# How did OHDSI do in 2024?

## OHDSI Community Call
## Dec. 10, 2024 • 11 am ET

# Upcoming Community Calls

| Date | Topic |
|---|---|
| Dec. 10 | How Did We Do In 2024? |
| Dec. 17 | Holiday-Themed Final Call of 2024 |
| Dec. 24 | No Call |
| Dec. 31 | No Call |
| Jan. 7 | What Can OHDSI Go In 2025? |

# Three Stages of The Journey

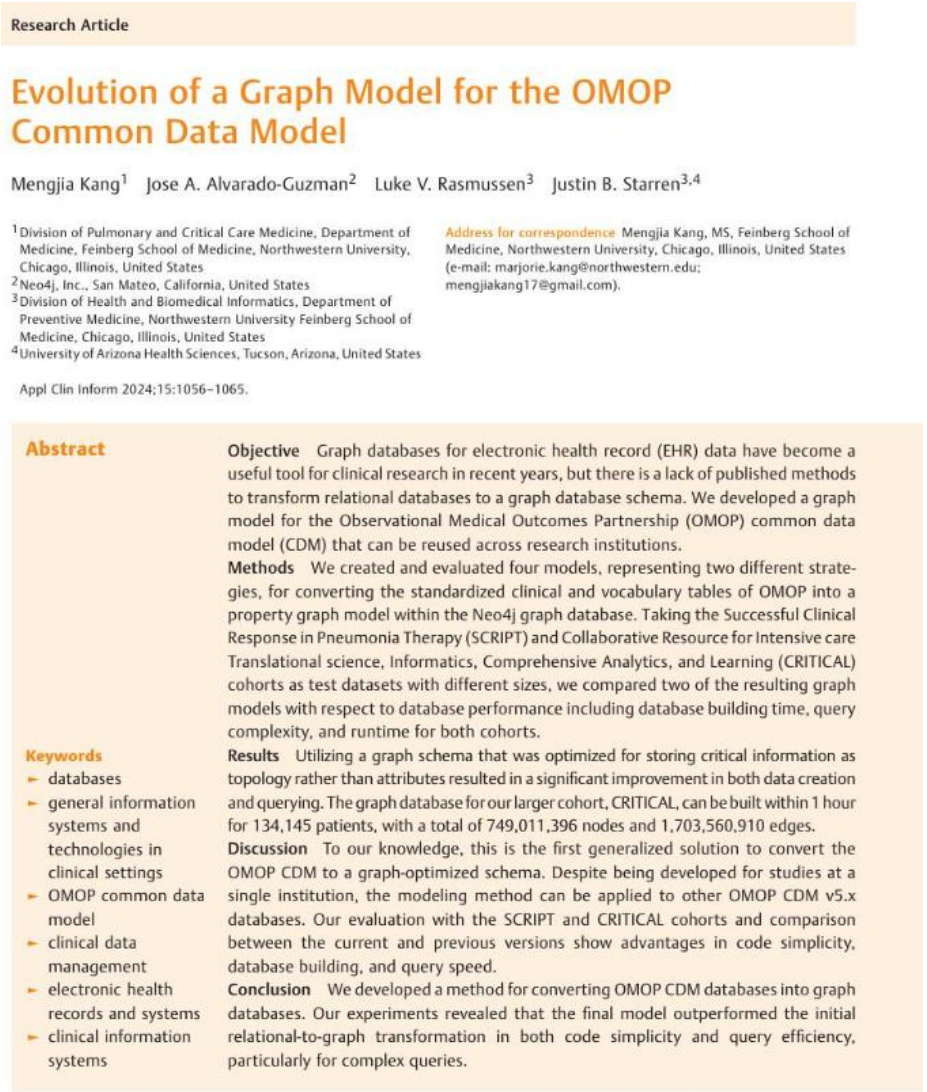
## Where Have We Been?

## Where Are We Now?

## Where Are We Going?

# OHDSI Shoutouts!

Congratulations to the team of **Mengjia Kang, Jose Alvarado-Guzman, Luke Rasmussen, and Justin Starren** on the publication of **Evolution of a Graph Model for the OMOP Common Data Model** in *Applied Clinical Informatics*.



Research Article

## Evolution of a Graph Model for the OMOP Common Data Model

Mengjia Kang[1] Jose A. Alvarado-Guzman[2] Luke V. Rasmussen[3] Justin B. Starren[3,4]

[1]Division of Pulmonary and Critical Care Medicine, Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, United States
[2]Neo4j, Inc., San Mateo, California, United States
[3]Division of Health and Biomedical Informatics, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States
[4]University of Arizona Health Sciences, Tucson, Arizona, United States

Address for correspondence Mengjia Kang, MS, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, United States (e-mail: marjorie.kang@northwestern.edu; mengjiakang17@gmail.com).

Appl Clin Inform 2024;15:1056–1065.

**Abstract**

**Objective** Graph databases for electronic health record (EHR) data have become a useful tool for clinical research in recent years, but there is a lack of published methods to transform relational databases to a graph database schema. We developed a graph model for the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) that can be reused across research institutions.

**Methods** We created and evaluated four models, representing two different strategies, for converting the standardized clinical and vocabulary tables of OMOP into a property graph model within the Neo4j graph database. Taking the Successful Clinical Response in Pneumonia Therapy (SCRIPT) and Collaborative Resource for Intensive care Translational science, Informatics, Comprehensive Analytics, and Learning (CRITICAL) cohorts as test datasets with different sizes, we compared two of the resulting graph models with respect to database performance including database building time, query complexity, and runtime for both cohorts.

**Results** Utilizing a graph schema that was optimized for storing critical information as topology rather than attributes resulted in a significant improvement in both data creation and querying. The graph database for our larger cohort, CRITICAL, can be built within 1 hour for 134,145 patients, with a total of 749,011,396 nodes and 1,703,560,910 edges.

