Book of OHDSI 2.0: How Do We Get There?

OHDSI Community Call
March 11, 2025 • 11 am ET

Upcoming Community Calls

Date	Topic
Mar. 4	Vocabulary Release Update, Winter 2025
Mar. 11	Book of OHDSI 2.0 Brainstorm and Planning Session
Mar. 18	OHDSI Evidence Network and Data Diagnostics Design
Mar. 25	Methods for Evaluating Data Fitness for Use
Apr. 1	Recent OHDSI Publications

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **ChulHyoung Park, So Hee Lee, Da Yun Lee, Seoyoon Choi, Seng Chan You, Ja Young Jeon, Sang Jun Park, and Rae Woong Park** on the publication of **Analysis of Retinal Thickness in Patients With Chronic Diseases Using Standardized Optical Coherence Tomography Data: Database Study Based on the Radiology Common Data Model** in *JMIR Medical Informatics.*

JMIR Publications
Advancing Digital Health & Open Science

Articles | Search articles

JMIR Medical Informatics | Journal Information | Browse Journal | Subm

Published on 21.02.2025 in Vol 13 (2025)

Preprints (earlier versions) of this paper are available at https://preprints.jmir.org/preprint/64422, first published July 18, 2024.

Analysis of Retinal Thickness in Patients With Chronic Diseases Using Standardized Optical Coherence Tomography Data: Database Study Based on the Radiology Common Data Model

ChulHyoung Park[1]; So Hee Lee[2]; Da Yun Lee[3]; Seoyoon Choi[4]; Seng Chan You[5]; Ja Young Jeon[6]; Sang Jun Park[4]; Rae Woong Park[1,2]

Article | Authors | Cited by | Tweetations | Metrics

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

Abstract

Background:
The Observational Medical Outcome Partners-Common Data Model (OMOP-CDM) is an international standard for harmonizing electronic medical record (EMR) data. However, since it does not standardize unstructured data, such as medical imaging, using this data in multi-institutional collaborative research becomes challenging. To overcome this limitation, extensions such as the Radiology Common Data Model (R-CDM) have emerged to include and standardize these data types.

Objective:
This work aims to demonstrate that by standardizing optical coherence tomography (OCT) data into an R-CDM format, multi-institutional collaborative studies analyzing changes in retinal thickness in patients with long-standing chronic diseases can be performed efficiently.

OHDSI Shoutouts!

Congratulations to the team of **Inessa Cohen, Zihan Diao, Pawan Goyal, Aarti Gupta, Kathryn Hawk, Bill Malcom, Caitlin Malicki, Dhruv Sharma, Brian Sweeney, Scott Weiner, Arjun Venkatesh and Andrew Taylor** on the publication of **Mapping Emergency Medicine Data to the Observational Medical Outcomes Partnership Common Data Model: A Gap Analysis of the American College of Emergency Physicians Clinical Emergency Data Registry** in the *Journal of the American College of Emergency Physicians Open*.

JACEP OPEN
JOURNAL OF THE AMERICAN COLLEGE OF EMERGENCY PHYSICIANS OPEN
A GLOBAL JOURNAL OF EMERGENCY MEDICINE
Open Access
American College of Emergency Physicians
ADVANCING EMERGENCY CARE

BRIEF REPORT
The Practice of Emergency Medicine
Check for updates

Mapping Emergency Medicine Data to the Observational Medical Outcomes Partnership Common Data Model: A Gap Analysis of the American College of Emergency Physicians Clinical Emergency Data Registry

Inessa Cohen MPH[1,2,3], Zihan Diao AB[1], Pawan Goyal MD, MHA[4], Aarti Gupta MBBS[4], Kathryn Hawk MD, MHS[1], Bill Malcom PMP[4], Caitlin Malicki MPH[1], Dhruv Sharma MS[4], Brian Sweeney[4], Scott G. Weiner MD, MPH[5], Arjun Venkatesh MD, MBA[1,6], R. Andrew Taylor MD, MHS[1,2]

[1]*Department of Emergency Medicine, Yale School of Medicine, New Haven, Connecticut, USA*
[2]*Section for Biomedical Informatics and Data Science, Yale University School of Medicine, New Haven, Connecticut, USA*
[3]*Program of Computational Biology and Bioinformatics, Yale University, New Haven, Connecticut, USA*
[4]*American College of Emergency Physicians, Washington, DC, USA*
[5]*Department of Emergency Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA*
[6]*Center for Outcomes Research and Evaluation (CORE), Section of Cardiovascular Medicine, Yale School of Medicine, New Haven, Connecticut, USA*

Correspondence
R. Andrew Taylor, MD, MHS, Department of Emergency Medicine, Yale School of Medicine, New Haven, 464 Congress Ave. Suite 260, New Haven, CT 06519, USA. Email: richard.taylor@yale.edu

Abstract
Objectives: This study aims to conduct a gap analysis to determine the feasibility of mapping electronic health record data from the Clinical Emergency Data Registry (CEDR) to the Observational Medical Outcomes Partnership Common Data Model (OMOP-CDM).

OHDSI Shoutouts!

