
# ATLAS Deepdive: Cohorts and Concept Sets

## OHDSI Community Call
## June 17, 2025 • 11 am ET

# Upcoming Community Calls

| Date | Topic |
|---|---|
| June 17 | ATLAS Deepdive: Cohorts and Conceptsets |
| June 24 | ATLAS Deepdive: Characterization, Cohort Pathways, Incidence |
| 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

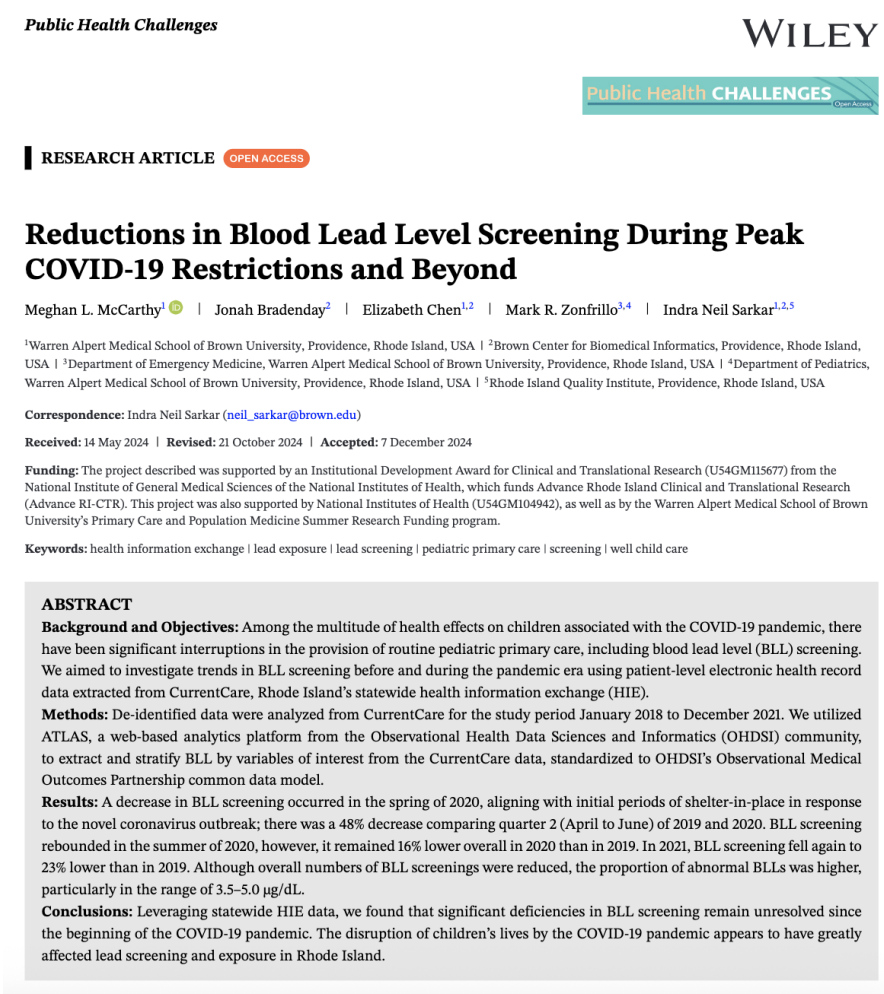
## Where Have We Been?

