



# ATLAS Deepdive: Technical and Administrative Capabilities

OHDSI Community Call  
July 1, 2025 • 11 am ET





# Upcoming Community Calls

Date	Topic
July 1	ATLAS Deepdive: Technical and Administrative Capabilities
July 8	No Meeting – Europe Symposium
July 15	Europe Symposium Review
July 22	OMOP/OHDSI Research Spotlight
July 29	Asia-Pacific Regional Updates
Aug. 5	No Meeting
Aug. 12	Newcomer Introductions



# Three Stages of The Journey

**Where Have We Been?**

**Where Are We Now?**

**Where Are We Going?**





# OHDSI Shoutouts!



Congratulations to the team of **Alexander Saelmans, Tom Seinen, Victor Pera, Aniek F. Markus, Egill Fridgeirsson, Luis H. John, Lieke Schiphof-Godart, Peter Rijnbeek, Jenna Reps, and Ross Williams** on the publication of **Implementation and Updating of Clinical Prediction Models: A Systematic Review** in *Mayo Clinic Proceedings: Digital Health*.



MAYO CLINIC PROCEEDINGS:  
DIGITAL HEALTH

REVIEW

## Implementation and Updating of Clinical Prediction Models: A Systematic Review

Alexander Saelmans, MD; Tom Seinen, PhD; Victor Pera, PharmD; Aniek F. Markus, PhD; Egill Fridgeirsson, PhD; Luis H. John, MSc; Lieke Schiphof-Godart, PhD; Peter Rijnbeek, PhD; Jenna Reps, PhD; and Ross Williams, PhD

### Abstract

**Objective:** To summarize the implementation approaches and updating methods of clinically implemented models and consecutively advise researchers on the implementation and updating.

**Patients and Methods:** We included studies describing the implementation of prognostic binary prediction models in a clinical setting. We retrieved articles from Embase, Medline, and Web of Science from January 1, 2010, to January 1, 2024. We performed data extraction, based on Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis and Prediction Model Risk of Bias Assessment guidelines, and summarized.

**Results:** The search yielded 1872 articles. Following screening, 37 articles, describing 56 prediction models, were eligible for inclusion. The overall risk of bias was high in 86% of publications. In model development and internal validation, 32% of the models was assessed for calibration. External validation was performed for 27% of the models. Most models were implemented into the hospital information system (63%), followed by a web application (32%) and a patient decision aid tool (5%). Moreover, 13% of models have been updated following implementation.

**Conclusion:** Impact assessments generally showed successful model implementation and the ability to improve patient care, despite not fully adhering to prediction modeling best practice. Both impact assessment and updating could play a key role in identifying and lowering bias in models.

© 2025 THE AUTHORS. Published by Elsevier Inc on behalf of Mayo Foundation for Medical Education and Research. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>) ■ Mayo Clin Proc Digital Health 2025;3(3):100228



# OHDSI Shoutouts!



Congratulations to the team of  
**Richard Noll, Alexandra Berger,  
Carlo Facchinello, Katharina  
Stratmann, Jannik Schaaf, and  
Holger Storf** on the publication  
of Enhancing diagnostic  
precision for rare diseases using  
case-based reasoning in *JAMIA*.

*Journal of the American Medical Informatics Association*, 2025, 1–14  
<https://doi.org/10.1093/jamia/ocaf092>  
Research and Applications

AMIA  
INFORMATICS PROFESSIONALS LEADING THE WAY

OXFORD

## Research and Applications

### Enhancing diagnostic precision for rare diseases using case-based reasoning

Richard Noll , MSc<sup>\*1</sup>, Alexandra Berger, PhD<sup>2</sup>, Carlo Facchinello, PhD<sup>3</sup>, Katharina Stratmann, PhD<sup>4</sup>, Jannik Schaaf, PhD<sup>1</sup>, Holger Storf, PhD<sup>1</sup>

<sup>1</sup>Institute of Medical Informatics, Goethe University Frankfurt, University Medicine, 60590 Frankfurt am Main, Germany, <sup>2</sup>Frankfurt Reference Centre for Rare Diseases, University Medicine, Goethe University Frankfurt, 60596 Frankfurt am Main, Germany, <sup>3</sup>Academic Teaching Practice of the Heinrich-Heine-University Düsseldorf, 40885 Ratingen, Germany, <sup>4</sup>Medical Clinic 1, Goethe University Frankfurt, University Medicine, 60596 Frankfurt am Main, Germany

\*Corresponding author: Richard Noll, MSc, Institute of Medical Informatics, Goethe University Frankfurt, University Medicine, Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany (noll@med.uni-frankfurt.de)

#### Abstract

**Objective:** This study aims to enhance the diagnostic process for rare diseases using case-based reasoning (CBR). CBR compares new cases with historical data, utilizing both structured and unstructured clinical data.

**Materials and Methods:** The study uses a dataset of 4295 patient cases from the University Hospital Frankfurt. Data were standardized using the OMOP Common Data Model. Three methods—TF, TF-IDF, and TF-IDF with semantic vector embeddings—were employed to represent patient records. Similarity search effectiveness was evaluated using cross-validation to assess diagnostic precision. High-weighted concepts were rated by medical experts for relevance. Additionally, the impact of different levels of ICD-10 code granularity on prediction outcomes was analyzed.

**Results:** The TF-IDF method showed a high degree of precision, with an average positive predictive value of 91% in the 10 most similar cases. The differences between the methods were not statistically significant. The expert evaluation rated the medical relevance of high-weighted concepts as moderate. The granularity of ICD-10 coding significantly influences the precision of predictions, with more granular codes showing decreased precision.

**Discussion:** The methods effectively handle data from multiple medical specialties, suggesting broad applicability. The use of broader ICD-10 codes with high precision in prediction could improve initial diagnostic guidance. The use of Explainable AI could enhance diagnostic transparency, leading to better patient outcomes. Limitations include standardization issues and the need for more comprehensive lab value integration.

**Conclusion:** While CBR shows promise for rare disease diagnostics, its utility depends on the specific needs of the decision support system and its intended clinical application.