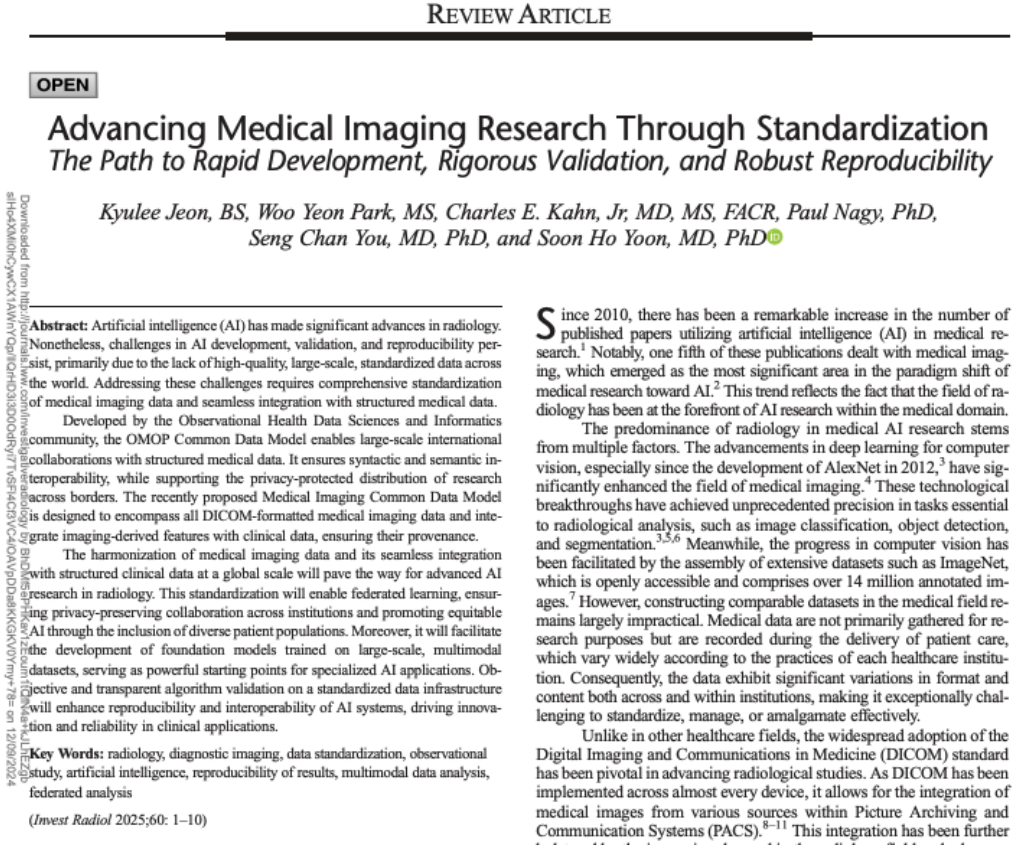
**Discussion** To our knowledge, this is the first generalized solution to convert the OMOP CDM to a graph-optimized schema. Despite being developed for studies at a single institution, the modeling method can be applied to other OMOP CDM v5.x databases. Our evaluation with the SCRIPT and CRITICAL cohorts and comparison between the current and previous versions show advantages in code simplicity, database building, and query speed.

**Conclusion** We developed a method for converting OMOP CDM databases into graph databases. Our experiments revealed that the final model outperformed the initial relational-to-graph transformation in both code simplicity and query efficiency, particularly for complex queries.

**Keywords**
- databases
- general information systems and technologies in clinical settings
- OMOP common data model
- clinical data management
- electronic health records and systems
- clinical information systems

# OHDSI Shoutouts!

Congratulations to the team of **Kyulee Jeon, Woo Yeon Park, Charles Kahn Jr, Paul Nagy, Seng Chan You, and Soon Ho Yoon** on the publication of **Advancing Medical Imaging Research Through Standardization: The Path to Rapid Development, Rigorous Validation, and Robust Reproducibility** in *Investigative Radiology.*



REVIEW ARTICLE

OPEN

## Advancing Medical Imaging Research Through Standardization

### *The Path to Rapid Development, Rigorous Validation, and Robust Reproducibility*

*Kyulee Jeon, BS, Woo Yeon Park, MS, Charles E. Kahn, Jr, MD, MS, FACR, Paul Nagy, PhD, Seng Chan You, MD, PhD, and Soon Ho Yoon, MD, PhD*



**Abstract:** Artificial intelligence (AI) has made significant advances in radiology. Nonetheless, challenges in AI development, validation, and reproducibility persist, primarily due to the lack of high-quality, large-scale, standardized data across the world. Addressing these challenges requires comprehensive standardization of medical imaging data and seamless integration with structured medical data.

Developed by the Observational Health Data Sciences and Informatics community, the OMOP Common Data Model enables large-scale international collaborations with structured medical data. It ensures syntactic and semantic interoperability, while supporting the privacy-protected distribution of research across borders. The recently proposed Medical Imaging Common Data Model is designed to encompass all DICOM-formatted medical imaging data and integrate imaging-derived features with clinical data, ensuring their provenance.

The harmonization of medical imaging data and its seamless integration with structured clinical data at a global scale will pave the way for advanced AI research in radiology. This standardization will enable federated learning, ensuring privacy-preserving collaboration across institutions and promoting equitable AI through the inclusion of diverse patient populations. Moreover, it will facilitate the development of foundation models trained on large-scale, multimodal datasets, serving as powerful starting points for specialized AI applications. Objective and transparent algorithm validation on a standardized data infrastructure will enhance reproducibility and interoperability of AI systems, driving innovation and reliability in clinical applications.

**Key Words:** radiology, diagnostic imaging, data standardization, observational study, artificial intelligence, reproducibility of results, multimodal data analysis, federated analysis

(*Invest Radiol* 2025;60: 1–10)

Since 2010, there has been a remarkable increase in the number of published papers utilizing artificial intelligence (AI) in medical research.[1] Notably, one fifth of these publications dealt with medical imaging, which emerged as the most significant area in the paradigm shift of medical research toward AI.[2] This trend reflects the fact that the field of radiology has been at the forefront of AI research within the medical domain.

The predominance of radiology in medical AI research stems from multiple factors. The advancements in deep learning for computer vision, especially since the development of AlexNet in 2012,[3] have significantly enhanced the field of medical imaging.[4] These technological breakthroughs have achieved unprecedented precision in tasks essential to radiological analysis, such as image classification, object detection, and segmentation.[3,5,6] Meanwhile, the progress in computer vision has been facilitated by the assembly of extensive datasets such as ImageNet, which is openly accessible and comprises over 14 million annotated images.[7] However, constructing comparable datasets in the medical field remains largely impractical. Medical data are not primarily gathered for research purposes but are recorded during the delivery of patient care, which vary widely according to the practices of each healthcare institution. Consequently, the data exhibit significant variations in format and content both across and within institutions, making it exceptionally challenging to standardize, manage, or amalgamate effectively.

