Building Organizational Support Within Your Community

OHDSI Community Call
Aug. 2, 2022 • 11 am ET

OHDSI
Observational Health Data Sciences and Informatics

Upcoming OHDSI Community Calls

Date	Topic
Aug. 9	Around The Asia-Pacific (APAC) Community
Aug. 16	OHDSI “Speed Dating”
Aug. 23	Workgroup Updates
Aug. 30	EHDEN Academy/EHDEN Portal
Sept. 6	OHDSI Studies
Sept. 13	Registries And Their Adoption To OMOP
Sept. 20	OHDSI2022 Preview
Sept. 27	HTA Challenge

Upcoming OHDSI Community Calls

Date	Topic
Aug. 9	Around The Asia-Pacific (APAC) Community
Aug. 16	OHDSI “Speed Dating”
Aug. 23	Workgroup Updates
Aug. 30	EHDEN Academy/EHDEN Portal
Sept. 6	OHDSI Studies
Sept. 13	Registries And Their Adoption To OMOP
Sept. 20	OHDSI2022 Preview
Sept. 27	HTA Challenge

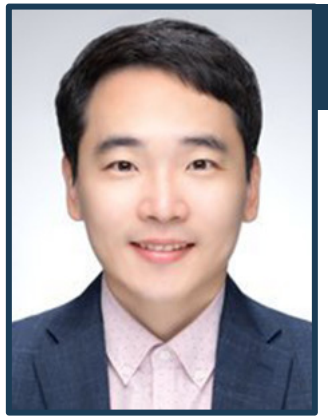
Aug. 9: Asia-Pacific (APAC) Regional Updates

Jason Hsu

Taiwan Chapter

Nicole Pratt

Australia Chapter



Seng Chan You

Korea Chapter

Mengling Feng

Singapore Chapter

Tatsuo Hiramatsu

Japan Chapter

Lei Liu

China Chapter

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **Karoline Bräuner, Andreas W. Rosen, Adamantia Tsouchnika, Julie S. Walbech, Mikail Gögenur, Viviane A. Lin, Johan Clausen, and Ismail Gögenur** on the publication of **Developing prediction models for short-term mortality after surgery for colorectal cancer using a Danish national quality assurance database** in the International Journal of Colorectal Disease.

International Journal of Colorectal Disease
https://doi.org/10.1007/s00384-022-04207-6

RESEARCH

Developing prediction models for short-term mortality after surgery for colorectal cancer using a Danish national quality assurance database

Karoline B. Bräuner[1] · Andreas W. Rosen[1] · Adamantia Tsouchnika[1] · Julie S. Walbech[1] · Mikail Gögenur[1] · Viviane A. Lin[1] · Johan S. R. Clausen[1] · Ismail Gögenur[1,2]

Accepted: 20 June 2022

Abstract

Purpose The majority of colorectal cancer surgeries are performed electively, and treatment is often decided at the multidisciplinary team conference. Although the average 30-day mortality rate is low, there is substantial population heterogeneity from young, healthy patients to frail, elderly patients. The individual risk of surgery can vary widely, and tailoring treatment for colorectal cancer may lead to better outcomes. This requires prediction of risk that is accurate and available prior to surgery.

Methods Data from the Danish Colorectal Cancer Group database was transformed into the Observational Medical Outcomes Partnership Common Data Model. Models were developed to predict the risk of mortality within 30, 90, and 180 days after colorectal cancer surgery using only covariates decided at the multidisciplinary team conference. Several machine-learning models were trained, but due to superior performance, a Least Absolute Shrinkage and Selection Operator logistic regression was used for the final model. Performance was assessed with discrimination (area under the receiver operating characteristic and precision recall curve) and calibration measures (calibration in large, intercept, slope, and Brier score).

Results The cohort contained 65,612 patients operated for colorectal cancer in the period from 2001 to 2019 in Denmark. The Least Absolute Shrinkage and Selection Operator model showed an area under the receiver operating characteristic for 30-, 90-, and 180-day mortality after colorectal cancer surgery of 0.871 (95% *CI*: 0.86–0.882), 0.874 (95% *CI*: 0.864–0.882), and 0.876 (95% *CI*: 0.867–0.883) and calibration in large of 1.01, 0.98, and 1.01, respectively.

Conclusion The postoperative short-term mortality prediction model showed excellent discrimination and calibration using only preoperatively known predictors.

Keywords Colorectal cancer · Machine learning · Prediction model · Mortality · Postoperative

OHDSI Shoutouts!

Congratulations to the team of **Carina Nina Vorisek, Moritz Lehne, Sophie Anne Ines Klopfenstein, Paula Josephine Mayer, Alexander Bartschke, Thomas Haese, and Sylvia Thun** on the publication of **Fast Healthcare Interoperability Resources (FHIR) for Interoperability in Health Research: Systematic Review** in JMIR Medical Informatics.

JMIR Publications
Advancing Digital Health & Open Science

Articles | Search articles

JMIR Medical Informatics | Journal Information | Browse Journal

Published on 19.7.2022 in Vol 10, No 7 (2022): July

Preprints (earlier versions) of this paper are available at https://preprints.jmir.org/preprint/35724, first published December 29, 2021.

Fast Healthcare Interoperability Resources (FHIR) for Interoperability in Health Research: Systematic Review

Carina Nina Vorisek [1]; Moritz Lehne [1]; Sophie Anne Ines Klopfenstein [1,2]; Paula Josephine Mayer [1]; Alexander Bartschke [1]; Thomas Haese [1]; Sylvia Thun [1]

Article | Authors | Cited by | Tweetations (3) | Metrics

- Abstract
- Introduction
- Methods
- Results
- Discussion
- References
- Abbreviations
- Copyright

Abstract

Background:
The standard Fast Healthcare Interoperability Resources (FHIR) is widely used in health information technology. However, its use as a standard for health research is still less prevalent. To use existing data sources more efficiently for health research, data interoperability becomes increasingly important. FHIR provides solutions by offering resource domains such as "Public Health & Research" and "Evidence-Based Medicine" while using already established web technologies. Therefore, FHIR could help standardize data across different data sources and improve interoperability in health research.

Objective:
The aim of our study was to provide a systematic review of existing literature and determine the current state of FHIR implementations in health research and possible future directions.