**Key words:** case-based reasoning; rare diseases; diagnostic techniques and procedures; medical informatics; interdisciplinary research.



# 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	ATLAS
Wednesday	8 am	Psychiatry
Thursday	11 am	Themis
Thursday	11 am	Industry
Thursday	12 pm	Methods Research
Thursday	1 pm	Oncology Vocabulary/Development Subgroup
Monday	9 am	Vaccine Vocabulary
Monday	10 am	Africa Chapter
Monday	10 am	Getting Started Subgroup
Tuesday	9 am	Oncology Genomic Subgroup
Tuesday	9:30 am	CDM Survey Subgroup





# #OHDSI2025 Submission Deadline: **TODAY**

The Collaborator Showcase submission deadline for the 2025 Global Symposium is **July 1 (8 pm ET)**.

The #OHDSI2025 Global Symposium will be held Oct. 7-9 in New Brunswick, NJ.

More information about the collaborator showcase, including links to the submission form and poster templates, can be found on the #OHDSI2025 homepage.







# #OHDSI2025 Tutorials

The 2025 Global Symposium will open with a day of tutorials (Oct. 7), providing opportunities for both OHDSI newcomers and veterans to learn more about the community and focused research areas.

An introductory tutorial will be a standalone session during the morning; while the afternoon will include five advanced tutorials. Learn more about each below; you can sign up for specific tutorials during the symposium registration process.





# #OHDSI2025 Tutorials

**Morning Session (8 am - 12 pm ET)**

## **An Introduction to the Journey from Data to Evidence Using OHDSI**

The journey from data to evidence can be challenging alone but is greatly enabled through community collaboration. In this half-day tutorial, we will introduce newcomers to OHDSI. Specifically, about the tools, practices, and open-science approach to evidence generation that the OHDSI community has developed and evolved over the past decade.



**Lead: Erica Voss**



# #OHDSI2025 Tutorials

**Afternoon Session (1 pm - 5 pm ET)**

## **Developing and Evaluating Your Extract, Transform, Load (ETL) Process to the OMOP Common Data Model**

In this tutorial, students will learn about the tools and practices developed by the OHDSI community to support the journey to establish and maintain an ETL to standardize your data to OMOP CDM and enable standardized evidence generation across a data network.



**Lead: Clair Blacketer**

## **Using the OHDSI Standardized Vocabularies for Research**

In this tutorial, students will learn how to take advantage of the OHDSI standardized vocabularies as an analytic tool to support your research, including searching for relevant clinical concepts, navigating concept relationships, creating conceptsets and understanding source codes that map within these expressions. Students will also learn where the OHDSI standardized vocabularies are used throughout OHDSI's standardized analytic tools.



**Lead: Anna Ostropolets**





# #OHDSI2025 Tutorials

## Clinical Characterization Applications to Generate Reliable Real-World Evidence

Clinical characterization—descriptive statistics to summarize disease natural history, treatment utilization, and outcome incidence—are the at heart of many real-world data applications, including study feasibility and quality improvement. In this tutorial, students will learn how to design and implement observational network studies for characterization, and how to apply tools and practices developed by the OHDSI community to ensure the evidence generated is reliable.



Lead: Patrick Ryan

## Population-Level Effect Estimation Applications to Generate Reliable Real-World Evidence

Population-level effect estimation—causal inference methods for comparative effectiveness and safety surveillance—enables researchers to understand how exposure to medical interventions are expected to impact health outcomes. In this tutorial, students will learn how to design causal inference studies and how to apply tools (such as CohortMethod) and practices (such as objective diagnostics) developed by the OHDSI community to ensure the evidence generated is reliable.



Lead: George Hripcsak

## Patient-Level Prediction Applications to Generate Reliable Real-World Evidence

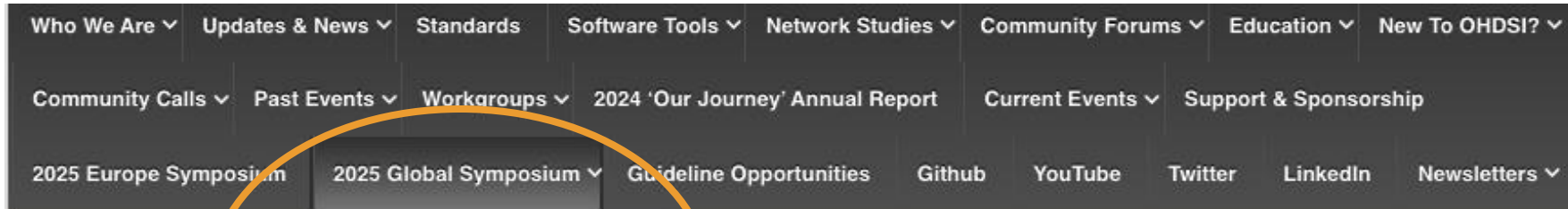
Patient-level prediction—the use of machine learning to train, test, and apply predictive models for disease interception and precision medicine—offers the potential to personalize healthcare by enabling individualized risk prediction based on personal health history. In this tutorial, students will learn how apply tools and practices developed by the OHDSI community, including the PatientLevelPrediction HADES R package, to design and implement network studies capable of learning and externally validating prediction models, and how to apply these models to your population.



Lead: Jenna Reps



# Global Symposium: Oct. 7-9

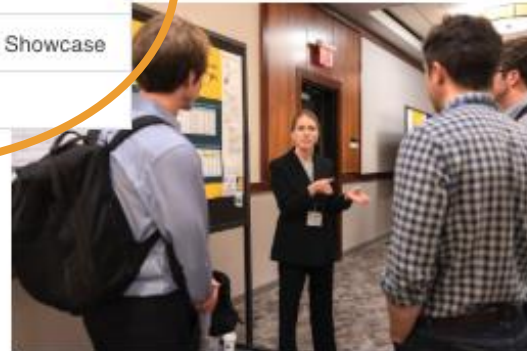


2025 Global Symposium Homepage

Register for OHDSI2025

OHDSI2025 Collaborator Showcase

OHDSI2025 Tutorials



## 2025 OHDSI Global Symposium

Oct. 7-9 • New Brunswick, N.J. • Hyatt Regency Hotel

There is nothing quite like the OHDSI Global Symposium, which welcomes hundreds of collaborators around the world who believe in the shared mission of improving health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care. We can't wait to return for our biggest event of the year this October in New Brunswick, N.J.