## Where Are We Now?

## Where Are We Going?

# OHDSI Shoutouts!

Congratulations to the team of **Meghan L. McCarthy, Jonah Bradenday, Elizabeth Chen, Mark R. Zonfrillo, and Indra Neil Sarkar** on the publication of **Reductions in Blood Lead Level Screening During Peak COVID-19 Restrictions and Beyond** in *Public Health Challenges.*



*Public Health Challenges*

WILEY

Public Health CHALLENGES

RESEARCH ARTICLE OPEN ACCESS

**Reductions in Blood Lead Level Screening During Peak COVID-19 Restrictions and Beyond**

Meghan L. McCarthy[1] | Jonah Bradenday[2] | Elizabeth Chen[1,2] | Mark R. Zonfrillo[3,4] | Indra Neil Sarkar[1,2,5]

[1]Warren Alpert Medical School of Brown University, Providence, Rhode Island, USA | [2]Brown Center for Biomedical Informatics, Providence, Rhode Island, USA | [3]Department of Emergency Medicine, Warren Alpert Medical School of Brown University, Providence, Rhode Island, USA | [4]Department of Pediatrics, Warren Alpert Medical School of Brown University, Providence, Rhode Island, USA | [5]Rhode Island Quality Institute, Providence, Rhode Island, USA

**Correspondence:** Indra Neil Sarkar (neil_sarkar@brown.edu)

**Received:** 14 May 2024 | **Revised:** 21 October 2024 | **Accepted:** 7 December 2024

**Funding:** The project described was supported by an Institutional Development Award for Clinical and Translational Research (U54GM115677) from the National Institute of General Medical Sciences of the National Institutes of Health, which funds Advance Rhode Island Clinical and Translational Research (Advance RI-CTR). This project was also supported by National Institutes of Health (U54GM104942), as well as by the Warren Alpert Medical School of Brown University's Primary Care and Population Medicine Summer Research Funding program.

**Keywords:** health information exchange | lead exposure | lead screening | pediatric primary care | screening | well child care

**ABSTRACT**

**Background and Objectives:** Among the multitude of health effects on children associated with the COVID-19 pandemic, there have been significant interruptions in the provision of routine pediatric primary care, including blood lead level (BLL) screening. We aimed to investigate trends in BLL screening before and during the pandemic era using patient-level electronic health record data extracted from CurrentCare, Rhode Island's statewide health information exchange (HIE).

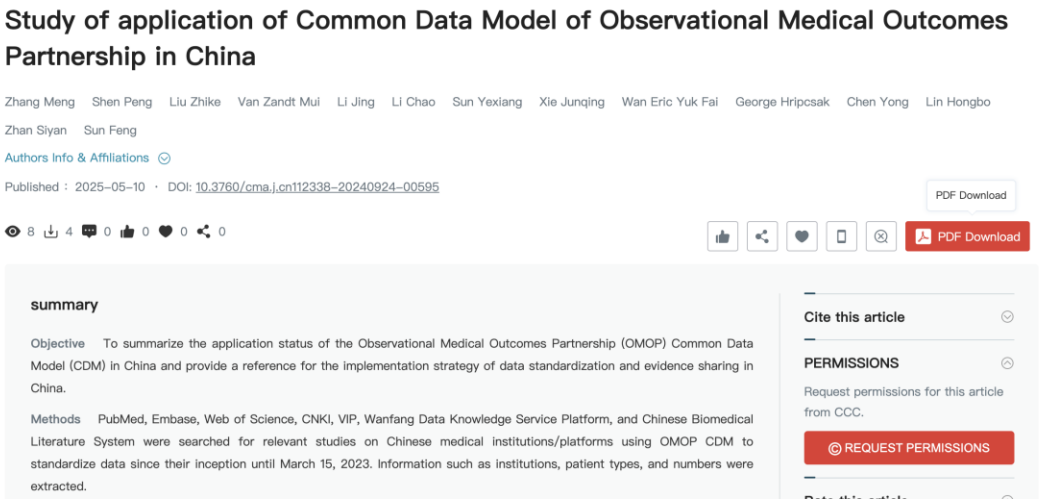
**Methods:** De-identified data were analyzed from CurrentCare for the study period January 2018 to December 2021. We utilized ATLAS, a web-based analytics platform from the Observational Health Data Sciences and Informatics (OHDSI) community, to extract and stratify BLL by variables of interest from the CurrentCare data, standardized to OHDSI's Observational Medical Outcomes Partnership common data model.