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

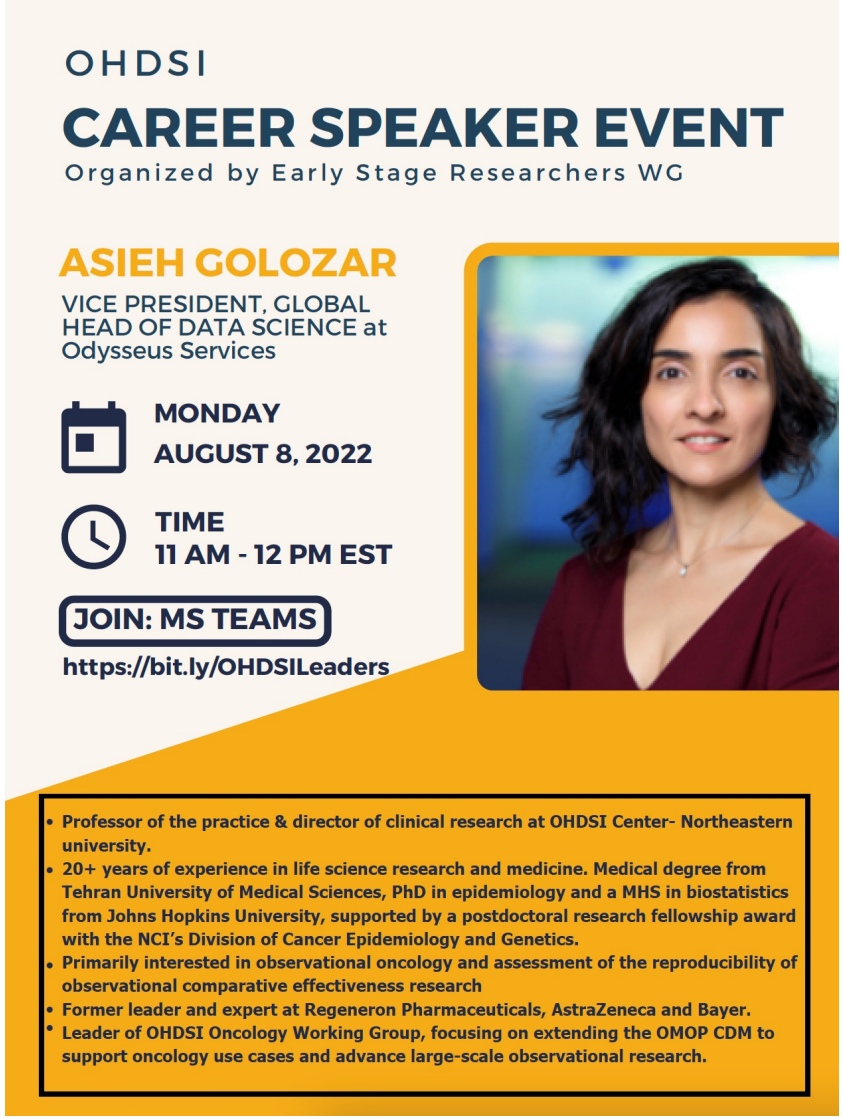
Date	Time (ET)	Meeting
Wednesday	2 am	Patient-Level Prediction/Population-Level Estimation
Wednesday	9 am	Psychiatry
Wednesday	10 am	FHIR and OMOP Digital Quality Measurements Subgroup (ZOOM)
Wednesday	12 pm	Health Equity
Wednesday	4 pm	FHIR and OMOP Data Model Harmonizations Subgroup (ZOOM)
Thursday	10 am	Medical Devices
Thursday	12 pm	Patient-Level Prediction/Population-Level Estimation
Thursday	12 pm	FHIR and OMOP Oncology Subgroup
Thursday	6 pm	FHIR and OMOP Terminologies Subgroup (ZOOM)
Friday	9 am	GIS-Geographic Information System Development
Friday	10:30 am	Clinical Trials
Monday	11 am	Early-Stage Researchers Speaker Series

www.ohdsi.org/upcoming-working-group-calls

Early-Stage Researcher Speaker Series

The first edition of the Early-Stage Researcher Career Speaker Series will take place Monday, Aug. 8 (11 am ET), with **Asieh Golozar** set to discuss her career in life science research and medicine, and how her OHDSI journey has impacted her career.

bit.ly/OHDSILeaders



Latest OHDSI Newsletter Is Available

Community Updates

Where Have We Been?

• Clair Blacketer and the CDM Workgroup team recently led a presentation about a new CDM Update Process, which includes a new decision tree and process that is implemented to both streamline the procedure and clarify how data model requests are made and codified. You can watch that presentation now.

• The Scientific Review Committee met recently to discuss a record-setting amount of submissions to the #OHDSI2022 Collaborator Showcase. If you sent in a brief report, you should hear soon whether your submission has been accepted.

Where Are We Now?

• Numerous peer-reviewed studies that include OHDSI tools and/or practices were published in July and shared later in this newsletter. The new OHDSI publication dashboard tracks community publications, citations, new authors and more, and is available on the front page of OHDSI.org.

• Nominations are OPEN for the 2022 Titan Awards, which recognize those who have gone above and beyond to foster community engagement, lead research and development efforts, and make significant contributions towards OHDSI's mission. If there are members or institutions who have made significant contributions that you would like to recognize, please nominate them before the Sept. 2 deadline!

August Video Podcast

In the latest On The Journey video, Patrick Ryan and Craig Sachson discuss the recent CDM Update Process presentation and its potential impact on research, the variety of studies coming out of the community in July, and the opening of Titan Award nominations in preparation for the 2022 Global Symposium.