# Is Semaglutide Associated with Yet Another Blinding Eye Disease?

JAMA Ophthalmology | **Original Investigation**

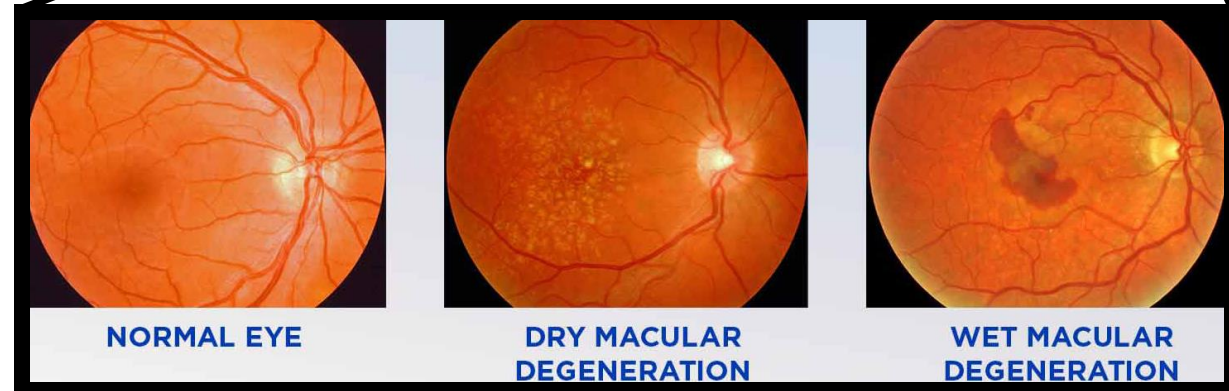
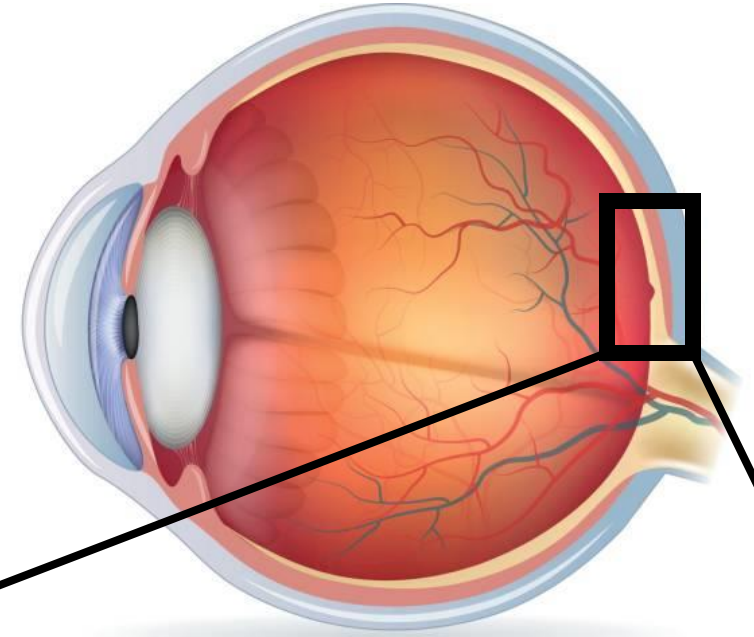
## Glucagon-Like Peptide-1 Receptor Agonists and Risk of Neovascular Age-Related Macular Degeneration

Reut Shor, MD; Andrew Mihalache, MD(C); Atefeh Noori, PhD; Renana Shor, MD; Radha P. Kohly, MD, PhD; Marko M. Popovic, MD, MPH; Rajeev H. Muni, MD, MSc

**Hazard Ratio of NVAMD 2.21 (95% CI 1.65 – 2.96)**

Linked claims + EHR data (Ontario Health Insurance Plan)

46,334 adults with diabetes exposed to GLP1-RA (>6mo) compared to 92,668 unexposed to GLP1-RA







# Europe Symposium Agenda

## Symposium Agenda – July 7, 2025

Time	Topic	
8:00 – 9:00	Registration & Coffee	
9:00 – 9:10	Welcome to the European OHDSI Journey ( <i>Speakers: Liesbet M. Peeters &amp; Peter Rijnbeek</i> )	
9:10 – 9:30	Journey of OHDSI: Where have we been and where can we go together? ( <i>Speaker: Patrick Ryan</i> )	
9:30 – 11:00	Impact of Leveraging OMOP CDM for Scalable and Reliable Evidence Generation Showcased by the National Nodes ( <i>Moderators: Renske Los &amp; Annelies Verbiest</i> )	
11:00 – 11:30	Coffee Break	
11:30 – 12:45	Collaborator Showcase: Rapid Fire Presentations ( <i>Moderator: TBC</i> )	
12:45 – 13:45	Lunch	
13:45 – 16:00	OHDSI Collaborator Showcase	Early Investigator Mentor Meeting (14:00 – 15:00)
16:00 – 17:10	Bridging Policy and Practice: OHDSI's Role in Implementing the European Health Data Space ( <i>Panel debate</i> ) ( <i>Confirmed speakers/moderators: Enrique Bernal-Delgado, Nick Marly, Talita Duarte-Salles, Patrick Ryan, Dipak Kalra</i> )	
17:10 – 17:30	Closing remarks ( <i>Speakers: Liesbet M. Peeters &amp; Peter Rijnbeek</i> )	