Congratulations to the team of **Seonji Kim, Subin Kim, Chungsoo Kim, Junhyuk Chang, Rae Woong Park, Kyung Won Kim, and Seng Chan You** on the publication of **Utility of Treatment Pattern Analysis Using a Common Data Model: A Scoping Review** in *Healthcare Informatics Research.*

Review Article

Healthc Inform Res. 2025 January;31(1):4-15.
https://doi.org/10.4258/hir.2025.31.1.4
pISSN 2093-3681 • eISSN 2093-369X

HIR
Healthcare Informatics Research

Utility of Treatment Pattern Analysis Using a Common Data Model: A Scoping Review

Eun-Gee Park[1], Min Jung Kim[2], Jinseo Kim[1], Kichul Shin[2], Borim Ryu[1]

[1]Center for Data Science, Biomedical Research Institute, Seoul Metropolitan Government–Seoul National University Boramae Medical Center, Seoul, Korea
[2]Division of Rheumatology, Department of Internal Medicine, Seoul Metropolitan Government–Seoul National University Boramae Medical Center, Seoul, Korea

Objectives: We aimed to derive observational research evidence on treatment patterns through a scoping review of common data model (CDM)-based publications. **Methods:** We searched the medical literature databases PubMed and EMBASE, as well as the Observational Health Data Sciences and Informatics (OHDSI) website, for papers published between January 1, 2010 and August 21, 2023 to identify research papers relevant to our topic. **Results:** Eighteen articles satisfied the inclusion criteria for this scoping review. We summarized study characteristics such as phenotypes, patient numbers, data periods, countries, Observational Medical Outcomes Partnership (OMOP) CDM databases, and definitions of index date and target cohort. Type 2 diabetes mellitus emerged as the most frequently studied disease, covered in five articles, followed by hypertension and depression, each addressed in four articles. Biguanides, with metformin as the primary drug, were the most commonly prescribed first-line treatments for type 2 diabetes mellitus. Most studies utilized sunburst plots to visualize treatment patterns, whereas two studies used Sankey plots. Various software tools were employed for treatment pattern analysis, including JavaScript, the open-source ATLAS by OHDSI, R code, and the R package "TreatmentPatterns." **Conclusions:** This study provides a comprehensive overview of research on treatment patterns using the CDM, highlighting the growing importance of OMOP CDM in enabling multinational observational network studies and advancing collaborative research in this field.

Keywords: Epidemiologic Methods, Cohort Studies, Drug Utilization, Scoping Review, Common Data Elements

OHDSI Shoutouts!

Congratulations to the team of **Yi Chai, Ivan Lam, Kenneth Man, Joseph Hayes, Eric Wan, Xue Li, Celine Chui, Wallis Lau, Xiaoyu Lin, Can Yin, Min Fan, Esther Chan, Ian Wong & Hao Luo** on the publication of **Psychiatric and neuropsychiatric sequelae of COVID-19 within 2 years: a multinational cohort study** in *BMC Medicine.*

RESEARCH Open Access

Psychiatric and neuropsychiatric sequelae of COVID-19 within 2 years: a multinational cohort study

Yi Chai[1,2,3†], Ivan C. H. Lam[2†], Kenneth K. C. Man[2,4,5], Joseph F. Hayes[6,7], Eric Y. F. Wan[2,5,8], Xue Li[2,5,9], Celine S. L. Chui[5,10,11], Wallis C. Y. Lau[2,4,5], Xiaoyu Lin[12], Can Yin[12], Min Fan[2], Esther W. Chan[2,5,13], Ian C. K. Wong[2,14] and Hao Luo[15*]

Abstract

Background The long-term psychiatric and neuropsychiatric sequelae of COVID-19 across diverse populations remain not fully understood. This cohort study aims to investigate the short-, medium-, and long-term risks of psychiatric and neuropsychiatric disorders following COVID-19 infection in five countries.

Methods This population-based multinational network study used electronic medical records from France, Italy, Germany, and the UK and claims data from the USA. The initial target and comparator cohorts were identified using an exact matching approach based on age and sex. Individuals diagnosed with COVID-19 or those with a positive SARS-CoV-2 screening test between December 1, 2019, and December 1, 2020, were included as targets. Up to ten comparators without COVID-19 for each target were selected using the propensity score matching approach. All individuals were followed from the index date until the end of continuous enrolment or the last healthcare encounter. Cox proportional hazard regression models were fitted to estimate the risk of incident diagnosis of depression, anxiety disorders, alcohol misuse or dependence, substance misuse or dependence, bipolar disorders, psychoses, personality disorders, self-harm and suicide, sleep disorders, dementia, and neurodevelopmental disorders within the first 6 months (short-term), 6 months to 1 year (medium-term), and 1 to 2 years (long-term) post-infection.

Results A total of 303,251 individuals with COVID-19 and 22,108,925 individuals without COVID-19 from five countries were originally included. Within the first 6 months, individuals with COVID-19 had a significantly higher risk of any studied disorders in all databases, with Hazard Ratios (HRs) ranging from 1.14 (95% CI, 1.07–1.22) in Germany to 1.89 (1.64–2.17) in Italy. Increased risks were consistently observed for depression, anxiety disorders, and sleep disorders across almost all countries. During the medium- and long-term periods, higher risks were observed only for depression (medium-term: 1.29, 1.18–1.41; long-term: 1.36, 1.25–1.47), anxiety disorders (medium-term: 1.29, 1.20–1.38; long-term: 1.37, 1.29–1.47), and sleep disorders (medium-term: 1.10, 1.01–1.21; long-term: 1.14, 1.05–1.24) in France, and dementia (medium-term: 1.65, 1.28–2.10) in the UK.