OHDSI

The Journey Newsletter (August 2022)

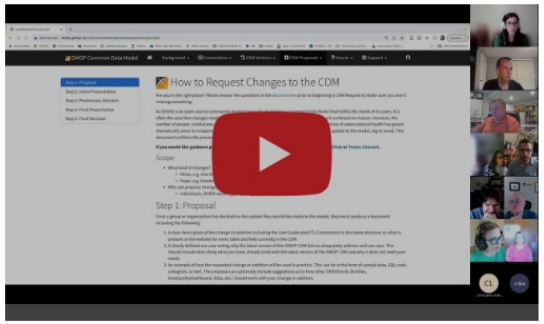
The OHDSI2022 Global Symposium in Bethesda, Md., is getting closer, and there are new details on the event. A new CDM Update Process was unveiled this month to help you figure out what to do when your data does not fit into the OMOP CDM. All that and numerous community updates are shared in our latest newsletter. **#JoinTheJourney**

New CDM Process Includes Decision Tree, Clarifies How New Requests Are Made, Codified

The July 26 community call featured a session led by **Clair Blacketer, Paul Nagy and Davera Gabriel** on what to do when your data does not fit into the OMOP CDM.

Our community continues to expand globally, and both individuals and organizations often look for new enhancements to the CDM. There will be a new decision tree and process implemented to try and streamline this procedure and to clarify how data model requests are made and codified, and these were presented and discussed during this meeting.

The presentation can be seen below, and it includes specific use cases presented by both Nagy and Gabriel to help show how the decision tree and process happens.



Start The CDM Request Decision Tree

July Publications

Schuemie Martijn J., Arshad Faaizah, Pratt Nicole, Nyberg Fredrik, Alshammari Thamir M, Hripcsak George, Ryan Patrick, Prieto-Alhambra Daniel, Lai Lana Y. H., Li Xintong, Fortin Stephen, Minty Evan, Suchard Marc A. Vaccine Safety Surveillance Using Routinely Collected Healthcare Data—An Empirical Evaluation of Epidemiological Designs. *Frontiers in Pharmacology.* Vol. 13. 2022. DOI: 10.3389/fphar.2022.893484

Molinaro, A., DeFalco, F. Empirical assessment of alternative methods for identifying seasonality in observational healthcare data. *BMC Med Res Methodol* 22, 182 (2022). DOI: 10.1186/s12874-022-01652-3

Delanerolle G, Williams R, Stipancic A, Byford R, Forbes A, Anand S, Bradley D, Tsang R, Murphy S, Akbari A, Bedston S, Lyons R, Owen R, Beggs J, Chuter A, Balharry D, Joy M, Sheikh A, Hobbs FR, de Lusignan S. Methodological Issues in Using a Common Data Model of COVID-19 Vaccine Uptake and Important Adverse Events of Interest: Feasibility Study of Data and Connectivity COVID-19 Vaccines Pharmacovigilance in the United Kingdom. *JMIR Formative Research.* 17/05/2022:37821 (forthcoming/in press). PMID: 35786634.

Voss, E.A., Ali, S.R., Singh, A. *et al.* Hip Fracture Risk After Treatment with Tramadol or Codeine: An Observational Study. *Drug Saf* 45, 791–807 (2022). https://doi.org/10.1007/s40264-022-01198-9

Ines Reinecke, Mirko Gruhl, Martin Pinnau, Fatma Betül Altun, Michael Folz, Michéle Zoch, Franziska Bathelt, Martin Sedlmayr. An OHDSI ATLAS Extension to Support Feasibility Requests in a Research Network. *Studies in Health Technology and Informatics, Vol. 295.* doi: 10.3233/SHTI220778

Emily Pfaff, Andrew Girvin, Tellen Bennett, Abhishek Bhatia, Ian Brooks, Rachel Deer, Jonathan Dekermanjian, Sarah Elizabeth Jolley, Michael Kahn, Kristin Kostka, Julie McMurry, Richard Moffitt, Anita Walden, Christopher Chute, Melissa A Haendel, and the N3C Consortium. Identifying who has long COVID in the USA: a machine learning approach using N3C data. *The Lancet Digital Health.* 2022. DOI: 10.1016/S2589-7500(22)00048-6

Tak YW, You SC, Han JH, Kim SS, Kim GT, Lee Y. Perceived Risk of Re-Identification in OMOP-CDM Database: A Cross-Sectional Survey. *J Korean Med Sci.* 2022 Jul;37(26):e205. DOI: 10.3346/jkms.2022.37.e205

Nicholas P. Giangreco, Nicholas P. Tatonetti, A database of pediatric drug effects to evaluate ontogenic mechanisms from child growth and development, *Med,* 2022, ISSN 2666-6340, DOI: 10.1016/j.medj.2022.06.001

Phuong J, Hong S, Palchuk MB, Espinoza J, Meeker D, Dorr DA, Lozinski G, Madlock-Brown C, Adams WG. Advancing Interoperability of Patient-level Social Determinants of Health Data to Support COVID-19 Research. *AMIA Annu Symp Proc.* 2022 May 23;2022:396-405. PMID: 35854720; PMCID:

Latest OHDSI Newsletter Is Available

OHDSI APAC Community Calls

The most recent OHDSI APAC community call provided updates on two of the four focus studies in 2022: **Comparison of mortality, morbidities & healthcare resources utilization between patients with and without a diagnosis of COVID-19**, and **Real world safety of treatments for multiple sclerosis**.

Date	Topic
July 14	APAC Study Quarterly Updates, Part 1
July 28	APAC Study Quarterly Updates, Part 2
Aug. 11	Working Group Updates #3
Aug. 25	Working Group Updates #4
Sept. 8	EU Chapter Sharing Session, Part 1
Sept. 22	EU Chapter Sharing Session, Part 2

ohdsi.org/APAC

Titan Awards Nominations Are Open

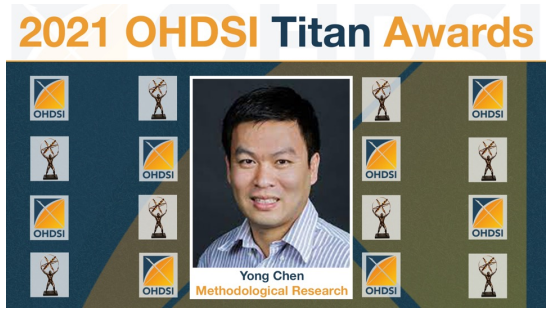
Nominations for the 2022 Titan Awards are now OPEN!
Please use the form below to nominate an individual or institution for a top contribution to the OHDSI community this past year!