## Agenda Saturday July 5, 2025

Time	Activity	Track 1A – Newcomers	Track 1B – Newcomers	Track 2 – Advanced	Track 3 – NN/WG
09:30 – 10:00		Registration + coffee			
10:00 – 12:30	Morning Session	<b>Introduction to OHDSI – Tutorial</b> Lead: Renske Los, Aniek Markus & Laura Verbeij (Erasmus MC) Overview of OHDSI, key concepts, and an introduction to the OMOP Common Data Model			<b>HADES hack-a-thon</b> Lead: Martijn Schuermie (J&J), Adam Black (Erasmus MC), Anthony Sena (Janssen R&D) Hands-on coding and tool development in HADES
12:30 – 13:30		Lunch break			
13:30 – 15:00	Afternoon Session I	<b>OMOP CDM &amp; ETL Conventions</b> Lead: Maxim Mainat (Erasmus MC), Sofia Bazakou & Anne van Winzum (The Hyve)	<b>OHDSI Standardized Vocabularies for Research – Part 1.1</b> Lead: Anna Ostropelets (Janssen R&D), Polina Talapova (Sciforce), Vlad Korsik & Oleg Zhuk (Odysseus) Concept sets & patient identification techniques.		
15:00 – 15:30		Coffee Break			
15:30 – 17:00	Afternoon Session II		<b>OHDSI Standardized Vocabularies for Research – Part 1.2</b> Lead: Anna Ostropelets (Janssen R&D), Polina Talapova (Sciforce), Vlad Korsik & Oleg Zhuk (Odysseus) Concept sets & patient identification techniques.		
17:15 – 18:45*		*Optional – guided city tour Hasselt (with local specialties)			

## Agenda Sunday July 6, 2025

Time	Activity	Track 1A – Newcomers	Track 1B – Newcomers	Track 2 – Advanced	Track 3 – NN/WG
09:30 – 10:00		Registration + coffee			
10:00 – 12:30	Morning Session		<b>OHDSI Standardized Vocabularies for Research – Part 2</b> (Janssen R&D), Polina Talapova (Sciforce), Vlad Korsik & Oleg Zhuk (Odysseus) <i>Final discussion &amp; application of concept sets.</i>	<b>NN All Actors Meet Parallel NN meetings</b>	
12:30 – 13:30		<b>Data Partners Lunch Break</b>			
13:30 – 15:00	Afternoon Session I	<b>Whirlwind Introduction to Open-Source Analytic Tools – Part 1</b> Lead: Martijn Schuermie (J&J), Adam Black (Erasmus MC), Anthony Sena (Janssen R&D) <i>Overview of HADES and other key OHDSI tools for analysis.</i>		<b>Running characterisation studies from beginning to end: a tutorial using DARWIN EU standardised analytics – Part 1</b> Lead: Daniel Prieto-Alhambra (Oxford University)	<b>NN All Actors Meet Parallel NN meetings</b>
15:00 – 15:30		<b>Coffee Break</b>			
15:30 – 17:00	Afternoon Session II	<b>Whirlwind Introduction to Open-Source Analytic Tools – Part 2</b> Lead: Martijn Schuermie (J&J), Adam Black (Erasmus MC), Anthony Sena (Janssen R&D) <i>Overview of HADES and other key OHDSI tools for analysis.</i>		<b>Running characterisation studies from beginning to end: a tutorial using DARWIN EU standardised analytics – Part 2</b> Lead: Daniel Prieto-Alhambra (Oxford University)	<b>OHDSI Europe NN leads meet</b> Lead: Renske Los (only NN leads/managers)
17:00 – 18:00*		<b>*Optional – networking drink</b>			



@OHDSI

www.ohdsi.org

#JoinTheJourney





# 2025 UK Symposium

The 2025 OHDSI UK Symposium will be held Sept. 26 in London.

Registration is now open!







# 2025 Global Symposium

The 2025 OHDSI Global Symposium will be held Oct. 7-9 in New Brunswick, N.J.

Registration for both the conference and tutorials is **OPEN!**





# 2025 Africa Symposium

The 2025 OHDSI Africa Symposium will be held Nov. 10-12 in Kampala, Uganda.

The abstract submission deadline will be August 25.

**OHDSI AFRICA Symposium**

Hosted by:  
Joint Clinical Research Centre  
**KAMPALA, UGANDA**

**Abstract Submission Deadline**  
August 25, 2025

*Save The Date*  
**NOV 10th - 12th 2025**

**About the Symposium:**  
Discover how interoperable healthcare data and federated networks can unlock real-world insights while safeguarding patient privacy. This innovative approach - bringing analyses to the data - empowers large-scale, global collaboration without compromising data ownership. A key highlight will be on advancing data-driven healthcare solutions, with a special focus on HIV.

Organised by JCRC  
in collaboration with: OHDSI

[www.jcrc.org.ug](http://www.jcrc.org.ug) **JCRC**

**OHDSI Uganda**





# 2025 APAC Symposium

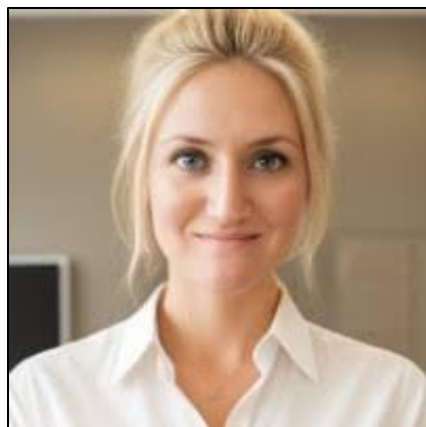
The 2025 OHDSI APAC Symposium will be held Dec. 6-7 in Shanghai, China.

Information on registration and abstract deadline will be posted when available.





# The Center for Advanced Healthcare Research Informatics (CAHRI) at Tufts Medicine welcomes:



**Tiffany Callahan, PhD**

*Senior Machine Learning Research Scientist at SandboxAQ*

**'Agentic Mixture-of-Workflows for Multi-Modal Chemical Search'**

July 31, 2025, 11am-12pm EDT

Virtually via [Zoom](#)

Please contact Marty Alvarez at [malvarez2@tuftsmedicalcenter.org](mailto:malvarez2@tuftsmedicalcenter.org) for calendar invite or questions.