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

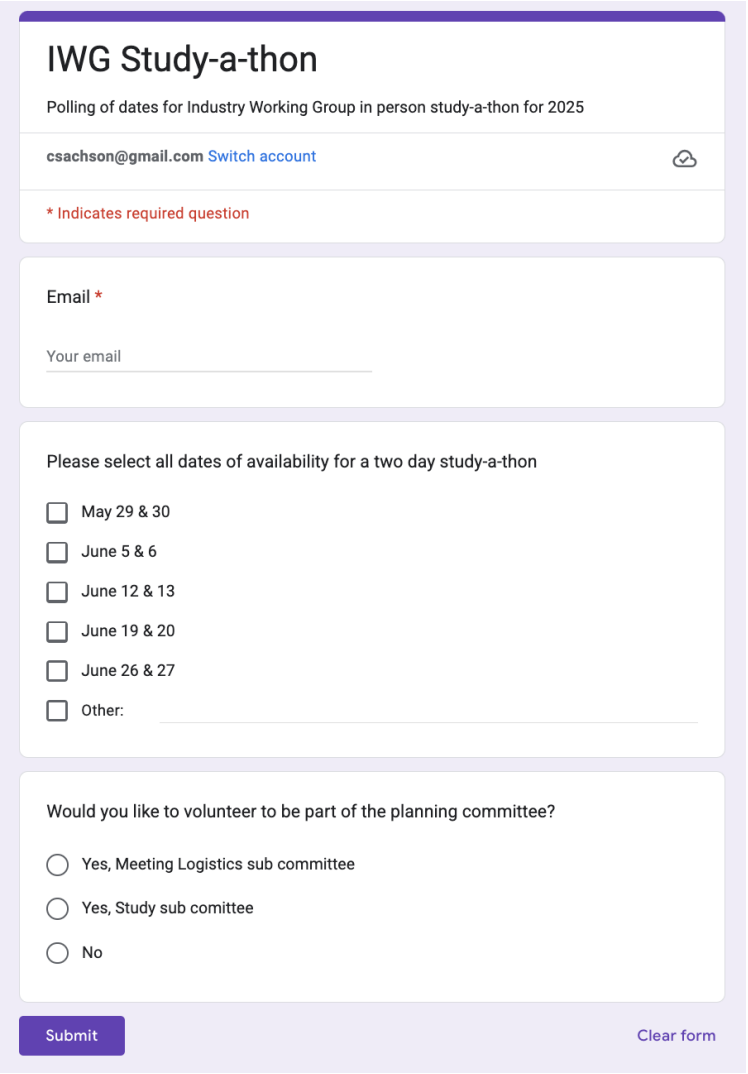
Upcoming Workgroup Calls

Date	Time (ET)	Meeting
Tuesday	12 pm	Generative AI and Analytics
Tuesday	12 pm	Atlas
Tuesday	3 pm	OMOP CDM Oncology Outreach/Research
Wednesday	9 am	Patient-Level Prediction
Wednesday	2 pm	Natural Language Processing
Wednesday	7 pm	Eyecare and Vision Research
Thursday	9:30 am	Network Data Quality
Thursday	10:30 am	Evidence Network
Thursday	12 pm	Strategus
Thursday	7 pm	Dentistry
Friday	9 am	Phenotype Development and Evaluation
Friday	10 am	GIS - Geographic Information System
Friday	11 am	Clinical Trials
Friday	11:30 am	Steering
Friday	11 pm	China Chapter
Monday	10 am	Getting Started Subgroup
Monday	10 am	Africa Chapter
Monday	11 am	Data Bricks User Group
Monday	2 pm	Electronic Animal Health Records

Industy WG Studyathon Questionnaire

The Industry WG is hosting an industry specific studaython (potential dates are May or June) and is looking for input on who might be interested and what dates work best.

If you are a member of the Industry WG or are considering joining the team, please fill out this survey by the end of this week.



IWG Study-a-thon

Polling of dates for Industry Working Group in person study-a-thon for 2025

csachson@gmail.com Switch account

* Indicates required question

Email *

Your email

Please select all dates of availability for a two day study-a-thon

- [ ] May 29 & 30
- [ ] June 5 & 6
- [ ] June 12 & 13
- [ ] June 19 & 20
- [ ] June 26 & 27
- [ ] Other:

Would you like to volunteer to be part of the planning committee?

- ( ) Yes, Meeting Logistics sub committee
- ( ) Yes, Study sub comittee
- ( ) No

Submit Clear form

OHDSI Europe Symposium - Save-the-date!

OHDSI BELGIUM

EUROPE

OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Save-the-date

5-7 July 2025

Location

Old Prison - Hasselt University
Martelarenlaan
Hasselt - BELGIUM

Global Symposium: Oct. 7-9

The 2025 OHDSI Global Symposium will return to the Hyatt Regency Hotel in New Brunswick, N.J., on Oct. 7-9.

The first Scientific Review Committee meeting will be held Thursday at 11 am.

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

Hongfang Liu, PhD

D. Bradley McWilliams Chair Professor of Biomedical Informatics, Vice President of Learning Health System, University of Texas Health Science Center at Houston

'A Translational Science Framework in Advancing Healthcare AI'

March 27, 2025, 11am-12pm EST
Virtually via Zoom

Please contact Marty Alvarez at malvarez2@tuftsmedicalcenter.org for calendar invite or questions.