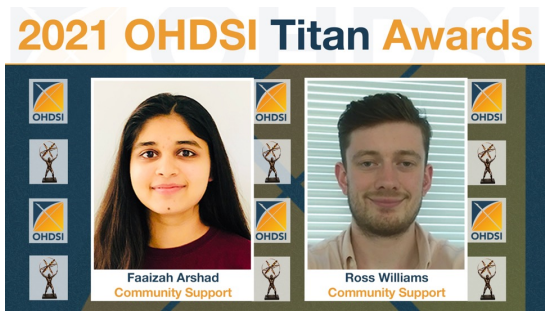
2022 Nomination Form

To recognize OHDSI collaborators (or collaborating institutions) for their contributions towards OHDSI's mission, the OHDSI Titan Awards were introduced at the 2018 Symposium and have been handed out at the U.S./Global Symposium each year since. Annually, community members are invited to nominate individuals or institutions they feel have made significant contributions towards advancing OHDSI's mission, vision and values. Once nominations are submitted, the OHDSI Titan Award Committee will select the award winners. Award winners will be announced before the networking reception at the annual symposium. The award categories, as well as all previous recipients, can be found below.

Titan Award for Data Standards – to recognize extraordinary contributions by an individual, organization, or team in development or evaluation in community data standards, including OMOP common data model and standardized vocabularies

- **2021 –** Maxim Moinat, The Hyve/Erasmus University Medical Center
- **2020 –** Clair Blacketer, Janssen Research and Development
- **2019 –** Oncology Workgroup (Michael Gurley, Northwestern Univ.; Rimma Belenkaya, Memorial Sloan Kettering Cancer Center; Robert Miller, Tufts CTSI)
- **2018 –** Vocabulary team (Christian Reich, IQVIA; Anna Ostropolets, Columbia Univ.; Dmitry Dymshyts, Odysseus Data Services)





ohdsi.org/titan-awards

Job Openings

Professor **Dani Prieto-Alhambra** and his team at the University of Oxford will be hiring two Research Assistants in Health Data Sciences.

The application deadline is August 8, 2022.

The link and more information will be available on the community calls page.

UNIVERSITY OF OXFORD

UK date and time: 11-July-2022 16:46

Applicant Options

- New Search
- Login
- Job Details
- Help
- Terms of Use & Privacy Policy

corehr

Job Details

Research Assistant in Health Data Sciences (2 posts)

Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, Windmill Road, Oxford

We have an exciting opportunity for an enthusiastic and dedicated Research Assistants in Health Data Sciences to join the Pharmaco- and Device epidemiology research group lead by Professor Daniel Prieto-Alhambra at the Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), Oxford.

As a Research Assistant in Health Data Sciences you will support the programming of analytical pipelines for the analysis of routinely collected data mapped to the OMOP Common Data Model. You will prepare analytical packages to run a number of pre-specified analyses, contribute to wider project planning, including ideas for new research projects and manage own research and administrative activities, within guidelines provided by senior colleagues.

You will hold a relevant post-graduate degree in Mathematics, Engineering, Health Data Sciences or Biostatistics. You will have an experience in biostatistics as well as experience in analysis of OMOP-mapped data. Knowledge of medical statistics and expertise in handling large patient level datasets, good knowledge of programming in R packages for statistical analyses and ability to communicate results effectively with colleagues in any discipline are essential. Expertise in pharmaco and/or vaccine epidemiology, experience working with electronic medical records/routinely collected data and experience of working within an academic environment are desirable.

This is a full-time fixed-term appointment for 2 years.

The closing date for this position is 12 noon on Monday 08/08/2022. You will be required to upload a CV and supporting statement as part of your online application.

Contact Person :	HR Team, NDORMS	**Vacancy ID :**	159236
Contact Phone :		**Closing Date & Time :**	08-Aug-2022 12:00
Pay Scale :	STANDARD GRADE 6	**Contact Email :**	hr@ndorms.ox.ac.uk
Salary (£) :	£29,614 to £36,326 p. a.		

Click on the link(s) below to view documents	Filesize
159236_JD	472

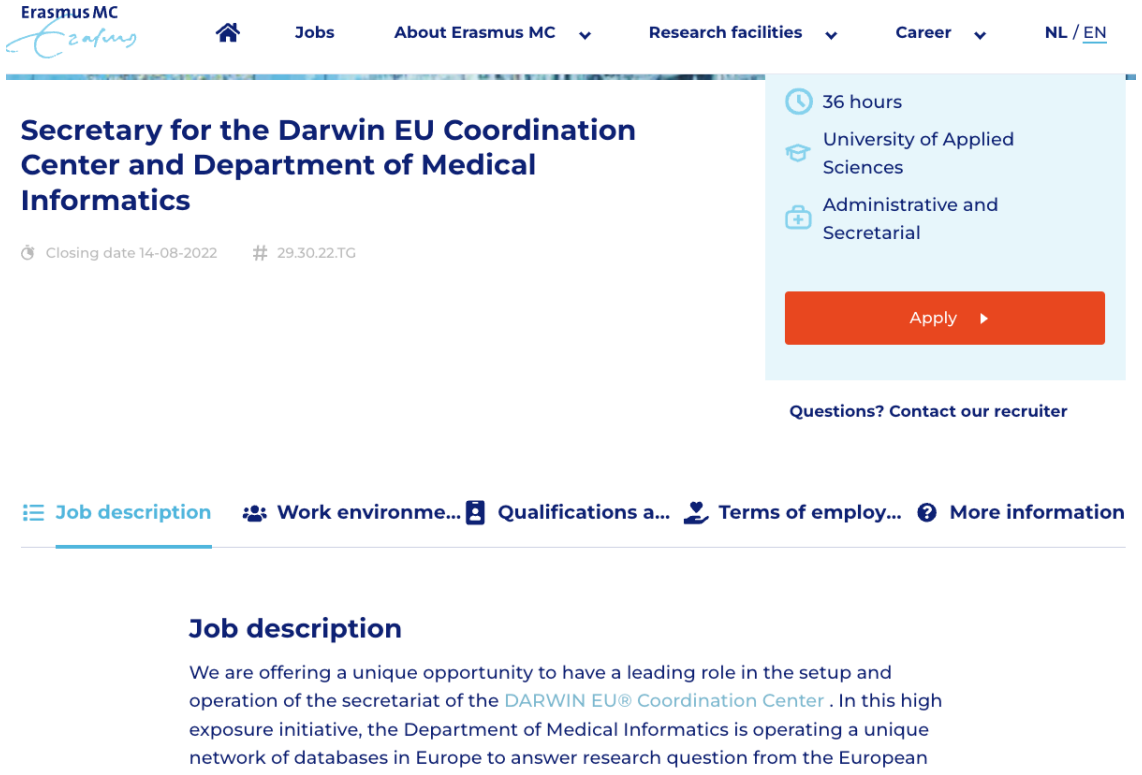
Return to Search Results | Apply Now

Job Openings

Peter Rijnbeek and his team at Erasmus University is hiring a Secretary for the Darwin EU Coordination Center and Department of Medical Informatics.