**Tufts**Medicine  
Tufts Medical Center

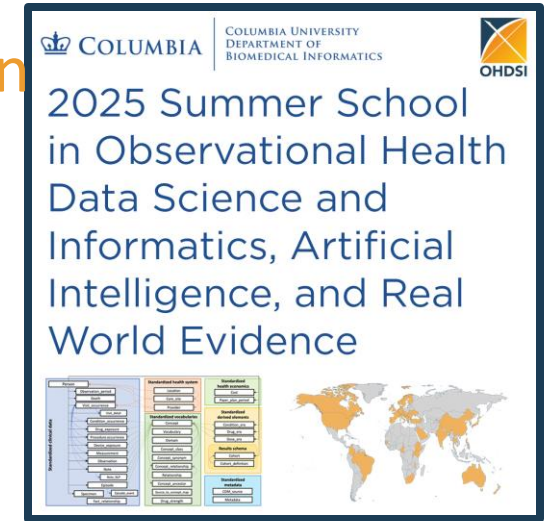




# Columbia Summer School on OHDSI

Registration is open for the first ever Columbia Summer School on OHDSI, held July 14-18, 2025, at the Columbia University Department of Biomedical Informatics in New York City.

The Columbia Summer School in Observational Health Data Science and Informatics, Artificial Intelligence, and Real World Evidence (RWE) offers health professionals, researchers and industry practitioners the opportunity to gain familiarity and hands-on experience with real world data and generating real world evidence. Participants will learn about the different types of healthcare data captured during routine clinical care, including electronic health records and administrative records, and how these data can be standardized to the OMOP Common Data Model to enable distributed data network research.



## Meet Our Faculty



**George Hripcsak, MD MS**  
Vivian Beaumont Allen  
Professor of Biomedical  
Informatics



**Patrick Ryan, PhD**  
Adjunct Assistant  
Professor of Biomedical  
Informatics



**Anna Ostropolets, MD PhD**  
Adjunct Assistant  
Professor of Biomedical  
Informatics



**Karthik Natarajan, PhD**  
Assistant Professor of  
Biomedical Informatics



# #OHDSISocialShowcase This Week

## Monday

# Research status of applying common data model in pharmaco-epidemiology: a systematic review

(Meng Zhang, Yongqi Zheng, Conghui Wang, Ling Gao, Feng Sun)



## Research status of applying common data model in pharmacoepidemiology: a systematic review

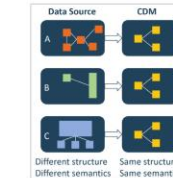
Meng Zhang<sup>1,2,\*</sup>, Yongqi Zheng<sup>1,2,\*</sup>, Conghui Wang<sup>1,3</sup>, Ling Gao<sup>1</sup>, Feng Sun<sup>1,2</sup>  
<sup>1</sup> Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China  
<sup>2</sup> Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing, China  
<sup>3</sup> Center for Pharmacovigilance of Inner Mongolia Autonomous Region, Inner Mongolia, China  
\* Contributed equally.



### Background

#### Multi-database analyses:

- larger sample size
- more generalizable evidence
- questions related to rare exposures and outcomes



- Information content varies greatly between databases, complicating analyses and interpretation of results across databases

- Common data models (CDMs) were developed to standardize data structures, format, and meaning

Table 1. 11 types of CDMs used in the world

CDMs	Institutions
VSD	CDC
CRN-VDR	CRN
i2b2	NIH
Sentinel	FDA
OMOP	OHDSI
ASPen	ACPE
FHIR	HL7
PCORnet	PCORI
PEDSnet	PCORI
CDASH	CDISC
ConcePTION	IMI

- ◆ The application of the CDMs in pharmacoepidemiologic research remain unclear

- To systematically summarize global work on the use of CDMs in pharmacoepidemiologic research

### Methods

#### Database

- Five English databases (PubMed, Web of Science, EMBASE, Scopus, Virtual Health Library) and four Chinese databases (CNKI, Wan-Fang Data, VIP, SinoMed)

#### Search strategy

- Combination of Chinese and English search terms: "common data model", "Observational Health Data Sciences and Informatics", "Observational Medical Outcomes Partnership", etc.

From database inception to Jan 2024

Language: English and Chinese

#### Inclusion criteria

- 1) The CDM (OMOP, CDASH, PCORnet, SDTM, i2b2, etc.) was used to answer questions in the field of drugs/vaccines/medical devices
- 2) Research areas included safety, effectiveness, utilization and accessibility, economic evaluation of all types of drugs, vaccines, and medical devices
- 3) The drug/vaccine/device must be the primary exposure or outcome

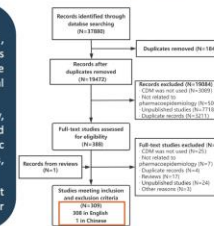


Figure 1. Flow diagram for study selection.

Contact: contact@ohdsi.org

### Results

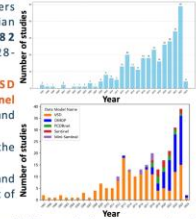
- 309 studies were from 36 countries with 1,522 authors, and published in 164 journals, covering 12 types of CDMs

- The median number of centers was 7 (IQR 4-8), with a median sample size of 2 67, 1 82 (interquartile range 16, 228, 1,531, 144)

- The top 5 CDMs used: VSD (52.8%), OMOP (24.3%), Sentinel (6.1%), Mini-sentinel (4.5%), and PCORnet (4.5%)

- 79.9% of the studies utilized the data sources from the US
- Korea (18.4%), China (2.6%), and Japan (2.6%) contributed most of Asian data

- Only 16 studies used reporting guidelines, and all of these employed the STROBE guidelines



#### Study type

- The focus was on vaccines in 56.0% of the studies, drugs in 43.4%, and devices/surgeries in 0.6%
- The most commonly studied vaccines were Influenza vaccines (54/173) and COVID-19 vaccines (21/173)
- Primary drugs included antidiabetic drugs (21/134) and antibiotics (12/134)