**Results:** A decrease in BLL screening occurred in the spring of 2020, aligning with initial periods of shelter-in-place in response to the novel coronavirus outbreak; there was a 48% decrease comparing quarter 2 (April to June) of 2019 and 2020. BLL screening rebounded in the summer of 2020, however, it remained 16% lower overall in 2020 than in 2019. In 2021, BLL screening fell again to 23% lower than in 2019. Although overall numbers of BLL screenings were reduced, the proportion of abnormal BLLs was higher, particularly in the range of 3.5–5.0 μg/dL.

**Conclusions:** Leveraging statewide HIE data, we found that significant deficiencies in BLL screening remain unresolved since the beginning of the COVID-19 pandemic. The disruption of children's lives by the COVID-19 pandemic appears to have greatly affected lead screening and exposure in Rhode Island.

# OHDSI Shoutouts!

Congratulations to the team of **Zhang Meng, Shen Peng, Liu Zhike, Mui Van Zandt, Li Jing, Li Chao, Sun Yexiang, Xie Junqing, Wan Eric Yuk Fai, George Hripcsak, Chen Yong, Lin Hongbo, Zhan Siyan, and Sun Feng** on the publication of **Study of application of Common Data Model of Observational Medical Outcomes Partnership in China** in the *Chinese Journal of Epidemiology.*



Study of application of Common Data Model of Observational Medical Outcomes Partnership in China

Zhang Meng Shen Peng Liu Zhike Van Zandt Mui Li Jing Li Chao Sun Yexiang Xie Junqing Wan Eric Yuk Fai George Hripcsak Chen Yong Lin Hongbo Zhan Siyan Sun Feng

Authors Info & Affiliations

Published：2025-05-10 · DOI: 10.3760/cma.j.cn112338-20240924-00595

PDF Download

8 4 0 0 0 0

PDF Download

summary

Objective To summarize the application status of the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) in China and provide a reference for the implementation strategy of data standardization and evidence sharing in China.

Methods PubMed, Embase, Web of Science, CNKI, VIP, Wanfang Data Knowledge Service Platform, and Chinese Biomedical Literature System were searched for relevant studies on Chinese medical institutions/platforms using OMOP CDM to standardize data since their inception until March 15, 2023. Information such as institutions, patient types, and numbers were extracted.