This position will be responsible for the day-to-day administrative tasks as the personal assistant for Peter Rijnbeek, and will also work as senior secretary for the Department of Medical Informatics.

The application deadline is Aug. 14.



Job Openings

Assistant professor **Brianne Oliveri-Mui** announced an opening for an Postdoctoral Fellow to work at the Roux Institute at Northeastern University.

If you are interested, please reach out to Dr. Mui at b.mui@northeastern.edu.

The link and more information will be available on the community calls page.

Observational Health Data Sciences and Informatics Postdoctoral Fellow

Apply

Portland, ME

Full time

Posted 30+ Days Ago

R105484

About the Opportunity

The Roux Institute at Northeastern University has one opening for a Postdoctoral Research Fellow beginning on or about September 1, 2022. The fellow will have an opportunity to conduct observational and administrative database research (e.g., analysis of existing datasets) on health outcomes for older adults with HIV or LGBT older adults, under the supervision of the PI. The fellow
will devote most of their time to independent research aligned with the PI's interests and across federated and local research models.
Position offers exceptional opportunity for collaboration at the OHDSI center on major projects in the U.S. and overseas. This research will directly improve our ability to use real world data to characterize under-represented and marginalized patient populations, construct population level estimates relating exposures to health outcomes, and to enhance clinical decision making through improved patient-level predictions. The term of fellowship appointment will be for two years, contingent on
continued funding. Stipend will be commensurate with experience, based on levels mandated by NIH.

The main research areas specific to older people with HIV or in the LGBTQ+ communities are as follows:

- Measurement of comorbidities, care quality, health outcomes and healthcare utilization patterns
- Risk assessment of multimorbidity, healthcare and prescription access

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

An Introductory Journey From Data To Evidence

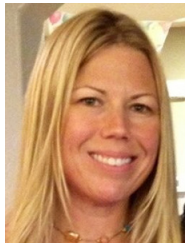
OHDSI2022 Tutorial • Saturday, Oct. 15 • Bethesda, Md.

The OHDSI Journey: Where Are We Going?

Patrick Ryan

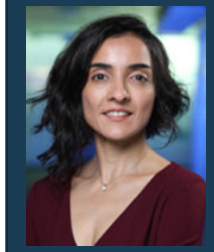
OMOP Common Data Model and Vocabulary

Clair Blacketer



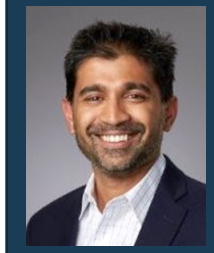
ETL – A Source Database Into OMOP CDM

Melanie Philofsky



Creating Cohort Definitions

Asieh Golozar



Phenotype Evaluations

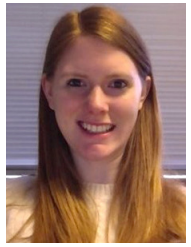
Gowtham Rao

Characterization

Kristin Kostka

Estimation

Martijn Schuemie



Prediction

Jenna Reps

The OHDSI Journey: Where Do We Go From Here?

George Hripcsak

Workgroup Activities

Saturday, Oct. 15, and Sunday, Oct. 16

Saturday, Oct 15					
Start Time (ET)	End Time (ET)				
800	900	Tutorial	HADES Hack-a-thon: Part 1	Oncology WG	FHIR-OMOP: Terminologies Subgroup, Part 1
900	1000				FHIR-OMOP: Increasing the Value of Data Through a Rich Set of Attributes
1000	1100				
1100	1200				
1200	1300		Lunch	Lunch	Lunch
1300	1400		Methods Research (PLE/PLP)	Oncology WG (continued)	FHIR-OMOP: Data Model Harmonization Subgroup
1400	1500				
1500	1600			Natural Language Processing	FHIR-OMOP: Oncology Subgroup
1600	1700				
1700	1800				FHIR-OMOP: Terminologies Subgroup, Part 2
1800	1900				
Sunday, Oct 16					
800	900	All-Hands Workgroup Meeting			
900	1000				
1000	1100				
1100	1200				
1200	1300	Lunch		Lunch	Lunch
1300	1400	Phenotype Evaluation	HADES Hack-a-thon: Part 2	Education	CDM and Data Quality
1400	1500				
1500	1600			Health Equity	
1600	1700				

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

Comparing Data Quality Dashboard results from two ETL iterations: three new utilities

PRESENTER: **Anne van Winzum**

Background

The Data Quality Dashboard (DQD) has been widely used to **evaluate the quality of an OMOP CDM** data set resulting from an ETL (extract, transform, load) process[1]. In practice, during the conversion to the OMOP CDM we perform several ETL iterations. However, interpreting the differences in quality is not always straightforward.

We developed three new utilities as part of mapping of the **UK Biobank** (UKB) data under the European Health Data Evidence Network (EHDEN) COVID19 rapid data partner call[2], and in collaboration with University College London.

