mechanisms** can allow for **comparison** of treatment effect estimation between **RWD** and **RCT** data.

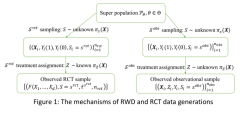


Figure 1: The mechanisms of RWD and RCT data generations

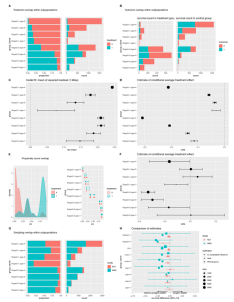


Figure 4: Estimates comparison between NKR and QUASAR trial using RCTrep. Subfigure (a)(b) diagnoses overlap of treatment within subgroups in NKR data and survival in treatment and control groups in NKR data. Figure (c)(d) diagnoses G-computation model fit and estimates of treatment effect in subgroups. Figure (e) (f) diagnoses propensity score overlap between treatment and control groups and estimates of treatment effect using inverse propensity score weighting. Figure (d)(h) diagnoses covariates balance between KNR and QUASAR trial and comparison of estimates from QUASAR and estimates from KNR with and without weighting.

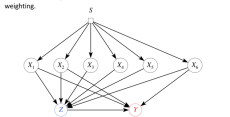


Figure 5: Illustration of adjustment sets in TEstimator and SEstimator. S is an indicator of selection into RCT and Z is an indicator of selection into treatment group.

Lingjie Shen[1*], Gijs Geleijnse[2], Maurits Kaptein[3]

[1]Department of Methodology and Statistics, Tilburg University,
[2]IKNL,
[3]Jheronimus Academy of Data Science,
[*]L.Shen@uvt.nl

JADS IKNL OHDSI

## FRIDAY

## A dashboard for visual comparison of OMOP CDM databases

**Leads: João Almeida, José Oliveira**

# 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?

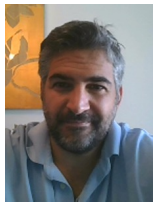
## Where Are We Going?

# Sept. 6: OHDSI Publications

**A database of pediatric drug effects to evaluate ontogenic mechanisms from child growth and development**
**Presenter:** Nick Giangreco

**Development and external validation of prediction models for adverse health outcomes in rheumatoid arthritis: A multinational real-world cohort analysis**
**Presenter:** Cynthia Yang

**Empirical assessment of alternative methods for identifying seasonality in observational healthcare data**
**Presenter:** Anthony Molinaro



**Phenotype Algorithms for the Identification and Characterization of Vaccine-Induced Thrombotic Thrombocytopenia in Real World Data: A Multinational Network Cohort Study**
**Presenter:** Azza Shoaibi

**TreatmentPatterns: An R package to facilitate the standardized development and analysis of treatment patterns across disease domains**
**Presenter:** Aniek Markus

Figure 4 **A**

# A database of pediatric drug effects using ATC and MedDRA

Giangreco NP, Tatonetti NP. A database of pediatric drug effects to evaluate ontogenic mechanisms from child growth and development. Med (N Y). 2022 Aug 12;3(8):579-595.e7. doi: 10.1016/j.medj.2022.06.001. Epub 2022 Jun 24. PMID: 35752163; PMCID: PMC9378670.

- We generated 500,000 drug safety signals for every drug (ATC5) and adverse event (PT) identified during childhood
- We compared drug safety signals to a null distribution to estimate statistical significance for each pediatric ADE
- We *grouped* ADEs into the ATC1-5 and PT-SOC categories
- We asked if the drug or adverse event class was enriched for significant ADEs at a stage, and the matrix shows how many
- The graphs show select drug ingredients enriched in systemic adverse events across childhood

| | ATC1 N=14 | ATC2 N=77 | ATC3 N=166 | ATC4 N=358 | ATC5 N=926 | Number of associated adverse event classes |
|---|---|---|---|---|---|---|
| SOC N=27 | 43 | 66 | 60 | 38 | 36 | 243 |
| HLGT N=334 | 23 | 15 | 9 | 3 | 2 | 52 |
| HLT N=1,472 | 3 | 3 | 2 | 2 | 1 | 11 |
| PT N=6,841 | 0 | 0 | 0 | 0 | | 0 |
| Number of associated drug classes | 69 | 84 | 71 | 43 | 39 | 306 Number of associated ADE classes |

**B**