Cite this article

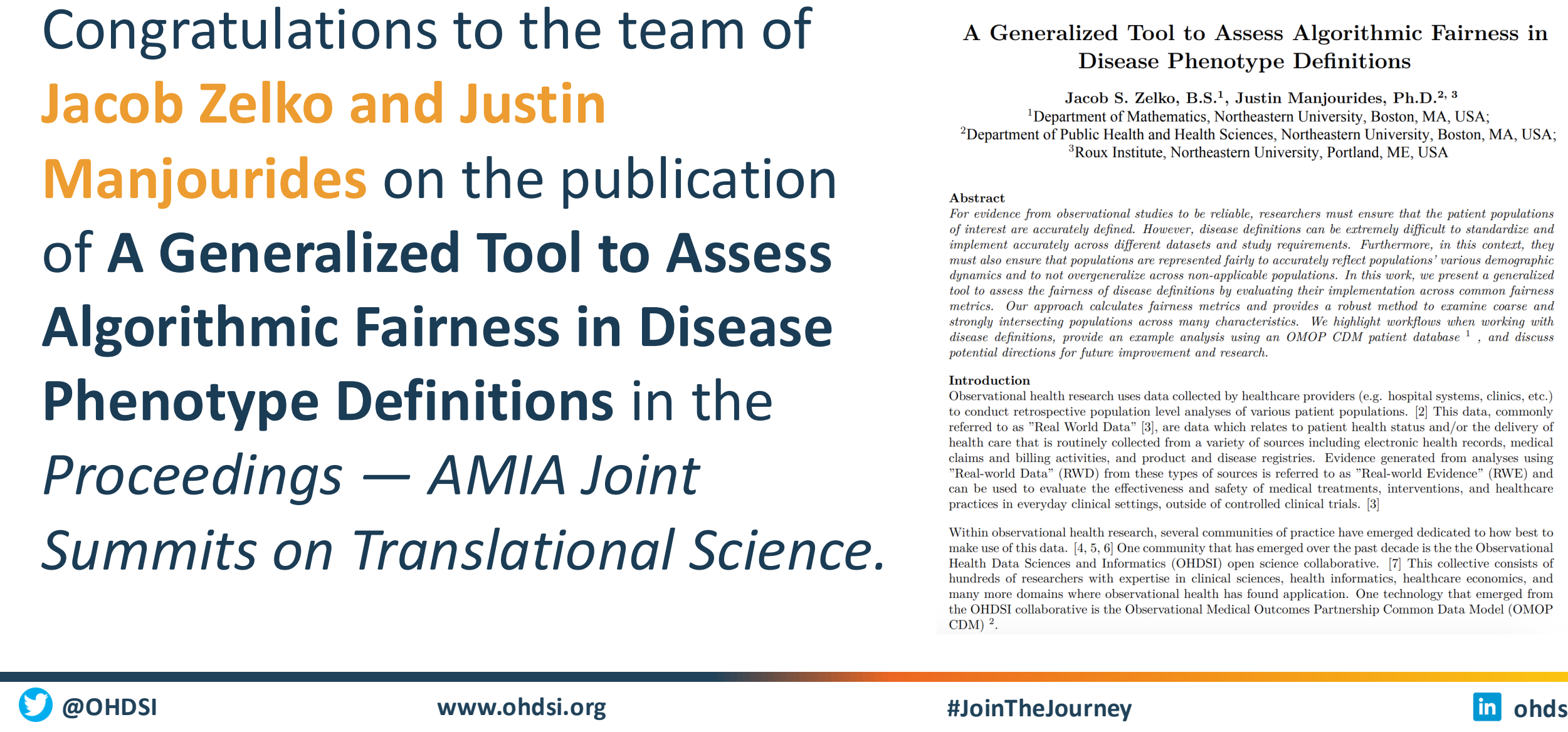
PERMISSIONS

Request permissions for this article from CCC.

REQUEST PERMISSIONS

# OHDSI Shoutouts!

Congratulations to the team of **Jacob Zelko and Justin Manjourides** on the publication of **A Generalized Tool to Assess Algorithmic Fairness in Disease Phenotype Definitions** in the *Proceedings — AMIA Joint Summits on Translational Science.*



**A Generalized Tool to Assess Algorithmic Fairness in Disease Phenotype Definitions**


Jacob S. Zelko, B.S.[1], Justin Manjourides, Ph.D.[2, 3]
[1]Department of Mathematics, Northeastern University, Boston, MA, USA;
[2]Department of Public Health and Health Sciences, Northeastern University, Boston, MA, USA;
[3]Roux Institute, Northeastern University, Portland, ME, USA



**Abstract**

*For evidence from observational studies to be reliable, researchers must ensure that the patient populations of interest are accurately defined. However, disease definitions can be extremely difficult to standardize and implement accurately across different datasets and study requirements. Furthermore, in this context, they must also ensure that populations are represented fairly to accurately reflect populations' various demographic dynamics and to not overgeneralize across non-applicable populations. In this work, we present a generalized tool to assess the fairness of disease definitions by evaluating their implementation across common fairness metrics. Our approach calculates fairness metrics and provides a robust method to examine coarse and strongly intersecting populations across many characteristics. We highlight workflows when working with disease definitions, provide an example analysis using an OMOP CDM patient database [1] , and discuss potential directions for future improvement and research.*