Unlike in other healthcare fields, the widespread adoption of the Digital Imaging and Communications in Medicine (DICOM) standard has been pivotal in advancing radiological studies. As DICOM has been implemented across almost every device, it allows for the integration of medical images from various sources within Picture Archiving and Communication Systems (PACS).[8–11] This integration has been further

# Three Stages of The Journey

## Where Have We Been?

## Where Are We Now?

## Where Are We Going?

# Upcoming Workgroup Calls

| Date | Time (ET) | Meeting |
|---|---|---|
| Tuesday | 12 pm | Generative AI and Analytics |
| Tuesday | 12 pm | Common Data Model Vocabulary Subgroup |
| Tuesday | 3 pm | OMOP CDM Oncology Outreach/Research Subgroup |
| Wednesday | 12 pm | Health Equity |
| Wednesday | 7 pm | Medical Imaging |
| Thursday | 8 am | Medical Devices |
| Thursday | 9:30 am | Network Data Quality |
| Thursday | 10:30 am | Evidence Network |
| Thursday | 12 pm | Strategus HADES Subgroup |
| Thursday | 6 pm | Eyecare and Vision Research |
| Friday | 9 am | Phenotype Development and Evaluation |
| Friday | 10 am | GIS-Geographic Information System |
| Friday | 11:30 am | Steering Group |
| Friday | 11 pm | China Chapter |
| Monday | 9 am | Vaccine Vocabulary |
| Monday | 11 am | Data Bricks User Group |
| Monday | 2 pm | Electronic Animal Health Records |

# 2024 APAC Symposium

## Dec. 4-8 • Marina Bay Sands & National University of Singapore (NUS)

**Day 2-3: Main conference**

**Day 2 Slides**

Day 1 Opening (Mengling 'Mornin' Feng)
OHDSI for Real-World Evidence (RWE) (Patrick Ryan)
Charting our APAC Journey: Lessons from the Past, Visions for the Future (Mui Van Zandt)
OHDSI APAC Regional Chapter Updates (Korea – Rae Woong Park, China – Wang Changran, Australia – Nicole Pratt, Japan – Tatsuo Hiramatsu, Taiwan – Jason Hsu, Singapore – Mengling 'Mornin' Feng, India – Parthiban Sulur)
2024 APAC ETL Project (Mui Van Zandt, Gyeol Song, Steven Yong, Satish Kumar Anbazhagan, Kosuke Tanaka, Santan Maddi)
OHDSI Evidence Network (Erica Voss)
Large Language Model and OHDSI: Part 1 (Hua Xu)
Large Language Model and OHDSI: Part 2 (Hyeonsik Kim)
HL7 Singapore and OHDSI Singapore Collaboration (Adam Chee, Mengling 'Mornin' Feng)




**Day 3 Slides**

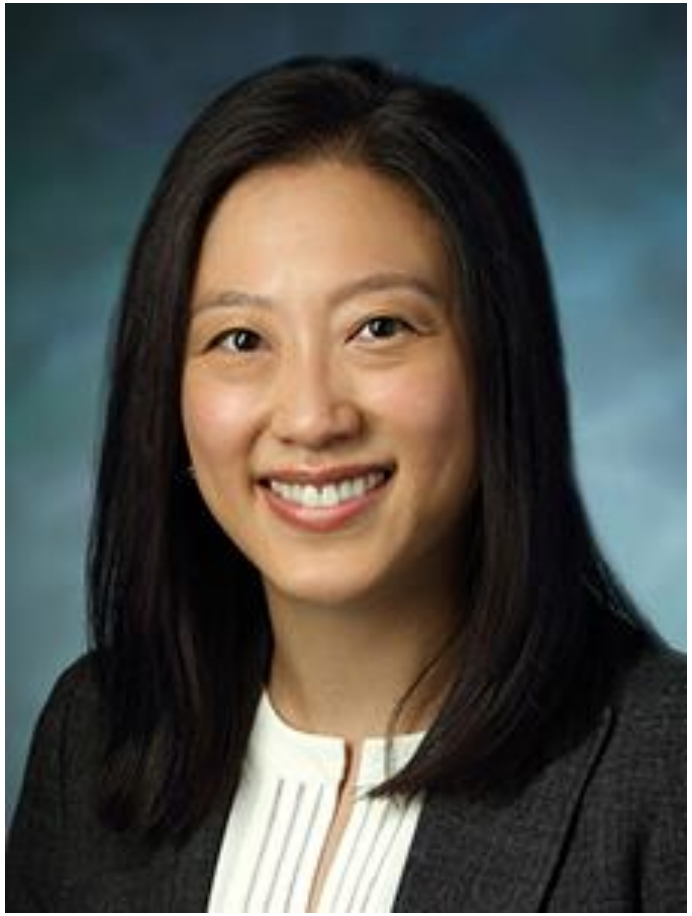
Overview of the International and Singapore Standards Ecosystem (Aik Lam Khor)
TRUST: Enabling Safe Data Exchange and Our OMOP Journey (Mingshi Koh)
OMOP Common Data Model: Journey Towards Singapore's National Data Standardization for Real-World Evidence Generation (Mukkesh Kumar)
Use of OHDSI to Evaluate Safety Signals (Mengling 'Mornin' Feng)
LEGEND-T2DM Study Introduction (Marc Suchard)
2024 APAC Study Introduction (Sreemanee Dorajoo)
2024 APAC Study: Journey from Data to Evidence (Evelyn Goh, Nicole Pratt)
Lightning Talks


**ohdsi.org/APAC2024**

# Collaborator Spotlight: Cindy Cai

**Dr. Cindy Cai** is the Jonathan and Marcia Javitt Rising Professor of Ophthalmology at Johns Hopkins University and a retina specialist seeing patients at the Wilmer Eye Institute's locations in the Baltimore, Maryland area. Her primary focuses are in medical and surgical retina treatments, including: diabetic retinopathy, diabetic macular edema, and age-related macular degeneration.

A co-lead of the Eyecare and Vision Research Workgroup, Cindy is currently leading another OHDSI network study focused on Semaglutide and Nonarteritic Anterior Ischemic Optic Neuropathy. The 2024 Titan Award for Clinical Applications honoree, she discusses her career journey, her experience running her first community network study, opportunities in vision research using real-world data, and plenty more in the latest collaborator spotlight.



**ohdsi.org/spotlight-Cindy-Cai**

# 2024 Global Symposium


## 2024 OHDSI Global Symposium

**Oct. 22-24 · New Brunswick, N.J. · Hyatt Regency Hotel**

The 10th annual OHDSI Global Symposium brought together more than 470 global collaborators for three days of sharing research, building new connections and pushing forward our mission of improving health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care.

This page will host all materials from OHDSI2024, including video presentations (when available) from the main conference and tutorials, slide decks, posters, demos and more.