TuftsMedicine
Tufts Medical Center

R/Medicine Conference

R/Medicine 2025

AN R CONSORTIUM VIRTUAL CONFERENCE

Call for Proposals

Submission is now open!

SUBMIT HERE

We are seeking abstracts for:

- **Lightning talks** (10 min, Thursday June 12 or Friday June 13) Must pre-record and be live on chat to answer questions
- **Regular talks** (20 min, Thursday June 12 or Friday June 13) Must pre-record and be live on chat to answer questions
- **Demos** (1 hour demo of an approach or a package, Tuesday June 10 or Wednesday June 11) Done live, preferably interactive
- **Workshops** (2-3 hours on a topic, Tuesday June 10 or Wednesday June 11) Detailed instruction on a topic, usually with a website and a repo, participants can choose to code along, include 5-10 min breaks each hour.

An abstract proposal should describe a presentation of methodology, a study or project, or an example or case study relevant to one (or more) of our fields of interest, with R-based tools having a substantial role in the work. The content of the presentation should be of interest to the R/Medicine community.

Dates to Remember

CFP Closes: Friday, April 11th at 11:59 PM EDT

CFP Notifications: Friday, April 25th

Schedule Announcement: Friday, May 9th

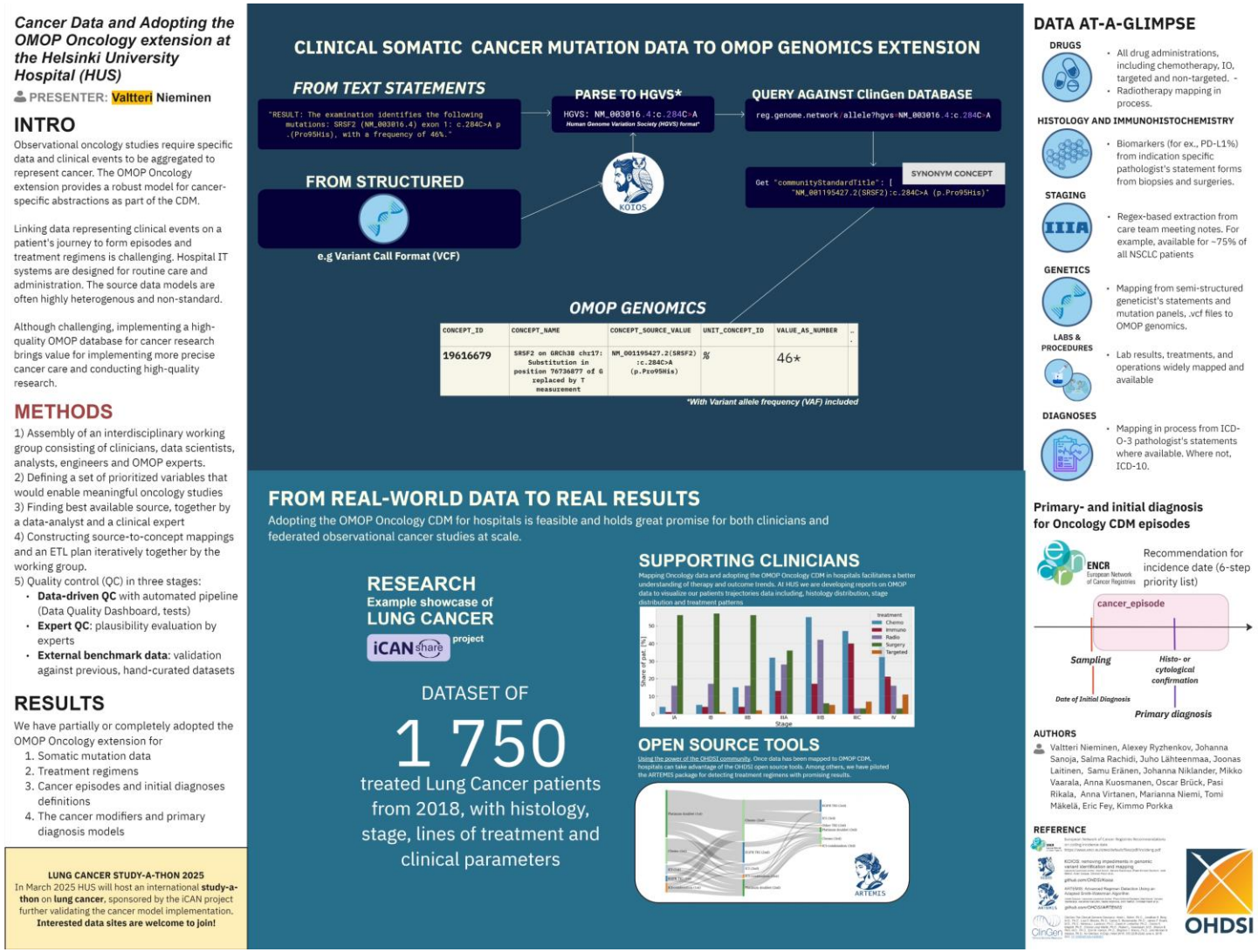
Pre-Recorded Video Submission: Monday, June 2nd

#OHDSISocialShowcase This Week

Monday

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)



#OHDSISocialShowcase This Week

Tuesday

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)

ETLing from your OMOP CDM

An efficient solution to vocabulary migration and data quality enhancement

PRESENTER: **Clair Blacketer**

INTRO:

- Updating the vocabulary of an existing CDM instance usually necessitates a complete rerun of the ETL process.
- This is especially problematic when access to the native data is restricted.
- We present a novel ETL method that allows for vocabulary updates without reverting to the native data.