#### Research directions

- Safety (77.3%), drug/vaccine utilization (17.2%), effectiveness (8.0%), and others
- Safety events were primarily concentrated on the nervous system diseases (58/237) and autoimmune diseases (49/237), maternal and infant outcomes (43/237)

### Conclusions

- Our study provides a comprehensive perspective on all CDMs, related drugs/vaccines and treatment areas, **identifying the feasibility of CDMs in pharmacoepidemiologic studies**

- The future direction of CDM applications still needs further expansion, with a focus on enhancing the standardization of research reports

- We call on researchers from Asian countries to focus on and increase their participation in CDM applications and multicentre research to generate more representative studies

- We have also compiled the common bias of CDMs in pharmacoepidemiologic studies, and the article is still in the process of being written

- Bibliometric analysis and more detailed analysis are still under study

#### Acknowledgment

The authors would like to thank the sponsorship from the Joint Real-World Evidence Research Lab founded by Peking University Health Science Center-AstraZeneca established since Oct 2022; this research is part of the internship program starting July 2024. (<https://news.bion.com/article/ceb829e26a8.html>)

VSD, Vaccine Safety Datalink; CDC, Center for Disease Control; CRN-VDR, Cancer Research Network-Virtual Data Warehouse; CRN, Cancer Research Network; i2b2, Informatics for Integrating Biology & the Bedside; NIH, National Institute of Health; FDA, Food and Drug Administration; OMOP, Observational Medical Outcomes Partnership; OHDSI, Observational Health Data Sciences and Informatics; ASPEN, Asian Pharmacoepidemiology Network; ACPE, Asian Conference on Pharmacoepidemiology; FHIR, Fast Healthcare Interoperability Resources; HL7, Health Level Seven; PCORnet, Patient-Centered Clinical Research Network; PCORI, Patient-Centered Outcomes Research Institute; CDASH, Clinical Data Acquisition Standards Harmonization; CDISC, Clinical Data Interchange Standards Consortium; IMI, Innovative Medicines Initiative.



# #OHDSISocialShowcase This Week

## Tuesday

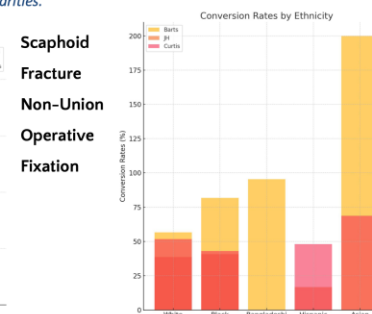
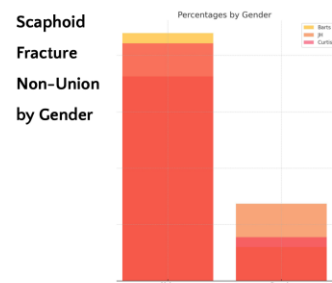
# Disparities in care for scaphoid fractures

(**Usama Rahman**, G Zhang, B Martin, P Nagy, A Giladi, D Laporte, D.S Edwards, X.L Griffin, J.C.E Lane)

*Those from minority ethnic backgrounds and those with male gender may be more likely to suffer a scaphoid fracture non-union and subsequent operative fixation.*

**Title:** An *international multi-centre federated-network data analysis using routine health data (EHR)* investigating *disparities in care for patients with scaphoid fractures*.

**Background:** Scaphoid fractures of the hand represent an injury that can be associated with significant morbidity if diagnosis or treatment is delayed. Patients may then present with a non-union of a fracture leading to disability and increased healthcare costs associated with treatment. Scaphoid fracture diagnosis and treatment are highly variable with age, gender, and ethnicity leading to disparities in care. We aimed to use routine health data (EHR) to investigate these disparities.



## Methods

- 1 Routine health data was extracted and converted to OMOP
- 2 Snomed CT and CPT-4 codes were used to create four cohorts:
  - scaphoid fracture
  - scaphoid fracture - primary operative intervention
  - scaphoid fracture non-union
  - scaphoid fracture non-union surgery
- 3 Federated analysis was undertaken across three international sites (secondary care) using the same OMOP code:
  - Barts NHS - UK
  - Johns Hopkins - USA
  - Curtis National Hand Centre - USA
- 4 Results were analysed descriptively and incidence rates were calculated based on the individual populations.

## Limitation:

- Variable coding between centres can limit the yield of information when conducting federated analyses
- Ethnicity coding varies between sites
- Scaphoid fractures may not be completely captured due to unmapped and unstructured data from outpatient departments



U Rahman 1,2, G Zhang 3, B Martin 3, P Nagy 4, A Giladi 3, D Laporte 4, D.S Edwards 1,2, X.L Griffin 1,2, J.C.E Lane 1,2

1 - Barts NHS Trust, London, United Kingdom  
2 - Barts Bone and Joint Health, Queen Mary University London, London, United Kingdom  
3 - Curtis National Hand Centre, MedStar Health, Baltimore, USA  
4 - Johns Hopkins Hospital, Baltimore, USA



@OHDSI

www.ohdsi.org

#JoinTheJourney



ohdsi



# #OHDSISocialShowcase This Week

Wednesday

## Standardisation of Multi-Agency Data to the OMOP Common Data Model

(Rachel Sippy, Anna Moore)

### Standardisation of Multi-Agency Data to the OMOP Common Data Model

Rachel Sippy, Anna Moore



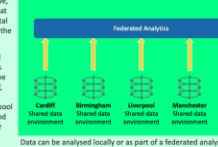
With funding and support from:  
NIHR, EPSRC, Wellcome, MRC, NIHR, EPSRC, Wellcome, MRC

#### CADRE: Children & Adolescent Data Resource

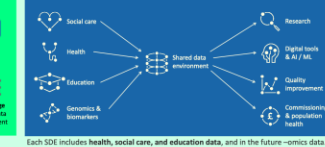
- An informatics platform to host live, data-led systems that enables child mental health research in the UK

- Network of shared data environments (SDEs) including five initial sites (Cambridge, Birmingham, Manchester, Liverpool and Cambridge) and extendable to other regions

##### Structure of the CADRE Network



##### Inputs & Outputs for Each Shared Data Environment



- Challenge**  
The OMOP-CDM is intended for health data  
Objective: fit non-health data into the OMOP-CDM  
What is the structure of social care data?  
What variables are in social care data?  
Can social care variables be mapped to OMOP?