### State of the Community

**Where Have We Gone and Where Are We Going?**
(George Hripcsak, Columbia University)

**Expand OHDSI Initiative for Eye Care and Ocular Imaging Challenge**
(Amberlynn Reed, Natiional Eye Institute)

**Titan Awards**
(George Hripcsak, Columbia University & Marc Suchard, UCLA)


State of the Community Slides

### Plenary: Value Proposition for Participating in OHDSI Network Studies like LEGEND-T2DM

**Introduction to OHDSI Evidence Network / Marketplace**
(Moderator: Clair Blacketer, Johnson & Johnson)

**Reflections from US Department of Veterans Affairs**
(Scott Duvall, VA)

**Reflections from SIDIAP (Spain)**
(Talita Duarte-Salles, IDIAP)

**Reflections from a Global Commercial Data Provider**
(Atif Adam, IQVIA)


Plenary: Value Proposition for Participating in OHDSI Network Studies like LEGEND-T2DM Slides

### Plenary Q&A: Lessons Learned on LEGEND-T2DM Journey

**Moderator:** Fan Bu, University of Michigan

**Panelists:** LEGEND-T2DM co-authors


Plenary Q&A: Lessons Learned on LEGEND-T2DM Journey Slides

### Plenary Panel: JACC-OHDSI Partnership

**Moderators:**
Nicole Pratt, University of South Australia
Marc Suchard, UCLA

**Panelists:**
Harlan Krumholz, Yale University
Seng Chan You, Yonsei University
Yuan Lu, Yale University


Plenary Panel: JACC-OHDSI Partnership Slides

**ohdsi.org/OHDSI2024**

### 2024 Global Collaborator Showcase

#### Observational Data Standards & Management

1 – Application of OMOP Common Data Model to Disease Registry Data (Vojtech Huser, Maria Rogozhkina, Vlad Korsik, Teresa A. Simon, Peter Moorthamer, Dan Kiselev, Teresa A. Simon, Anastasia Vakhmistrova, Eugene Paulenkovich, Alexander Davydov, Michel Van Speybroeck)
2 – Best Practices for Developing Disease-Specific Federated Networks: Insights from a Systemic Lupus Erythematosus Study (Clair Blacketer, Frank DeFalco, Gowtham A Rao, Anna Sheahan, Michel Van Speybroeck, Martine Lewi, Federico Zazzetti)
3 – Standardizing Rare Disease Patient Registry data to the OMOP-CDM (Parag Shiralkar, Radhika Lakireddy, Sushma Ghanta, Sanket Kalyankar)
4 – PHederation – the federated network of Pulmonary Hypertension registries (Eva-Maria Didden, Valerie van Baalen, Michel van Speybroeck, Monika Brand)
5 – Lessons from mapping cancer information from European hospitals to ICD-O-3 conditions in OMOP (Lars Halvorsen, Olivier Bouissou, Elisabeth Ross, Stelios Theophanous, Joëlle Thonnard, Piers Mahon)
6 – SMEs optimization with high precision data ingestion of CAPriCORN CDM onto OMOP at AllianceChicago (Andrew Hamilton, Amro Hassan, Davera Gabriel, Guy Tsafnat)
7 – Process of Conversion of Ukrainian Medical Data to OMOP CDM Format (Bohdan Khilchevskyi, Denys Kaduk, Maksym Trofymenko, Polina Talapova, Tetiana Nesmiian, Max Ved, Inna Ageeva, Pavlova Olga, Holovko Tetiana, Shevchenko Natalia)
8 – An evaluation of the transformation of large German EHR database to OMOP CDM (Andreas Ochs, Milou Brand, Jack Brewster, Methosdios Typou, Meda Sandu, Joe Maskell, Meghan Pettine, Atif Adam, George Kafatos)
9 – Adopting the OMOP Oncology CDM at the Helsinki University Hospital (Valtteri Nieminen, Alexey Ryzhenkov, Johanna Sanoja, Salma Rachidi, Juho Lähteenmaa, Joonas Laitinen, Samu Eränen, Tomi Mäkelä, Eric Fey, Kimmo Porkka)
10 – Going global, redeeming the local: an innovative approach to implement the OMOP CDM in two countries of the Global South (Valentina Martufi, Emma Kalk, Enny S. Cruz, Juliana Araújo Prata de Faria, Adalton do Anjos Fonseca, Maurício L. Barreto, Maria Yury Travassos Ichihara, Jessica Gammon, Nicki Tiffin, Chris Fourie, Danilo Luis Cerqueira Dias, Denise Moraes Pimenta, Tsaone Tamuhla, Andrew Boulle, Themba Mutemaringa, Juan-Paul Hynek, Muzzammil Ismail, Julio Barbour Oliveira, Ricardo Felix Monteiro Neto, Júlia Pescarini, Fernanda Revoredo de Sousa, Marianne Costa e Silva Lage, Adam Loff, Melvin Moodley, Elzo Pereira Pinto Junior)
11 – Transforming Clinical Trial Data to the OMOP CDM (Cynthia Sung, Mike Hamidi, Zhen Lin, Tom Walpole, Rebecca Baker, Melissa Cook, Shital Desai, Priya Gopal, Dan Hartley, Vojtech Huser, Priya Meghrajani, Tra Nguyen, Paul Orona12, Katy Sadowski, Sebastiaan van Sandijk, Philip Solovyev, Ramona Walls, Kenneth J. Wilkins, Qi Yang)
12 – Streamlining Research Data Standardization: AI-READI Survey Instrument Data Elements and MoCA Measurement Data Elements are curated and mapped utilizing a Standardized Value Set Mapping Table for transformation into the OMOP Common Data Model (Stephanie S. Hong, James Cavallon, Yvette Chen, Monique Bangudi, Jessica Mitchell, Dawn Matthies, Steven Chamberlin, Aaron Cohen, Julie Owens, Abigail Lucero, Sally Baxter, Christopher G Chute, Cecilia S. Lee, Aaron Lee, AI-READI consortium)
13 – Institutionalizing data interoperability and the application of common data models in a health data and research center: CIDACS' experience in Brazil (Valentina Martufi, Juliana Araújo Prata de Faria, Danilo Luis Cerqueira Dias, Elzo Pereira Pinto Junior, Roberto Carreiro, Pablo Ivan Ramos, Maurício L. Barreto)
14 – OMOP GIS Vocabulary Package for Observational Studies in Health Care and Public Health (Maksym Trofymenko, Polina Talapova, Andrew Williams)
15 – Enhancing Infectious Disease Data Integration and management through OMOP-CDM in South Korea (Min Ho An, Seok Kim, ByungJin Choi,Sooyoung Yoo,Rae Woong Park,Ji Seon Oh)
16 – FHIR to OMOP Cookbook – Mapping mCODE FHIR Resources for Observational Research (Qi Yang, Guy Livne, Sebastian van Sandijk, May Terry)
17 – Towards Reproducible Imaging Research: Implementation of DICOM to OMOP CDM (Woo Yeon Park, Ben Martin, Gabriel Salvador, Blake Dewey, Teri Sippel Schmidt, Paul Nagy)
18 – Leveraging UDI for Advanced Medical Device Tracking in OMOP-CDM (Seojeong Shin, Yiju Park, Sujeong Eom, Kyulee Jeon, Seng Chan You)
19 – Inclusion of intraocular pressure data into the University of California Health Data Warehouse (William Halfpenny*, Shahin Halla*, Ayan Patel, Catherine Q. Sun, Kerry Goetz, Michelle Hribar, Sally L. Baxter, on behalf of the OMOP Eye Care & Vision Research Workgroup)
20 – A Collaborative Analytic Enclave for the Metabolic Dysregulation and Obesity Cancer Risk Program (MeDOC) Consortium: Extensions of the OMOP Common Data Model for Translational Research (Madhan Subramanian, Nisha Grover, Maddie Wheeler, Marinella Temprosa)
21 – Expanding the OMOP Common Data Model to support Extracorporeal Life Support research (Clemens Rieder, Oleg Zhuk, Ahmed Said, Peta M.A. Alexander, Dominik J. Hoechter)
22 – ETIng from your OMOP CDM to your OMOP CDM? An efficient solution to vocabulary migration (Clair Blacketer, Anton Ivanov, Evanette Burrows, Dmitry Dymshyts, Frank DeFalco)
23 – Evaluating the impact of different vocabulary versions on cohort definitions and CDM (Dmitry Dymshyts, Frank DeFalco, Anna Ostropolets, Gowtham Rao, Azza Shoaibi, Clair Blacketer)