**Introduction**

Observational health research uses data collected by healthcare providers (e.g. hospital systems, clinics, etc.) to conduct retrospective population level analyses of various patient populations. [2] This data, commonly referred to as "Real World Data" [3], are data which relates to patient health status and/or the delivery of health care that is routinely collected from a variety of sources including electronic health records, medical claims and billing activities, and product and disease registries. Evidence generated from analyses using "Real-world Data" (RWD) from these types of sources is referred to as "Real-world Evidence" (RWE) and can be used to evaluate the effectiveness and safety of medical treatments, interventions, and healthcare practices in everyday clinical settings, outside of controlled clinical trials. [3]

Within observational health research, several communities of practice have emerged dedicated to how best to make use of this data. [4, 5, 6] One community that has emerged over the past decade is the the Observational Health Data Sciences and Informatics (OHDSI) open science collaborative. [7] This collective consists of hundreds of researchers with expertise in clinical sciences, health informatics, healthcare economics, and many more domains where observational health has found application. One technology that emerged from the OHDSI collaborative is the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM) [2].

# OHDSI Shoutouts!

Congratulations to the team of **Seok Kim, Dachung Boo, Sooyoung Yoo, Borham Kim, Kyubo Kim, Kwangsoo Kim, Eunhye Song, Junmo Kim, Hyun Gee Ryoo, Jin Chul Paeng, In Young Choi, SooJeong Ko, Ie Ryung Yoo, Rae Woong Park, and Ho-Young Lee** on the publication of **Secondary Cancer Risk in Breast Cancer with and without Radiotherapy: The Observational Health Data Sciences and Informatics (OHDSI) Cohort Study** in *Cancer Research and Treatment.*

CANCER RESEARCH AND TREATMENT

pISSN: 1598-2998 eISSN: 200

Keyword, Title, Abstract, or Author | Search | Advanced Search

ABOUT | BROWSE ARTICLES | FOR AUTHORS AND REVIEWERS

HOME > J Korean Cancer Assoc > Accepted articles > Article

Original Article

**Secondary Cancer Risk in Breast Cancer with and without Radiotherapy: The Observational Health Data Sciences and Informatics (OHDSI) Cohort Study**

Seok Kim[1], Dachung Boo[2], Sooyoung Yoo[1], Borham Kim[1], Kyubo Kim[3], Kwangsoo Kim[4,5], Eunhye Song[6], Junmo Kim[7], Hyun Gee Ryoo[8,9], Jin Chul Paeng[9], In Young Choi[10], SooJeong Ko[10], Ie Ryung Yoo[11], Rae Woong Park[12], Ho-Young Lee[1,8]

DOI: https://doi.org/10.4143/crt.2024.968 [Accepted]
Published online: June 5, 2025

Check for updates

Author information | Article notes

# OHDSI Shoutouts!

Congratulations to the team of **Kim López-Güell, Martí Català, Daniel Dedman, Talita Duarte-Salles, Raivo Kolde, Raúl López-Blasco, Álvaro Martínez, Gregoire Mercier, Alicia Abellan, Johnmary T. Arinze, Theresa Burkard, Edward Burn, Zara Cuccu, Antonella Delmestri, Dominique Delseny, Sara Khalid, Chungsoo Kim, Ji-woo Kim, Kristin Kostka, Cora Loste, Miguel A. Mayer, Jaime Meléndez-Cardiel, Núria Mercadé-Besora, Mees Mosseveld, Akihito Nishimura, Hedvig ME. Nordeng, Jessie O. Oyinlola, Roger Paredes, Laura Pérez-Crespo, Marta Pineda-Moncusí, Juan Manuel Ramírez-Anguita, Nhung TH. Trinh, Anneli Uusküla, Bernardo Valdivieso, Daniel Prieto-Alhambra, Junqing Xie, Lourdes Mateu, and Annika M. Jödicke** on the publication of **Clusters of post-acute COVID-19 symptoms: a latent class analysis across 9 databases and 7 countries** in the *Journal of Clinical Epidemiology.*

Journal of Clinical Epidemiology
Available online 13 June 2025, 111867
In Press, Journal Pre-proof What's this?