#### Social Care Data Structure



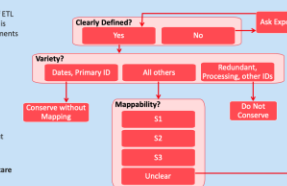
#### Efficient Processing of Social Care Variables

Choice of common data elements is part of ETL design, and understanding the source data is necessary to choose the common data elements

Variables were processed by these criteria:

- Clear variable definition available
- Variety
  - Identifiers
  - Processing
  - Dates
- Mappability
  - S1: directly mappable to OMOP concept
  - S2: closely mappable OMOP concept
  - S3: not represented in OMOP

Experts include data managers and social care staff involved in data collection



Social Care Variables			
Conserve	31%	32%	32%
Unclear	11%	11%	11%
No Mapping	7%	7%	7%
Do Not Conserve	42%	42%	42%

Of 135 social care variables, 58% are conserved and 42% are not conserved. Of the conserved variables, 22% can be mapped (S1 or S2). Mapping is unclear for 11% of variables – these are either ill-defined or have tentative mappings that need to be confirmed with social care data experts. The OMOP Vocabulary will need to be expanded to accommodate 7% of variables (S3).

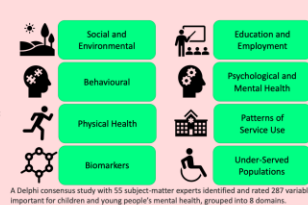
#### Prioritising Variables

For variables that are unclear or cannot be mapped using the current OMOP Vocabulary, additional effort is needed:

- Unclear variables need their definitions and mapping to be confirmed with social care experts
- The OMOP Vocabulary needs to be expanded for those variables that cannot be mapped

There are 22 unclear variables and 14 un-mappable variables.

Given that this additional processing is resource-intensive, we prioritise variables based on their importance for children's mental health research.



- Unclear Variables**  
Unclear variables may lack a definition, or the definition may not be understood by the data mapping team. There may be preliminary mapping that needs additional input. Important variables that are unclear include:
  - Those related to separation from family (e.g. fostering)
  - Those capturing entry into the social care system
- In the social care data, this information is given as outcomes from contacts/meetings. Clarity is needed for the level of specificity for these outcomes (e.g. agency versus non-agency adoption)**
- Unmappable Variables**  
Important variables that need to be added to the OMOP Vocabulary include:
  - Lower Super Output Area
  - Reasons for referral and
  - Provision of short break care

Webster A, Rimmer C, Schibye A, Gell W. Feasibility of Mapping Autism Health Claims Data to the OMOP Common Data Model. J Med Syst. 2019 Oct;43(10):354.





# #OHDSISocialShowcase This Week

## Thursday

# Lessons Learned from Mapping UK Pain Datasets to the OMOP CDM

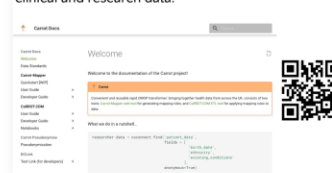
(**Gordon Milligan**, Erum Masood, Phil Appleby, Phil Quinlan, Sam Cox, Armando Mendez Villalon, Tom Giles, Calum MacDonald, Christian Cole)

Carrot Tools can make standardising data to OMOP more consistent to improve data reusability, interoperability and reduce time to map data. There is a need for a pain-specific standard vocabulary.

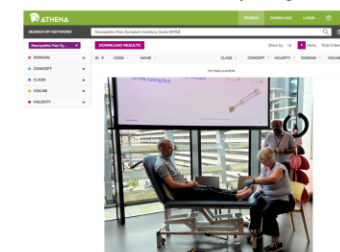
*Lessons Learned from Mapping UK Pain Datasets to the OMOP CDM*

**Background:** Chronic Pain affects up to 28M people in the UK and is poorly represented in electronic healthcare data. The aim of the Alleviate Pain Data Hub is to improve visibility and accessibility of pain data and ultimately facilitate access to research data across the UK. We have developed open-source tools which have supported the mapping of 5M records from across the UK and have found opportunities for improvement with pain data.

**Result 1:** We use Carrot tools with our mapping expertise to transform datasets to OMOP-CDM. The tools have improved the efficiency of mapping of clinical and research data.

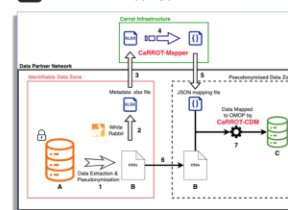


**Result 2:** We identified a lack of standard vocabulary representations for pain specific data (terms, scales such as Neuropathic Pain Symptom Inventory (NPSI) and tests such as Quantitative Sensory Testing (QST))



## Methods

1 Carrot Tools mapping process.



2 Mapping data UI and rules



**Note:** The tools are open-source, under continued development and are having more features added as the project progresses to address the needs of those performing the data mapping.