# CDM Survey Subgroup Landscape Assessment

**The CDM Survey Subgroup** invites colleagues who have or are going to design, develop, and/or implement research surveys and use them with the OMOP CDM to share information about those efforts by completing this survey. Your completion of this 10-15 minute survey will provide information to the CDM workgroup about OMOP utilization among survey research teams. The CDM Survey subgroup is a collaborative effort, led by a team at the National Cancer Institute, to develop standardized approaches and best practices for helping research teams better integrate survey data elements into the OMOP common data model.

**The survey deadline is Dec. 31, 2024.**

## LANDSCAPE ASSESSMENT

- **Activities**
  - Invite representatives from cohorts with experience using the CDM for survey data to share their knowledge and challenges.
  - Conduct a community survey to gather information on experiences and needs related to survey data in the CDM.
  - Review the most used Common Data Elements (CDMs) as a foundation for developing standards, tools, and best practices.
- **Key Result**
  - A comprehensive report summarizing survey CDM mapping resources, challenges, and identified development priorities (vocabulary, standards, tools, best practices) to be shared with the OHDSI community.

## WHO SHOULD PARTICIPATE

- You have survey data and you've mapped it to the OMOP CDM
- You have survey data and you would like to map it to the OMOP CDM
- You are in the process of developing a survey(s) and plan to map to the OMOP CDM
- Multiple perspectives from the same team
- Multiple surveys from the same person

# #OHDSISocialShowcase This Week

## Monday

## Best Practices for Developing Disease-Specific Federated Networks: Insights from a Systemic Lupus Erythematosus Study

(Clair Blacketer, Frank DeFalco, Gowtham A Rao, Anna Sheahan, Michel Van Speybroeck, Martine Lewi, Federico Zazzetti)

***Best Practices for Developing Disease-Specific Federated Networks***
*Insights from a Systemic Lupus Erythematosus Study*

PRESENTER: **Clair Blacketer**

**INTRODUCTION**

- Federated Networks (FN) offer a unique opportunity for collaborative data analysis, particularly for rare diseases like Systemic Lupus Erythematosus (SLE).
- This poster explores the development and best practices of a federated network using data from five global sources, all standardized to the OMOP Common Data Model (CDM).

**METHODS**

We identified challenges through the process of developing a disease-specific federated network focused on SLE.

***DATA STANDARDIZATION***

- Defining clinical characteristics of patients at registry enrollment was initially hindered by the differing methods used to represent the same questionnaires across the five data sources.

***DATA QUALITY and CHARACTERIZATION***

- Registry partners struggled to interact meaningfully with the data once it was standardized to the CDM which led to challenges in identifying and addressing data quality concerns.

***GENERATING EVIDENCE***

- Generating evidence using standardized registry data must address the inherent variability and complexity of the data sources.
- Differences in granularity between registries may result in varying levels of detail for similar data points.

***COLLABORATIVE ENGAGEMENT***

- Network partners must be able to understand the standard definitions and methodologies behind analyses.

By adhering to these **best practices**, disease-specific **federated networks** can achieve more **reliable**, **accurate**, and **meaningful outcomes.**

Take a picture to **download** the **short report**

**RECOMMENDATIONS FOR FUTURE STUDIES**

Based on the experiences and insights from the SLE federated network study, several recommendations can be made for future efforts:

**Enhanced Data Quality Investigations:** Comprehensive data quality presented in an accessible format.

**Leverage Standard Vocabularies:** Utilize standard vocabularies and mapping to align the data sources.

**Address Data Variability:** Manage differences in data granularity, temporality, and registry inclusion criteria.

**Foster Collaborative Engagement:** Engage registry partners in ongoing discussions to ensure a common understanding.

**Iterative Improvements:** Continuously refine data standardization efforts and methodologies based on feedback.

**Clair Blacketer[1,2], Frank DeFalco[1], Gowtham A Rao[1], Anna Sheahan[1], Michel Van Speybroeck[1], Martine Lewi[3], Federico Zazzetti[4]**

[1]Janssen Research & Development, Raritan, NJ, 2 Department of Medical Informatics, Eramus MC, Rotterdam, NL, 3 Janssen Global Services, LLC, Raritan, NJ, USA, 4 Janssen Global Services, LLC, a Johnson & Johnson Company, Immunology Medical Affairs, Horsham, PA, USA

Johnson&Johnson

Erasmus MC University Medical Center Rotterdam

OHDSI

# #OHDSISocialShowcase This Week

## Tuesday

# A Computable Phenotype for Time Toxicity of Elective Tracheostomy

(Abigail Martin, Ben Martin, Jen Park, Khyzer Aziz)

***Title: Developing a Computable Phenotype for Planned Tracheostomy***

PRESENTER: **Abby Martin**

**INTRO:**

- Tracheostomy is a surgical procedure that is conducted for both emergent and elective clinical indications.
- It will be important to separate emergent tracheostomy cases from elective ones for characterizing elective tracheostomy.
- Objective: develop a computable phenotype for patients receiving an elective tracheostomy to use in large-scale analysis of time toxicity.

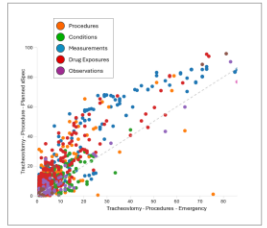
**METHODS:**

- Calculate standardized difference of covariate prevalence observed in the short-term window (30 days) prior to explicitly classified emergency vs. planned tracheostomy cohorts to evaluate classification criteria (***Figure 1***).
- A comparison of intubation start date in relation to tracheostomy start date between emergency vs. planned trach (***Table 3***).
- Use these criteria from the evaluation of explicitly coded trach procedures to classify remaining unclassified trach procedures (***Table 1***).
- Evaluate classification accuracy using Phevaluator and a set of cohorts with iterative addition of classification criteria (***Table 2***).

**Take a picture** to **download** the **full paper**

A **computable phenotype** for **planned tracheostomy** can use **intubation time** to eliminate **emergency tracheostomy** cases.