METHODS

Source Concept ID Re-Mapping

Original Source Concept ID → Corrected Source Concept ID

source code + valid dates

Concept Code	Concept Name	Valid Dates	Mapped Concept
J0180	Vinblastine	Valid until 2018	1582881
J0180	Trastuzumab	Valid from 2020	761883

Data Quality Rules Applied

Implement logic to correct any data quality issues observed in original CDM.

Standard Concept Mapping & Movement

- The primary focus of the ETL was the event tables, where most vocabulary mappings occur.
- The re-mapped source codes from step 1 were converted to standard concepts and the corresponding event tables were populated based on the target concept domain.
- This required all primary keys to be reissued.

Operationalize ETL Logic

- Developed code to operationalize the CDM-to-CDM ETL logic.
- Transferred any custom vocabularies from the original CDM.

Quality Assurance

The DQD was used post generation of the new CDM instance with updated vocabulary to identify any errors or issues in the concept mappings.

Application

The method was applied to IQVIA® Adjudicated Health Plan Data to update the vocabulary version from 31-Aug-2023 to 29-Feb-2024.

You can keep your **vocabulary** up-to-date by ETLing your current **CDM** to a new **CDM** using our tool.

Take a picture to download the short report

RESULTS

Link to full CDM-to-CDM ETL specifications

- **13,556** persons without gender were removed
- **379,586** standard concepts were added to the OHDSI Vocabularies between versions 31-Aug-2023 and 29-Feb-2024.

Figure 1: Number of records in IQVIA Pharmetrics Plus mapped to new concepts between vocabulary versions 31-Aug-2023 and 29-Feb-2024, by table.

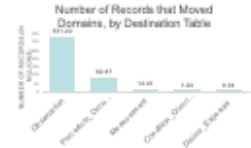


Figure 2: Number of records in IQVIA Pharmetrics Plus that moved domains between vocabulary versions 31-Aug-2023 and 29-Feb-2024, by destination table.

Key considerations

1. Domain changes
2. Multi-domain mappings
3. Primary key re-issuance
4. Database-specific vocabulary integration
5. Data Quality Dashboard

Clair Blacketer[1,2], Anton Ivanov[1], Evanette Burrows[1], Dmitry Dymshyts[1], Frank DeFalco[1]

[1]Janssen Research & Development, Raritan, NJ
[2]Department of Medical Informatics, Erasmus MC, Rotterdam, NL

Johnson&Johnson

Erasmus MC

OHDSI

#OHDSISocialShowcase This Week

Wednesday

Comparative Analysis of OMOP CDM Database Profiles Across Institutions and Future Research Implications

(Haeun Lee, Snehil Gupta, Clair Blacketer, Michael Cook, Shinji Naka, Ruochong Fan, Benjamin Martin, Khyzer Aziz, Linying Zhang, Paul Nagy)

Comparative Analysis of OMOP CDM Database Profiles Across Institutions and Future Research Implications

PRESENTER: Haeun Lee

INTRO

- The OHDSI community has facilitated network studies through a federated approach, yet assessing institutional data attributes and suitability within this distributed framework remains challenging.
- Database profiles provide detailed insights into OMOP CDM databases, enabling the evaluation of data characterization and study eligibility for multicenter studies.
- This study aims to examine Database profiles, assess data patterns, and explain their implications for future research within the OHDSI community.

METHODS

1. Data sources and Tools:
Utilized DbDiagnostics (v1.3.1) to analyze OMOP CDM databases from Johns Hopkins School of Medicine and Washington University School of Medicine in St.Louis.
2. Analysis:
Generated summary statistics for various CDM domains and compared common concepts, data density, and data completeness between datasets.
3. Evaluation:
Assessed data characterization and consistency to evaluate OMOP CDM implementation across institutions.

Table 1. Summary of Common Concepts Across Two Institutions

Concept Category	Site A (Concept ID)	Site B (Concept ID)
Race		
White	8527	8527
Black or African American	8516	8516
Asian	8515	8515
American Indian or Alaska Native	8657	8657
Other Pacific Islander	38003613	38003613
Visit		
Inpatient	9201	9201
Outpatient	9202	9202
Emergency room visit	9203	9203
Intensive care visit	32037	32037
Condition		
Essential hypertension	320128	320128
Type 2 diabetes mellitus without complication	4193704	4193704
Obesity	433736	433736
Gastroesophageal reflux disease without esophagitis	4144111	4144111
Procedure		
Cesarean section	2110316	4015701
Total knee replacement	2105103	43531648
Cholecystectomy	2109368	4242997
Laparoscopic Appendectomy	2109144	4243973
Measurement		
Body height	3036277	3036277
Body weight	3025315	3025315
Body mass index	40762636	3038553
Body temperature	3020891	3020891
Systolic blood pressure	3004249	3004249
Diastolic blood pressure	3012888	3012888
Respiratory rate	3024171	4313591
Heart rate (Pulse rate)	3027018	4301868