Original Research

## Clusters of post-acute COVID-19 symptoms: a latent class analysis across 9 databases and 7 countries

Kim López-Güell [1 #], Martí Català [1 #], Daniel Dedman [2], Talita Duarte-Salles [3 4], Raivo Kolde [5], Raúl López-Blasco [6], Álvaro Martínez [7], Gregoire Mercier [9 10], Alicia Abellan [3], Johnmary T. Arinze [4], Theresa Burkard [1], Edward Burn [1], Zara Cuccu [2], Antonella Delmestri [1], Dominique Delseny [9], Sara Khalid [1], Chungsoo Kim [11], Ji-woo Kim [12], Kristin Kostka [1 13], Cora Loste [8 21 23]... Annika M. Jödicke [1 †]

Show more

Add to Mendeley Share Cite

https://doi.org/10.1016/j.jclinepi.2025.111867 Get rights and content



### Highlights

- 787,078 persons with PCC were included, which makes this study the largest international PCC study to our knowledge
- Complex multi-symptomatic clusters included anxiety-depression, abdominal-gastrointestinal symptoms, and respiratory problems with fatigue and joint pain

# Three Stages of The Journey

## Where Have We Been?

## Where Are We Now?

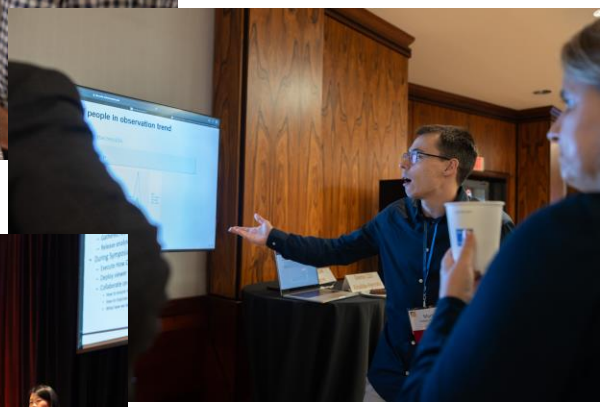
## Where Are We Going?

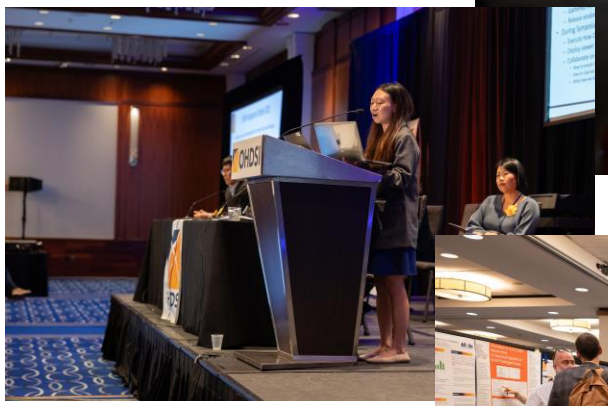
# Upcoming Workgroup Calls

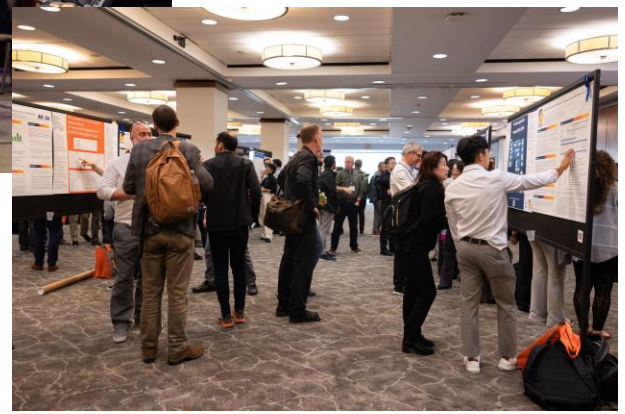
| Date | Time (ET) | Meeting |
|---|---|---|
| Tuesday | 12 pm | ATLAS |
| Thursday | 8 am | India Community Call |
| Thursday | 9 am | Oncology Vocabulary/Development Subgroup |
| Thursday | 11 am | Themis |
| Thursday | 12 pm | HADES |
| Friday | 10 am | GIS-Geographic Information System |
| Friday | 11:30 am | Steering |
| 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 |