**Figure 1.** Covariate prevalence comparison




**RESULTS:**

- There were 3,552 patients with a tracheostomy procedure performed at Johns Hopkins Medicine from Jan. 2017- March 2024
- 2591 or 72.9% were explicitly coded as "emergency" or "planned"
  - 2484 planned (95.87%)
  - 107 emergency (4.13%)
- 961 (27.1%) were unclassified - not explicitly coded
- PheValuator results are outlined in ***Table 2***.

**Table 2.** PheValuator results for Planned Tracheostomy phenotype definitions tested on all patients in the JH OMOP database

| ID | Cohort Description | Sensitivity | PPV | Specificity | NPV | F1 Score | TP | TN | FP | FN |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Planned trach codes + unclassified | 0.980 | 0.628 | 1.000 | 1.000 | 0.766 | 1113 | 1486430 | 658 | 24 |
| 2 | Planned trach codes + (unclassified - Intubation exclusions) | 0.953 | 0.637 | 1.000 | 1.000 | 0.764 | 1103 | 1486470 | 626 | 54 |
| 3 | Cohort 2 criteria + Emergency-biased exclusions | 0.723 | 0.637 | 1.000 | 1.000 | 0.677 | 785 | 1486603 | 445 | 301 |
| 4 | Cohort 3 criteria + Planned-biased inclusions (at least 1) | 0.720 | 0.640 | 1.000 | 1.000 | 0.678 | 781 | 1486610 | 437 | 304 |
| 5 | Cohort 3 criteria + Planned-biased inclusions (at least 2) | 0.684 | 0.668 | 1.000 | 1.000 | 0.675 | 742 | 1486679 | 368 | 343 |
| 6 | Cohort 3 criteria + Planned-biased inclusions (at least 3) | 0.650 | 0.685 | 1.000 | 1.000 | 0.667 | 699 | 1486712 | 322 | 376 |

**CONCLUSIONS:**

- The phenotype using tracheostomy on the same day as intubation start as exclusion criteria provided the best ratio of PPV increase and sensitivity preservation.
- Additional criteria had much higher impact on sensitivity with marginal improvement in PPV.

**Table 1.** Tracheostomy patient counts: classified vs. unclassified

| Tracheostomy Concept Set | Patient Count | Percent |
|---|---|---|
| All tracheostomy procedure codes | 3552 | |
| Emergency + Planned tracheostomy codes | 2591 | |
| Unclassified tracheostomy procedures | 961 | 27.1% |
| | | |
| Emergency + Planned tracheostomy codes | 2591 | |
| Emergency | 107 | 4.13% |
| Planned | 2484 | 95.87% |

**Table 3.** Intubation start vs. Tracheostomy occurrence

| Intubation Days Before Trach | Emergency Trach | | Planned Trach | | Std. Diff Of Mean |
|---|---|---|---|---|---|
| | count | percent | count | percent | |
| Intubation same day | 27 | 71.05 | 77 | 6.59 | 1.764 |
| Intubation 1 day before | 0 | 0.00 | 51 | 4.36 | -0.302 |
| Intubation 2 days before | 0 | 0.00 | 86 | 7.36 | -0.399 |
| Intubation 3 days before | 1 | 2.63 | 94 | 8.04 | -0.242 |
| Intubation 4 days before | 1 | 2.63 | 113 | 9.67 | -0.296 |
| Intubation 5 days before | 4 | 10.53 | 121 | 10.35 | 0.006 |
| Intubation 6 days before | 3 | 7.89 | 98 | 8.38 | -0.018 |
| Intubation 7 days before | 1 | 2.63 | 97 | 8.30 | -0.251 |

Abby Martin, MPH MS[1]; Benjamin Martin, PhD[2]; Jen Wooyeon Park, MS[2]; Khyzer Aziz, MD[2]

1. Medical University of South Carolina, College of Medicine
2. Johns Hopkins University, School of Medicine

MUSC Health Medical University of South Carolina

Johns Hopkins School of Medicine

OHDSI

# #OHDSISocialShowcase This Week

## Wednesday

**Comparative Study of Informer, Prophet, and SARIMA Time Series Forecasting Models for Predicting Pneumonia-Related Hospitalizations and Emergency Room Visits in Elderly Patients Using OMOP-CDM**

(Seonghwan Shin, Junhyuk Chang, Min-Gyu Kim, Byungjin Choi, Rae Woong Park)

## Comparative Study of Informer, Prophet, and SARIMA Time Series Forecasting Models for Predicting Pneumonia-Related Hospitalizations and Emergency Room Visits in Elderly Patients Using OMOP-CDM


Seonghwan Shin, PharmD[1], Junhyuk Chang, PharmD[1], Min-Gyu Kim, MD[2], Byungjin Choi, MD[2], Rae Woong Park, MD, Ph.D.[1,2]

[1]Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, South Korea

[2]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, South Korea


### Background

- Pneumonia in elderly patients often presents fewer symptoms, making timely treatment difficult, which can lead to increased morbidity and mortality.
- As a result, sudden hospitalization and emergency room (ER) visits occur, placing a burden on healthcare resource management.
- Therefore, accurately predicting pneumonia-related hospitalizations is crucial for both patient care and efficient resource allocation.
- To address this need, this study aims to predict the daily number of pneumonia-related hospitalizations in the elderly using Prophet, SARIMA, and Informer time series forecasting models.

### Methods

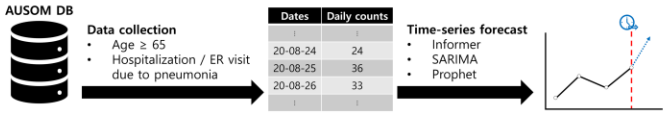



**Figure 1.** Framework and workflow of this study

**1. Data collection**
- **Database**
  - Ajou University School of Medicine (AUSOM) database (OMOP-CDM format)
- **Inclusion criteria for study population**
  - Patient records (2018-2023)
  - Age ≥ 65
  - Hospitalized or visited ER
  - Diagnosed as pneumonia within 24 hours of hospitalization or ER visit

**2. Preprocessing**
- Aggregated the daily counts of hospitalization and ER visits for the study population
- Missing dates are filled with 0
- Split: 80% for training / 20% for testing

**3. Model development**
- **Three models**
  - Prophet
  - SARIMA
  - Informer
- **Test period: 2 weeks (14 days)**
  - Compared to the actual observed counts during the test period

**4. Evaluation Metrics**
- Metrics used
  - Mean absolute error (MAE)
  - Root mean square error (RMSE)
- Lower metric values indicate better model performance
- Compared each model's accuracy using metrics above

### Conclusion

- Informer outperformed other models.
- We confirmed the potential of advanced time series forecasting models in predicting pneumonia-related hospitalizations and ER visits in elderly patients


Contact: contact@ohdsi.org


### Results

- A total of 31,338 patients, and 12,037 hospitalizations and ER visits were included.
- Informer demonstrated the lowest RMSE (1.089) and MAE (0.778), indicating superior performance.
- SARIMA followed with an RMSE of 2.595 and an MAE of 2.227.
- Prophet exhibited the highest error values, with an RMSE of 4.776 and an MAE of 4.489, reflecting the least favorable performance (Table 1, Figure 2).