Table 2. Key Data Domains Across Institutions

Domain	Site A	Site B
Person	Y	Y
Visit Occurrence	Y	Y
Condition Occurrence	Y	Y
Procedure Occurrence	Y	Y
Measurement	Y	Y
Drug Exposure	Y	Y
Observation	Y	Y
Care Site	N	Y
Device Exposure	Y	N

Table 3. Top 5 Most Common Concepts by Domain and Concept Ids Across Institutions

Domain	Site A	Site B
Condition	Essential hypertension (320128)	Patient encounter procedure (4203722)
	Hyperlipidemia (432867)	History of event (1340204)
	Type2 diabetes mellitus without complication (4193704)	Essential hypertension (320128)
	Gastroesophageal reflux disease without esophagitis (4144111)	Hyperlipidemia (432867)
	Postoperative state (438485)	Type2 diabetes mellitus without complication (4193704)
Measurement	Heart rate (3027018)	Hemoglobin [Mass/volume] in Blood (3000963)
	Diastolic blood pressure (3012888)	Body weight (3025315)
	Systolic blood pressure (3004249)	Platelets [#/volume] in Blood by Automated count (3024929)
	Respiratory rate (3024171)	No matching concept (0)
	Oxygen saturation in Arterial blood by Pulse oximetry (40762499)	Respiratory rate (4313591)
	Potassium chloride 0.004 MEQ (19135374)	No matching concept (0)
Medication	SARS-COV-2 (COVID-19) vaccine (724906)	100 ML sodium chloride 9 MG/ML Injection (40220357)
	SARS-COV-2 (COVID-19) vaccine (724907)	10 ML sodium chloride 9 MG/ML Prefilled Syringe (19127213)
	oxycodone hydrochloride 5 MG Oral Tablet (40232756)	100 ML sodium chloride 9 MG/ML Injection (40221385)
	2 ML ondansetron 2 MG/ ML Injection (35605482)	2 ML ondansetron 2 MG/ML Injection (35605482)

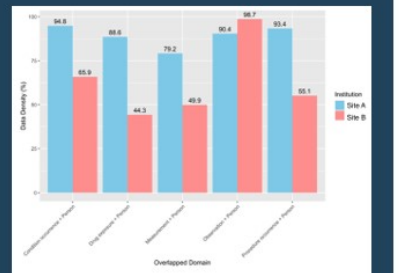


Figure 1. Comparison of Data Density for Different Domains Between Institutions

Scan QR to go to OHDSI Evidence Network Community

RESULTS

- The study analyzed data from 2.1 M patients at Site A and 8 M patients at Site B, revealing data quality issues and gaps in demographic information.
- While both institutions used SNOMED CT codes for condition concepts, differences were observed in vocabularies for common procedures and measurements.
- Essential data domains were present at both sites, though care site information was excluded from Site A's OMOP instance.
- Data density varied between Site A and Site B across overlapped domains, with Site A generally showing higher densities except in Observations.

CONCLUSIONS

- Database profiles are crucial for understanding institutional characteristics and their ETL and mapping techniques.
- Addressing variations between institutions is essential to improve data consistency and provide critical insights for evaluating study eligibility based on specific research questions.
- Sharing database profiles enhances network studies through improved collaboration, and cross-institutional transparency.

Haeun Lee, Snehil Gupta, Clair Blacketer, Michael Cook, Shinji Naka, Ruochong Fan, Benjamin Martin, Khyzer Aziz, Linying Zhang, Paul Nagy

Johns Hopkins School of Medicine · Johnson&Johnson · Washington University in St.Louis School of Medicine · OHDSI

#OHDSISocialShowcase This Week

Thursday

Atlas2AoU: Enabling Comparison of OHDSI Phenotype Library Phenomic Profiles in All of Us and the UK Biobank

(Abigail Newbury, Xinzhuo Jiang, Karthik Natarajan, Gamze Gürsoy)

Atlas2AoU:
Enabling Comparison of OHDSI Phenotype Library Phenomic Profiles in All of Us and the UK Biobank

PRESENTER: **Abigail** Newbury

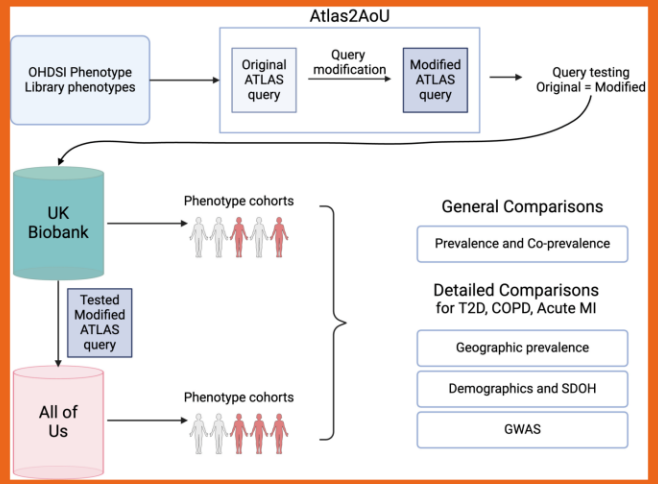
INTRO:

- **Biobanks facilitate precision medicine efforts to improve understanding of disease etiology**
- **Biobanks are increasingly providing clinical data in OMOP CDM format, which allows researchers to use complex rule-based phenotyping algorithms**
- **These complex phenotyping algorithms can result in more accurate patient cohorts**
- **Due to technical limitations in the All of Us (AoU) Researcher Workbench, direct application of OHDSI Phenotype Library (PL) phenotype algorithms (constructed by ATLAS) has been challenging**
- **We propose a tool Atlas2AoU to enable OHDSI PL cohort creation in the AoU workbench**

METHODS

1. Atlas2AoU modifies ATLAS queries by calling a temporary observation period table created by taking the minimum and maximum dates of recorded clinical events across 9 OMOP CDM tables
2. We illustrate the use of Atlas2AoU by comparing the prevalence of 423 cohorts from the OHDSI Phenotype Library v3.1.6 between AoU and the UK Biobank (UKBB)
3. We emphasize the importance of cross-biobank comparison by comparing geographic disease prevalence, demographics and social determinants of health (SDOH), and results of genome-wide association studies (GWAS) between biobanks for Chronic Obstructive Pulmonary Disease (COPD).

Atlas2AoU gives researchers the ability to create **ATLAS phenotypes** in the **All of Us Workbench** which furthers **precision medicine** efforts & enables **cross-biobank comparison**



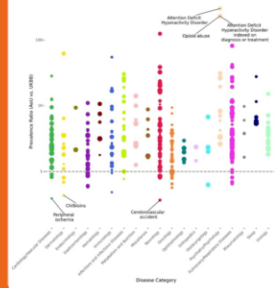
Take a picture to download the full paper

Impact of Atlas2AoU

Phenotype	Original query false positives	Original query false negatives	Original query cohort size	Atlas2AoU query cohort size
T2D	370	43,111	7,503	50,244
COPD	488	23,044	2,256	24,812
Acute MI	0	6,131	194	6,325

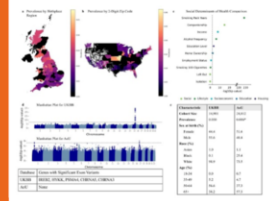
Atlas2AoU query modification leads to **accurate cohorts** and **increases sample size**

Prevalence comparison



AoU has significantly higher prevalence compared to the UKBB for 335 of the 423 OHDSI PL phenotypes

Detailed COPD comparison



- Significantly **higher prevalence** in AoU
- **Regions of high prevalence:** UKBB: West Dunbartonshire, Glasgow City, and West Lothian; AoU: Kansas and Missouri
- **SDOH variables:** Smoking pack years, companionship, income, alcohol frequency and education level significantly different among cohorts
- **GWAS:** variants in the exons of HYKK, CHRNA3, CHRNA4, and CHRNA significantly associated with COPD in the UKBB

Abigail Newbury, Xinzhuo Jiang, Karthik Natarajan, Gamze Gürsoy

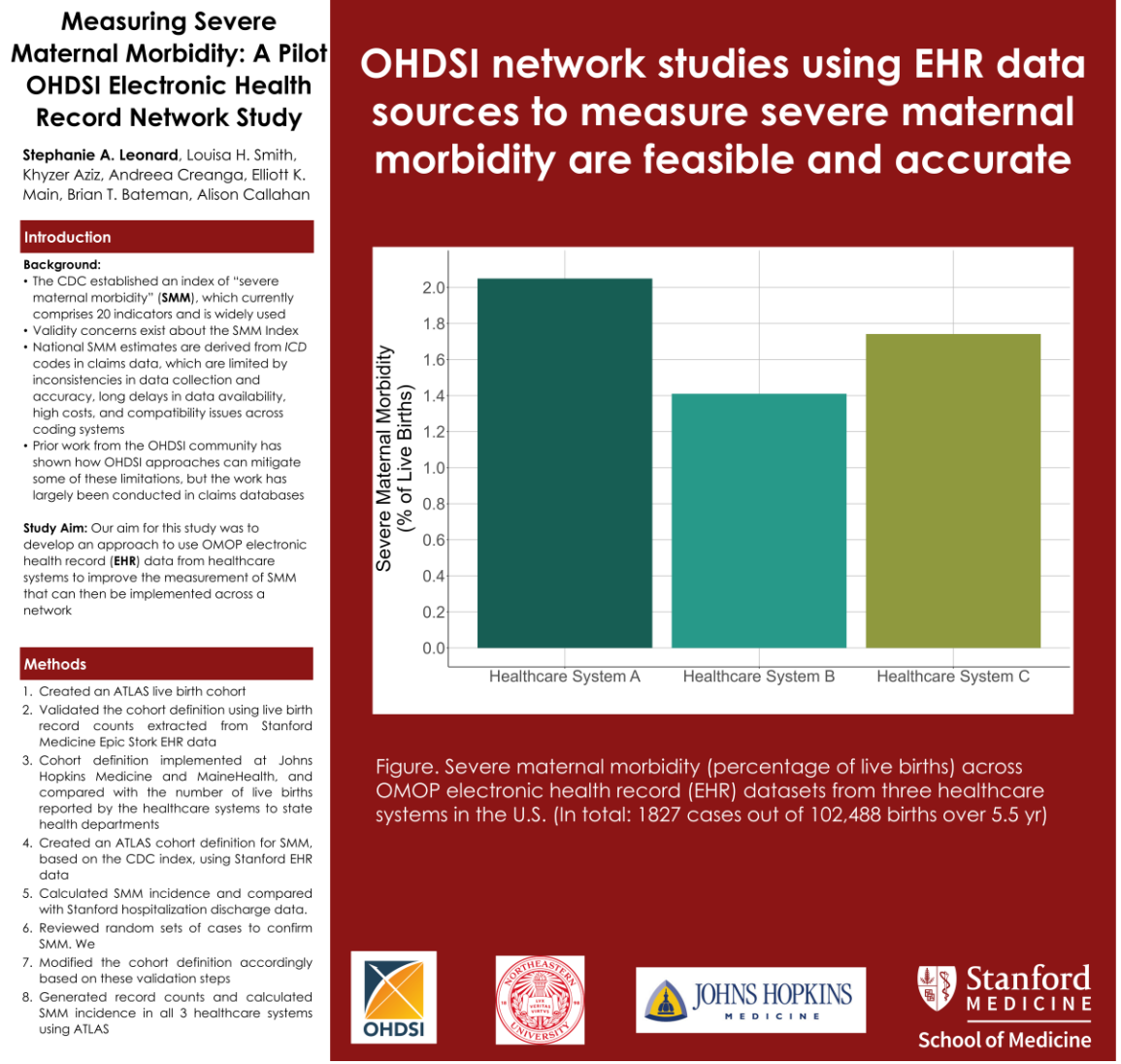
NEW YORK GENOME CENTER · COLUMBIA UNIVERSITY DEPARTMENT OF BIOMEDICAL INFORMATICS · OHDSI