Methods

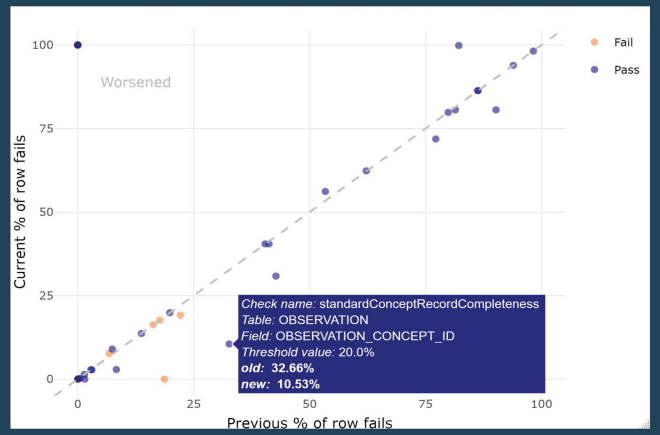
Thresholds editing

As part of the ETL iterations, we needed to change the fail-thresholds of individual checks. We created a separate table to **list the changed thresholds in a user-friendly way**. Our utility script takes this new table to produce a customized thresholds file accepted by the default DQD scripts.

Comparison of DQD results

This visualization script selects the checks for which the percentage of records that satisfy said check has changed between ETL iterations. Here, the percentage of records satisfying the checks had modestly improved (Figure 1). As an example, there is an outlier (top left corner) that prompted us to investigate and update the ETL accordingly.

Visualising and comparing DQD results is an important step to interpret the data quality and to find actionable data quality issues.



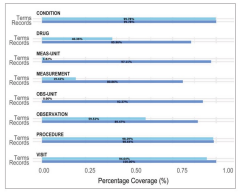
▲ *Figure 1: Each dot represents one check that has a different percentage of row fails between the iterations. This percentage of row fails is marked in the x-axis for the earliest run, and in the y-axis for the latest run.*

Scan QR to see our blogpost

On the other hand, we improved the standard record completeness in the observation table to be above 80%. Both visualizations are produced directly from DQD output.

Coverage per domain

An important part of the conversion quality is the concept mapping coverage. It is hard to get this overview from the DQD result tables alone. The new bar plot shows the **concept mapping coverage across all OMOP domains**. This ETL iteration achieved a high coverage throughout all domains and units (Figure 2) in terms of records mapped to standard concepts. The number of unique terms mapped was low for measurement and observation units (1.82% and 1.00%) and for measurement (19.42%).



▲ *Figure 2: Barplot for the mapping coverage in an ETL. In light blue: the percentage of distinct terms mapped to a standard OMOP concept; In darker blue: the percentage of records mapped to a standard OMOP concept.*

References
[1] Blacketer C, Defalco F, Ryan P, Rijnbeek P. Increasing trust in real-world evidence through evaluation of observational data quality. medRxiv 2021.
[2] EHDEN COVID19 Rapid Data Partner Call, April 2020, retrieved May 2021.https://www.ehden.eu/open-calls/04-2020-covid19-data-partner-call/

Elena G. Lara, Maxim Moinat, Anne van Winzum

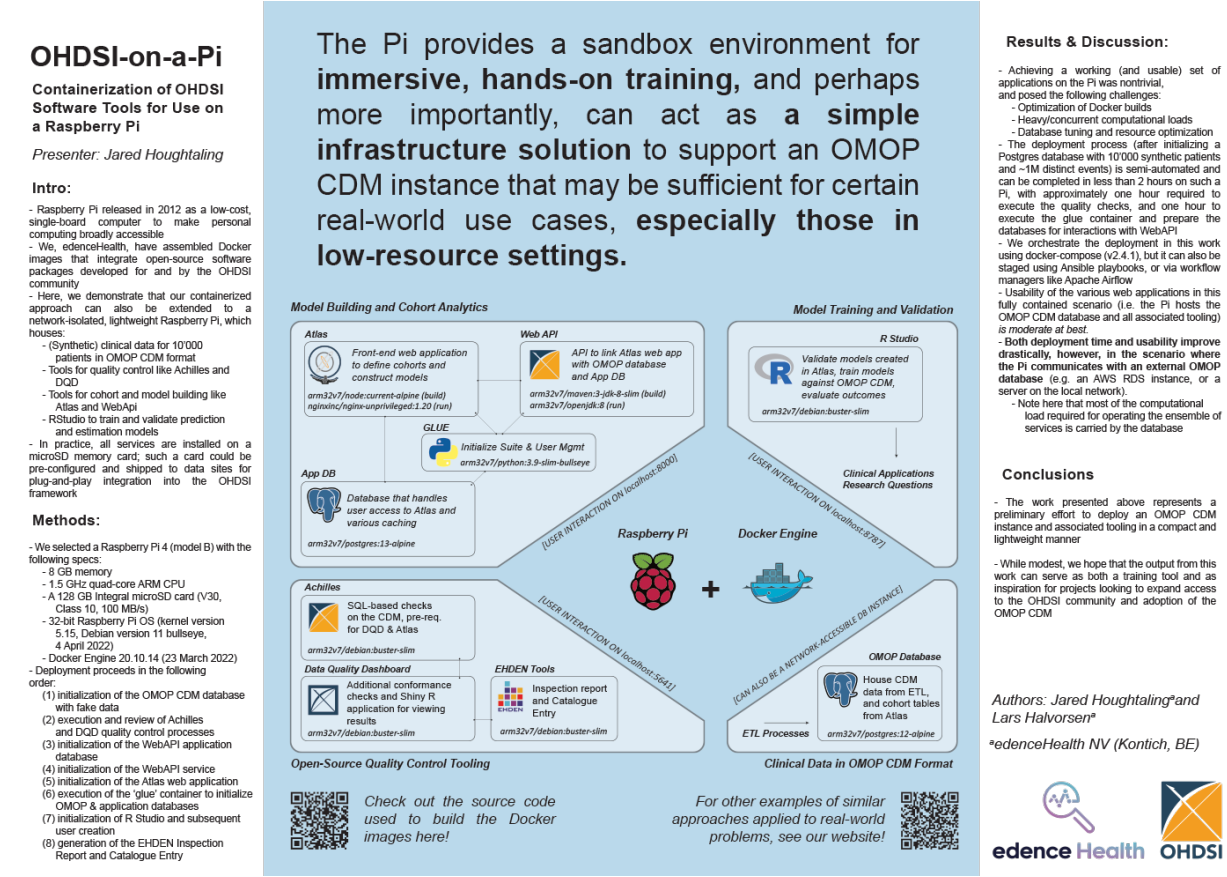
The Hyve OHDSI EHDEN

MONDAY

Comparing Data Quality Dashboard results from consecutive ETL iterations: two new visualizations and one utility script