**Table 1.** Performance metrics of the models

| Models | MAE | RMSE |
|---|---|---|
| **Informer** | **0.778** | **1.089** |
| **SARIMA** | 2.227 | 2.595 |
| **Prophet** | 4.489 | 4.776 |

*Note: Bold values indicate the best performance for each metric.

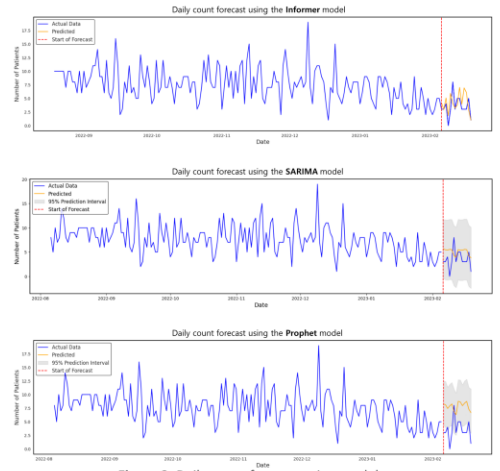



**Figure 2.** Daily count forecast using models


### Acknowledgements

This research was funded a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HR16C0001) and this research was supported by a Government-wide R&D Fund project for infectious disease research (GFID), Republic of Korea (grant number: HG22C0024, KH124685).

# #OHDSISocialShowcase This Week

## Thursday

## Comparative Effectiveness Research of Aflibercept and Bevacizumab in Patients with Diabetic Macular Edema: A Bayesian Causal Inference Study Using Real-world Data to Update Evidence from the Randomized Controlled Trial

(Kyungseon Choi, Sang Jun Park, Seng Chan You, Semin Jang, Haesun Suh)

***Comparative Effectiveness Research of Aflibercept and Bevacizumab in Patients with Diabetic Macular Edema: A Bayesian Causal Inference Study Using Real-world Data to Update Evidence from the Randomized Controlled Trial***

PRESENTER: Kyungseon Choi
Contact: kyungseon.choi@khu.ac.kr

**INTRO:**

- This study aims to evaluate aflibercept and bevacizumab effectiveness and efficacy in diabetic macular edema using frequentist and Bayesian statistics to inform clinical and regulatory decisions based on updated and synthesized evidence.

**METHODS (continued):**

- A frequentist-based post-hoc analysis of individual patient data from randomized controlled trial (RCT) was conducted to derive efficacy for updating evidence as prior belief and synthesizing evidence.
- For evidence updating, we employed a Bayesian causal inference with a retrospective cohort study design and utilized OMOP-CDM-transformed real-world data (RWD) from Bundang Seoul National University Hospital with 1.90 million patients, as it was the only available ophthalmology CDM fit for use.
- The study population was defined as patients aged over 18 who received intraocular aflibercept or bevacizumab treatment between June 1, 2015, and December 31, 2019.
- To fully adopt the RCT inclusion criteria, eligible participants have had a diagnosis of diabetes mellitus or diabetic retinopathy prior to treatment and a central subfield thickness of ≥300 μm.
- Propensity score was estimated using Bayesian additive regression tree and inverse probability weighting (IPTW) with standardized mortality ratio weighting was employed to correct for selection bias, considering covariates such as age, sex, disease status, measurements, and drug factors. The covariates over 0.20 standardized mean difference were considered as unmatched covariates and we checked negative control.

**Aflibercept showed superior efficacy compared to bevacizumab in patients with DME according to a post-hoc analysis of RCT.**

**Outcome:** Blindness-free survival (ln(Hazard ratio))

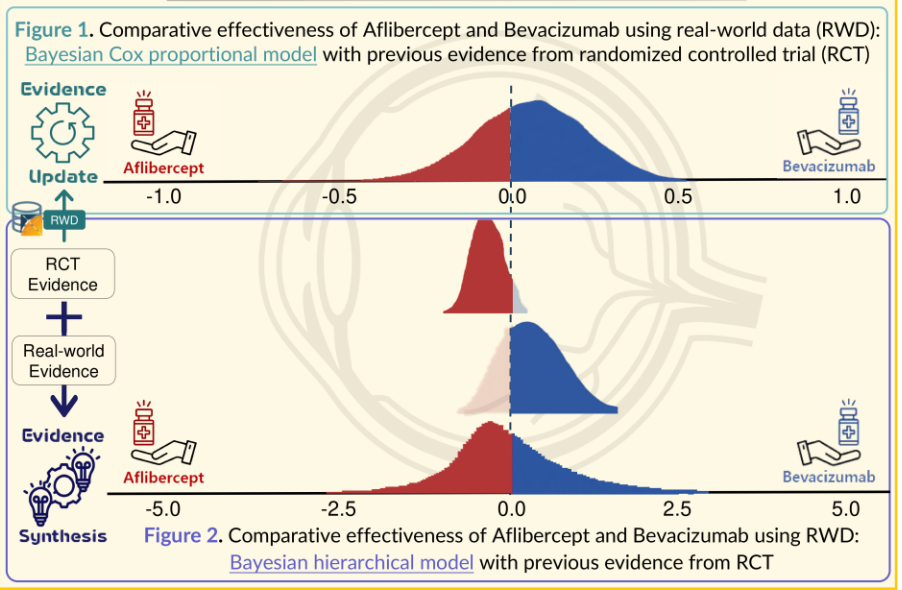



Figure 1. Comparative effectiveness of Aflibercept and Bevacizumab using real-world data (RWD): Bayesian Cox proportional model with previous evidence from randomized controlled trial (RCT)

Figure 2. Comparative effectiveness of Aflibercept and Bevacizumab using RWD: Bayesian hierarchical model with previous evidence from RCT

$$h(t) = h0(t)*exp(\beta1*drug+\beta2*age+\beta3*sex+\beta4*cardiovascular\ disease+\beta5*cerebrovascular\ disease+\beta6*renal\ disease+\beta7*hypertension+\beta8*proliferative\ diabetic\ retinopathy+\beta9*glaucoma+\beta10*cancer+\beta11*HbA1c+\beta12*Anti\text{-}VEGF+\beta13*best\ corrected\ visual\ acuity+\beta14*central\ subfield\ thickness)$$

**Although previous superior evidence from the RCT was included as prior probability in a Bayesian model, the superiority was not significant with the RWD.**

Take a picture to access the Brief report / connect on LinkedIn

This research was supported by KHIDI and a grant (21153MFDS601) from Ministry of Food and Drug Safety in 2023.