#OHDSISocialShowcase This Week

Friday

Measuring Severe Maternal Morbidity: A Pilot OHDSI Electronic Health Record Network Study

(Stephanie A. Leonard, Louisa H. Smith, Khyzer Aziz, Andreea Creanga, Elliott K. Main, Brian T. Bateman, Alison Callahan)



Measuring Severe Maternal Morbidity: A Pilot OHDSI Electronic Health Record Network Study

Stephanie A. Leonard, Louisa H. Smith, Khyzer Aziz, Andreea Creanga, Elliott K. Main, Brian T. Bateman, Alison Callahan

Introduction

Background:
- The CDC established an index of "severe maternal morbidity" (**SMM**), which currently comprises 20 indicators and is widely used
- Validity concerns exist about the SMM Index
- National SMM estimates are derived from *ICD* codes in claims data, which are limited by inconsistencies in data collection and accuracy, long delays in data availability, high costs, and compatibility issues across coding systems
- Prior work from the OHDSI community has shown how OHDSI approaches can mitigate some of these limitations, but the work has largely been conducted in claims databases

Study Aim: Our aim for this study was to develop an approach to use OMOP electronic health record (**EHR**) data from healthcare systems to improve the measurement of SMM that can then be implemented across a network

Methods

1. Created an ATLAS live birth cohort
2. Validated the cohort definition using live birth record counts extracted from Stanford Medicine Epic Stork EHR data
3. Cohort definition implemented at Johns Hopkins Medicine and MaineHealth, and compared with the number of live births reported by the healthcare systems to state health departments
4. Created an ATLAS cohort definition for SMM, based on the CDC index, using Stanford EHR data
5. Calculated SMM incidence and compared with Stanford hospitalization discharge data.
6. Reviewed random sets of cases to confirm SMM. We
7. Modified the cohort definition accordingly based on these validation steps
8. Generated record counts and calculated SMM incidence in all 3 healthcare systems using ATLAS

OHDSI network studies using EHR data sources to measure severe maternal morbidity are feasible and accurate

Figure. Severe maternal morbidity (percentage of live births) across OMOP electronic health record (EHR) datasets from three healthcare systems in the U.S. (In total: 1827 cases out of 102,488 births over 5.5 yr)

OHDSI · Johns Hopkins Medicine · Stanford Medicine School of Medicine

Results

- We successfully used ATLAS to generate SMM cohorts using OMOP EHR data from 3 healthcare systems
- Each system had 24,000-40,000 births (total N = **102,488 births**)
- SMM incidence from January 1, 2019 to May 31, 2024 at Healthcare System A was 2.05%, at Healthcare System B was 1.41%, and at Healthcare System C was 1.74% (**Figure**)
- Across the 3 systems, the live birth cohort definition **captured 92-99% of live births** as reported to state health departments and in Stork EHR data
- The incidence of SMM using the cohort definition at Stanford was **within 0.05 percentage points** of the annual hospital-reported incidence

Conclusion

- This pilot study generated evidence that supports the feasibility and accuracy of OHDSI network studies using EHR data sources to measure SMM
- Challenges unique to OMOP obstetric EHR data were identified and addressed
- Our approaches are informative for perinatal health OHDSI network studies and could facilitate an impactful addition to claims databases or aggregated multi-site EHR data
- Based on our results, we are expanding the study to a broad national network of healthcare systems with ATLAS and EHR data in the OMOP CDM
- Leveraging this network, we will evaluate how proposed modifications to the SMM Index would affect incidence and we will characterize patient cohorts identified under the proposed modified cohort definitions.

Author Affiliations

- Stanford University: SAL, EKM, BTB, AC
- Johns Hopkins University: KA, AC
- Northeastern University: LHS

OHDSI Workgroup

This work is part of the OHDSI Perinatal and Reproductive Health Group (**PRHeG**)

Contact **leonard@ohdsi.org** if you are interested in joining the workgroup and/or this network study

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-2025