Gordon Milligan, Erum Masood, Phil Appleby, Phil Quinlan, Sam Cox, Armando Mendez Villalon, Tom Giles, Calum MacDonald, Christian Cole



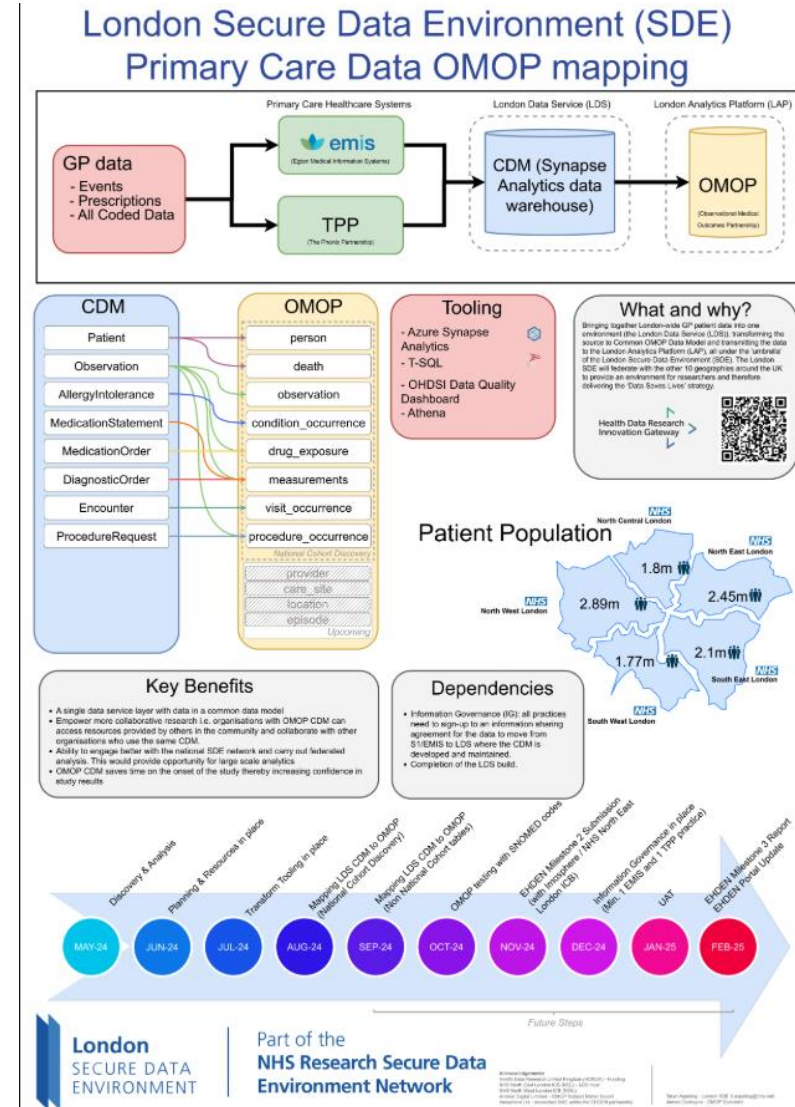


# #OHDSISocialShowcase This Week

Friday

## London Secure Data Environment (SDE) Primary Care Data OMOP Mapping

(Taryn Aspeling, James Cockayne)







# 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?**





# ATLAS Roadmap Homepage



Who We Are ▾	Updates & News ▾	Standards	Software Tools ▾	Network Studies ▾	Community Forums ▾	Education ▾	New To OHDSI? ▾
Community Calls ▾	Past Events ▾	Workgroups ▾	2024 'Our Journey' Annual Report	Current Events ▾	Support & Sponsorship		
2025 Europe Symposium	2025 Global Symposium ▾	Guideline Opportunities	Github	YouTube	Twitter	LinkedIn	Newsletters ▾

## ATLAS Deepdive: Learn About the Current Tool, Help Develop the Roadmap for Future Versions

ATLAS is an open-source, web-based tool that enables researchers to conduct scientific analyses on standardized observational health data. Our June community calls focused on both educating users about ATLAS and shaping its future roadmap. Watch the videos below to learn more—and be sure to complete the surveys to help guide the next phase of ATLAS development.

ATLAS workgroup lead **Christopher Knoll** guided the community through the month, while tool collaborators **Peter Hoffmann**, **Alexey Manoylenko**, **Richard Boyce** and **Konstantin Iaroshovets** provided demos on various aspects of ATLAS, including data sources and vocabularies, concept sets and cohorts, and characterization, incidence and treatment pathways.

One of the most widely used research tools in the community, the ATLAS team is now considering future versions and is seeking global input. What tools are the most important to you? How often do you use them? This is your opportunity to have your voice heard as we develop the roadmap for future versions of ATLAS!

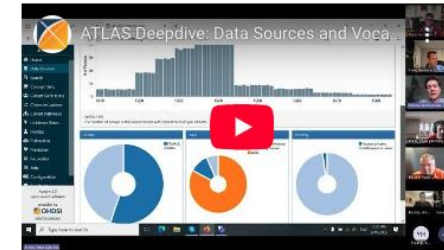


### The Journey of ATLAS: Introduction to the Tool



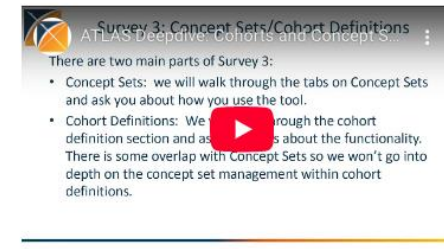
Survey 1: Overview

### ATLAS Deepdive: Data Sources/Vocabularies



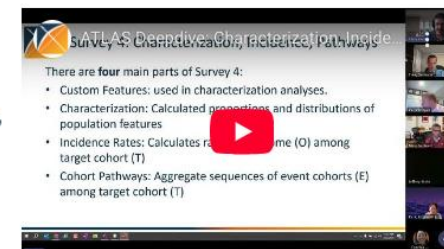
Survey 2: Data Sources/Vocabulary

### ATLAS Deepdive: Concept Sets/Cohorts



Survey 3: Concept Sets/Cohorts

### ATLAS Deepdive: Characterization, Incidence, Treatment Pathways



Survey 4: Characterization, Incidence, Treatment Pathways

[ohdsi.org/atlas-roadmap-2025](https://ohdsi.org/atlas-roadmap-2025)



# July 1: ATLAS Deepdive

## Technical and Administrative Functions



### Christopher Knoll

Director, Observational Health Data Analytics  
Janssen Research and Development  
ATLAS Workgroup Co-Lead



### Konstantin Iaroshovets

Product Ops Manager  
Odysseus Data Services, Inc.  
ATLAS Workgroup Co-Lead

Take our  
surveys to help  
create the  
roadmap for  
ATLAS!





**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-2025](https://ohdsi.org/community-calls-2025)**