**METHODS (continued):**

- Blindness free survival (BFS), the outcome, was defined as the occurrence of death or blindness, with blindness specified as a best-corrected visual acuity of 20/200 or worse in the included eye according to legal criteria.
- For evidence synthesis, we employed a Bayesian hierarchical model to address heterogeneity between estimates from the frequentist-based Cox proportional models using RCT and RWD.

**RESULTS:**

- The study included 442 patients with DME from RCT IPD data set and 504 patients with DME from RWD after IPTW.
- Evidence update indicated a 34% probability that aflibercept would be superior in BFS (0.0728, SD 0.1771).
- In the evidence synthesis from efficacy and effectiveness, despite integrating RCT results, no significant differences were found between bevacizumab and aflibercept in BFS (-0.11, SD: 1.06, (59% probability of superiority for aflibercept))

**CONCLUSIONS:**

- Aflibercept showed superior efficacy in certain measures compared to bevacizumab. However, in the RWD, the superiority was not significant. Using a Bayesian model with a 95% threshold, the synthesized and updated evidence indicated no significant difference between aflibercept and bevacizumab in BFS.
- The findings support that bevacizumab may be as effective as aflibercept, suggesting policy implications for cost-effective drug reimbursement decisions.
- Bayesian approach, integrating new data, would enhance regulatory science decision-making, particularly for high-cost drugs, by providing a comprehensive view of efficacy and effectiveness. Furthermore, utilizing HERMES, a cost analysis tool for the OMOP-CDM, could expand to the RWD economic evaluation with OMOP-CDM.

Kyungseon Choi, Sang Jun Park, Seng Chan You, Semin Jang, Hae Sun Suh

SEOUL NATIONAL UNIVERSITY KYUNG HEE UNIVERSITY

# #OHDSISocialShowcase This Week

## Friday

## Determinants and persistence of medication adherence and its influence on health outcomes based on national health database

(Kerli Mooses, Marek Oja, Johannes Holm, Maarja Pajusalu, Hanna Keidong, Maria Malk, Sirli Tamm, Helene Loorents, Nikita Umov, Raivo Kolde)

***Determinants and persistence of medication adherence and its influence on health outcomes based on national health database***

PRESENTER: **Raivo Kolde**

**KNOWLEDGE GAP**

- Existing evidence base on medication adherence determinants is fragmented and conflicting, obtained by **small sample** studies, observing **single drug** and **small number of determinants** (Kardas, 2013).
- Need for **comprehensive analysis**, covering **all medications** on **population level.**

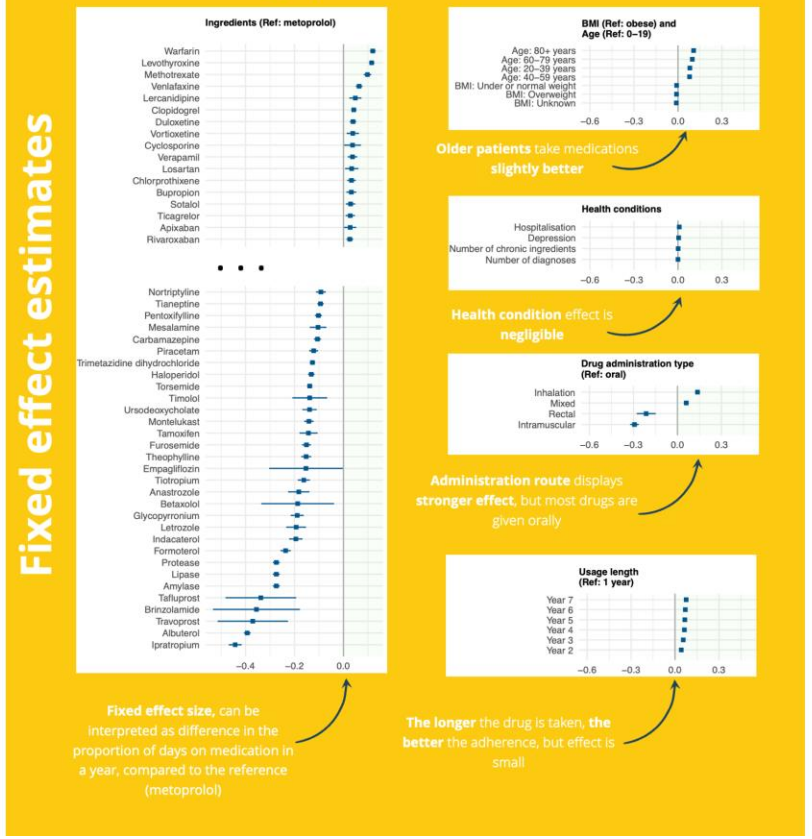
**METHODS**

- Medication adherence was calculated over **137 most prescribed** chronic use medications
- The dataset covered 150K individuals (**10% random sample of Estonian population**) out of whom **64K** (Table 1) had two consecutive prescriptions of at least one of the chronic medications.
- The medication adherence was estimated according to the **CMA5 measure**, taking into account gap times and refill banking. The calculations were done using AdhereR package (Dima, 2017) on an **OMOP** database.
- **CMA values** are between **0 and 1**, showing **how much of the year was covered with prescriptions**
- The CMA5 was calculated **yearly** for every prescribed **drug, per person**.
- The effect of the determinants was modeled using **linear mixed model**. (LMM)

$$CMA \sim AgeGroup + BMI + Diagnoses + Route + UsageLength + Drug + Drug*Diagnosis + (1|PersonId)$$

health-informatics.cs.ut.ee

Nationwide **prescription registry** analysis reveals **high correlation of medication adherence across drugs**, but surprisingly **little predictive power** of common **demographic** and **clinical** factors




**11.6%**
Of **variation** in LMM explained by **fixed effects**

**22.0%**
Of **variation** in LMM explained by **individual level effects** (random effect)

**Table 1. Cohort characteristics**

| Total | N = 64 837 |
|---|---|
| **Gender** | |
| Female | 37,111 (57.2%) |
| Male | 27,726 (42.8%) |
| **Body mass index category** | |
| Obese | 10,690 (16.5%) |
| Overweight | 9,396 (14.5%) |
| Under or normal weight | 8,003 (12.3%) |
| Unknown | 36,748 (56.7%) |
| **Age (years)** | 56.50 (21.75) |
| **CMA** | 0.75 (0.21) |

Dima AL, *et al* (2017) Plos One
Kardas P, *et al* (2013) Frontiers in Pharmacology

Kerli Mooses, Marek Oja, Johannes Holm, Maarja Pajusalu, Hanna Keidong, Maria Malk, Sirli Tamm, Helene Loorents, Nikita Umov, Sulev Reisberg, Jaak Vilo, Raivo Kolde
raivo.kolde@ut.ee
UNIVERSITY OF TARTU
1632

# Where Are We Going?

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

# Three Stages of The Journey

## Where Have We Been?

## Where Are We Now?

## Where Are We Going?


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

# Everybody is invited!

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