Lead: Anne van Winzum

#OHDSISocialShowcase This Week



OHDSI-on-a-Pi

Containerization of OHDSI Software Tools for Use on a Raspberry Pi

Presenter: Jared Houghtaling

Intro:

- Raspberry Pi released in 2012 as a low-cost, single-board computer to make personal computing broadly accessible
- We, edenceHealth, have assembled Docker images that integrate open-source software packages developed for and by the OHDSI community
- Here, we demonstrate that our containerized approach can also be extended to a network-isolated, lightweight Raspberry Pi, which houses:
  - (Synthetic) clinical data for 10'000 patients in OMOP CDM format
  - Tools for quality control like Achilles and DQD
  - Tools for cohort and model building like Atlas and WebApi
  - RStudio to train and validate prediction and estimation models
- In practice, all services are installed on a microSD memory card; such a card could be pre-configured and shipped to data sites for plug-and-play integration into the OHDSI framework

Methods:

- We selected a Raspberry Pi 4 (model B) with the following specs:
  - 8 GB memory
  - 1.5 GHz quad-core ARM CPU
  - A 128 GB Integral microSD card (V30, Class 10, 100 MB/s)
  - 32-bit Raspberry Pi OS (kernel version 5.15, Debian version 11 bullseye, 4 April 2022)
  - Docker Engine 20.10.14 (23 March 2022)
- Deployment proceeds in the following order:
  (1) initialization of the OMOP CDM database with fake data
  (2) execution and review of Achilles and DQD quality control processes
  (3) initialization of the WebAPI application database
  (4) initialization of the WebAPI service
  (5) initialization of the Atlas web application
  (6) execution of the 'glue' container to initialize OMOP & application databases
  (7) initialization of R Studio and subsequent user creation
  (8) generation of the EHDEN Inspection Report and Catalogue Entry

The Pi provides a sandbox environment for **immersive, hands-on training,** and perhaps more importantly, can act as **a simple infrastructure solution** to support an OMOP CDM instance that may be sufficient for certain real-world use cases, **especially those in low-resource settings.**

Check out the source code used to build the Docker images here!

For other examples of similar approaches applied to real-world problems, see our website!

Results & Discussion:

- Achieving a working (and usable) set of applications on the Pi was nontrivial, and posed the following challenges:
  - Optimization of Docker builds
  - Heavy/concurrent computational loads
  - Database tuning and resource optimization
- The deployment process (after initializing a Postgres database with 10'000 synthetic patients and ~1M distinct events) is semi-automated and can be completed in less than 2 hours on such a Pi, with approximately one hour required to execute the quality checks, and one hour to execute the glue container and prepare the databases for interactions with WebAPI
- We orchestrate the deployment in this work using docker-compose (v2.4.1), but it can also be staged using Ansible playbooks, or via workflow managers like Apache Airflow
- Usability of the various web applications in this fully contained scenario (i.e. the Pi hosts the OMOP CDM database and all associated tooling) *is moderate at best.*
- **Both deployment time and usability improve drastically, however, in the scenario where the Pi communicates with an external OMOP database** (e.g. an AWS RDS instance, or a server on the local network).
  - Note here that most of the computational load required for operating the ensemble of services is carried by the database

Conclusions

- The work presented above represents a preliminary effort to deploy an OMOP CDM instance and associated tooling in a compact and lightweight manner

- While modest, we hope that the output from this work can serve as both a training tool and as inspiration for projects looking to expand access to the OHDSI community and adoption of the OMOP CDM

Authors: Jared Houghtaling[a] and Lars Halvorsen[a]

[a]edenceHealth NV (Kontich, BE)

edence Health OHDSI

TUESDAY

OHDSI-On-A-Pi: Containerization of OHDSI Software Tools for Use on a Raspberry Pi

Lead: Jared Houghtaling

#OHDSISocialShowcase This Week

WEDNESDAY

Integration prospects of the Ukrainian healthcare system with OMOP CDM

Lead: Mariia Kolesnyk

#OHDSISocialShowcase This Week

Title: *Patient Treatment Trajectory Modeling With Markov Chains*

PRESENTER: **Markus Haug**

INTRO:

- Treatment trajectories give us a foundation to find out the best healthcare practices, evaluate the economics of treatment patterns and model the treatment paths.
- Two R packages (**Cohort2Trajectory** & **TrajectoryMarkovAnalysis**) were developed.

METHODS

- **Cohort2Trajectory**

1. Importing relevant target and state cohorts.
2. Resolving cohort overlap conflicts.
3. Choosing the trajectory creation settings.
4. Output: CSV with trajectories.

- **TrajectoryMarkovAnalysis**

1. Importing treatment trajectories.
2. Using the them to produce discrete or continuous time Markov chain models.
3. Querying data from specific domains for state cost analysis.
4. Synthetic trajectories can be generated from the assembled Markov models.
5. Output: Markov model, state cost statistics, synthetic medical data.

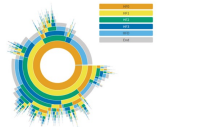
RESULTS

- To showcase the functionalities of the R packages we reproduced the study of heart failure carried out in the UK (Thokala et al., 2020) on data supplied by the Estonian Health Insurance Fund.
- The packages can be implemented in large-scale studies with regard to patient treatment trajectories.

Constructing **patient treatment trajectories** and producing **Markov chain** models for **cost analysis.**

CASE STUDY:

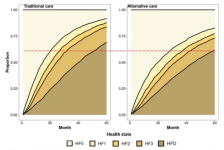
- Study by Thokala et al. for comparing traditional care with additional telemonitoring use among heart failure patients was reproduced using the packages.
- Defined five notating Markov states for heart failure progression and death.
- Markov and cost-effectiveness analysis were conducted.



Chronological transitions on heart failure study

	Results from Estonia					Results from the UK				
	HF0	HF1	HF2	HF3	HFD	HF0	HF1	HF2	HF3	HFD
HF0	0.95	0.04	0.00	0.00	0.01	0.98	0.01	0.00	0.00	0.01
HF1	0.06	0.78	0.12	0.02	0.02	0.07	0.87	0.02	0.00	0.04
HF2	0.00	0.07	0.72	0.18	0.03	0.01	0.09	0.78	0.04	0.08
HF3	0.00	0.00	0.04	0.92	0.04	0.00	0.00	0.10	0.78	0.12
HFD	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00

- When using telemonitoring the disease progression was diminished, but the cost of a QALY was high:
  - 60735.87 €/QALY in Estonia;
  - 56316.94 €/QALY in the UK.