# TWO Weeks Remaining

**The submission deadline for the 2025 Global Symposium Collaborator Showcase is July 1.**







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

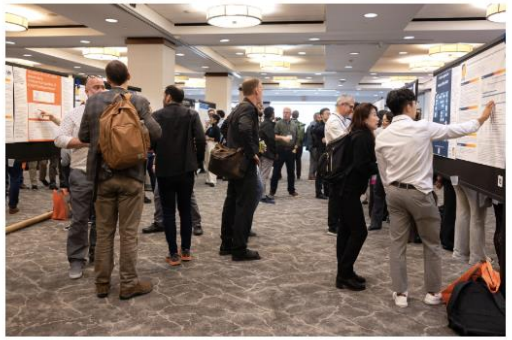
# Global Symposium: Oct. 7-9

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

- 2025 Global Symposium
- 2025 Global Collaborator Showcase

## 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.

# 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 *(Speakers: Liesbet M. Peeters & Peter Rijnbeek)* | |
| 9:10 - 9:30 | Journey of OHDSI: Where have we been and where can we go together? *(Speaker: Patrick Ryan)* | |
| 9:30 - 11:00 | Impact of Leveraging OMOP CDM for Scalable and Reliable Evidence Generation Showcased by the National Nodes *(Moderators: Renske Los & Annelies Verbiest)* | |
| 11:00 - 11:30 | *Coffee Break* | |
| 11:30 - 12:45 | Collaborator Showcase: Rapid Fire Presentations *(Moderator: TBC)* | |
| 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 *(Panel debate)* *(Confirmed speakers/moderators: Enrique Bernal-Delgado, Nick Marly, Talita Duarte-Salles, Patrick Ryan, Dipak Kalra)* | |
| 17:10 - 17:30 | Closing remarks *(Speakers: Liesbet M. Peeters & Peter Rijnbeek)* | |

## Agenda Saturday July 5, 2025

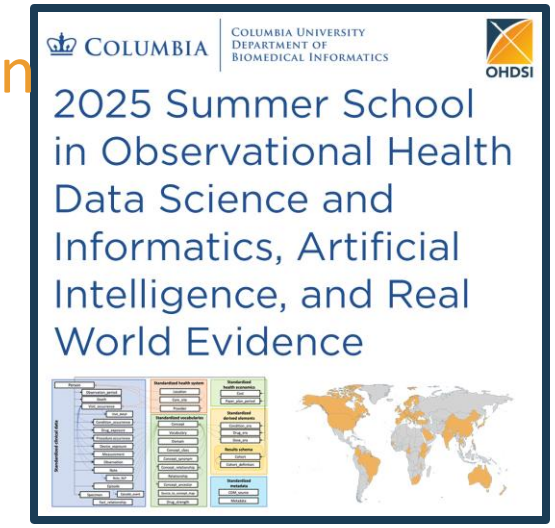
| Time | Activity | Track 1A - Newcomers | Track 1B - Newcomers | Track 2 - Advanced | Track 3 - NN/WG |
|---|---|---|---|---|---|
| 09:30 - 10:00 | *Registration + coffee* | | | | |
| 10:00 - 12:30 | Morning Session | **Introduction to OHDSI – Tutorial** Lead: Renske Los, Aniek Markus & Laura Verbeij (Erasmus MC) *Overview of OHDSI, key concepts, and an introduction to the OMOP Common Data Model* | | | **HADES hack-a-thon** Lead: Martijn Schuemie (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 | **OMOP CDM & ETL Conventions** Lead: Maxim Moinat (Erasmus MC), Sofia Bazakou & Anne van Winzum (The Hyve) | **OHDSI Standardized Vocabularies for Research – Part 1.1** Lead: Anna Ostropolets (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 | | **OHDSI Standardized Vocabularies for Research – Part 1.2** Lead: Anna Ostropolets (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 specialities)* | | | | |