State distributions of synthetic data comparing usual care and alternative care

AUTHORS:
Markus Haug, Marek Oja, Maarja Pajusalu, Raivo Kolde

THURSDAY

Patient treatment trajectory modeling with Markov chains

Lead: Markus Haug

#OHDSISocialShowcase This Week

Assessing treatment effect heterogeneity using the RiskStratifiedEstimation R-package

PRESENTER: Alexandros Rekkas

INTRODUCTION

In the presence of a truly effective treatment, effect heterogeneity should always be anticipated on some scale, as baseline risk is bound to vary across the study population.

METHODS

- We translated existing RCT guidelines to the observational setting.
- We created an open-source software package for the application of our framework.
- **Hypertension**: We compared ACE inhibitors to beta blockers with regard to 3 main outcomes (acute myocardial infarction-MI, hospitalization with heart failure, and stroke) and 6 safety outcomes (abnormal weight gain, angioedema, cough, hyperkalemia, hypokalemia, and hypotension) in IBM MarketScan Commercial Claims and Encounters (CCAE), IBM MarketScan Medicaid (MDCD), and IBM MarketScan Medicare Supplemental Beneficiaries (MDCR).
- **Osteoporosis**: We compared the effect of teriparatide to oral bisphosphonates in female patients over the age of 50, diagnosed with osteoporosis in CCAE, Optum De-Identified Electronic Health Record Dataset (Optum-EHR), and Optum De-Identified Clinformatics Data Mart Database-Date of Death (Optum-DOD).

RESULTS

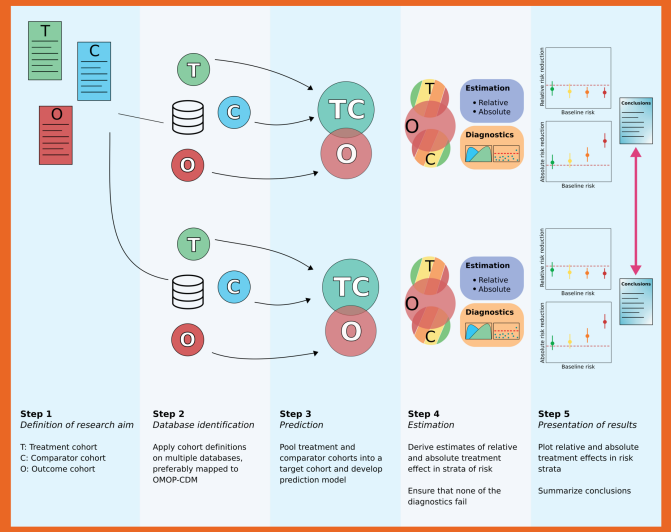
Hypertension

- With increasing acute MI risk we observed increasing absolute benefits in terms of acute MI and hospitalization with heart failure with ACE inhibitors.
- We also observed large cough risk increase with ACE inhibitors across all acute MI risk strata.

Osteoporosis

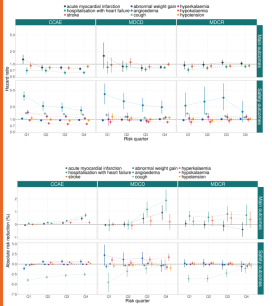
- Strong evidence of residual confounding limits our ability to draw conclusions.

Large-scale evaluation of treatment effect heterogeneity is feasible

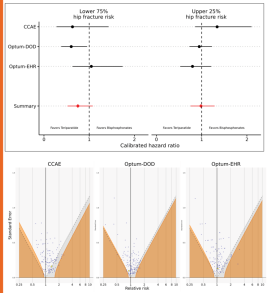


Scan QR code
Link to github

Hypertension



Osteoporosis



Alexandros Rekkas, David van Klaveren, Peter R. Rijnbeek

Erasmus MC
OHDSI

FRIDAY

Assessing treatment effect heterogeneity using the RiskStratifiedEstimation R-package

Lead: Alexandros Rekkas

Where Are We Going?

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

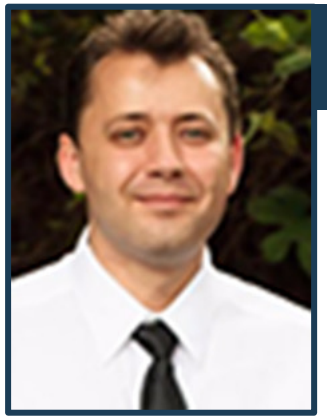
Three Stages of The Journey

Where Have We Been?

Where Are We Now?

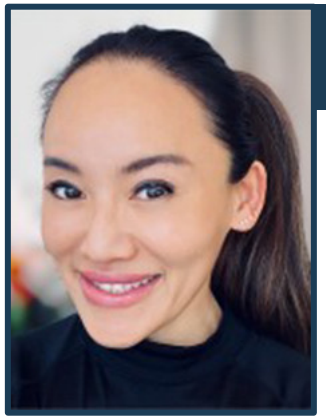
Where Are We Going?

Aug. 2 Community Call: Building Organizational Support Within Your Community



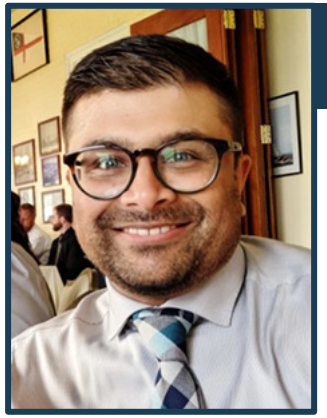
Greg Klebanov

CTO/SVP • Odysseus Data Services, Inc.



Keran Moll

Director, HEOR Real World Data & Analytics Research • Regeneron



Ajit Londhe

Senior Manager, Center for Observational Research • AMGEN

Paul Nagy

Program Director for Graduate Training in Biomedical Informatics and Data Science • Johns Hopkins University