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

# 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

# OHDSI on Bluesky

## OHDSI is now on Bluesky!

**You can now get updates on all community activities and see all global research through the #OHDSISocialShowcase on Bluesky.**

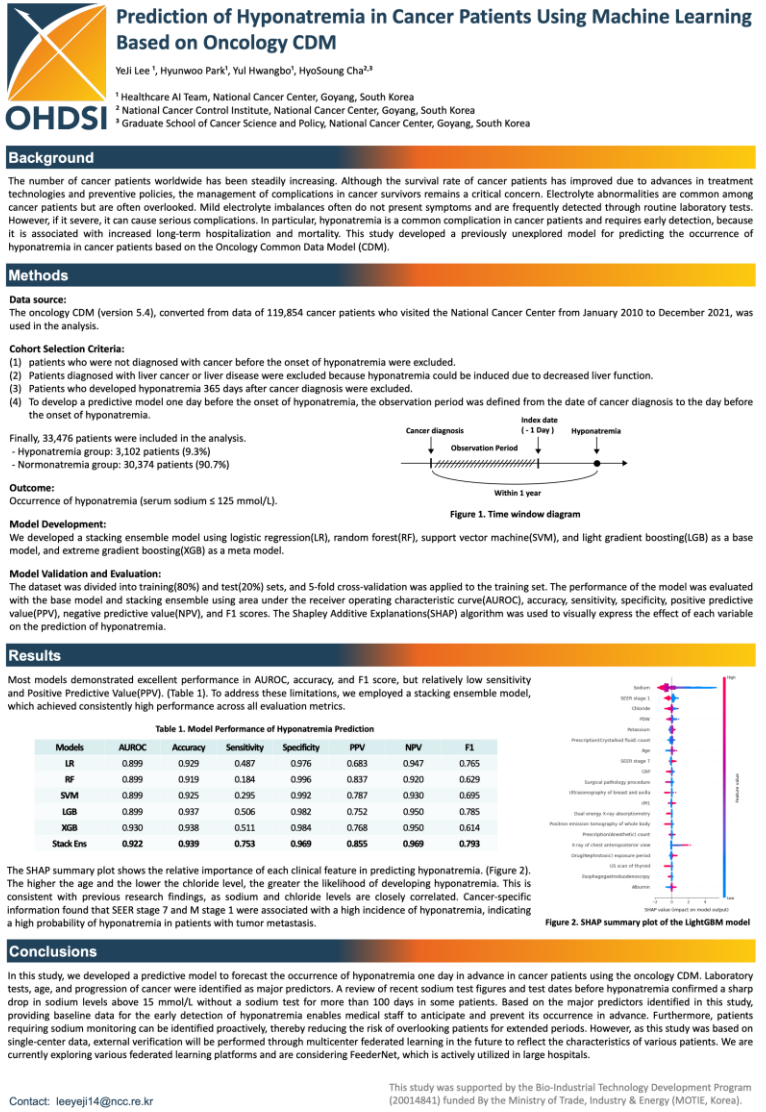
**bsky.app/profile/ohdsi.bsky.social**

# #OHDSISocialShowcase This Week

## Monday

## Prediction of Hyponatremia in Cancer Patients Using Machine Learning Based on Oncology CDM

(Yeji Lee, Hyunwoo Park, Yul Hwangbo, HyoSoung Cha)



### Prediction of Hyponatremia in Cancer Patients Using Machine Learning Based on Oncology CDM


Yeji Lee [1], Hyunwoo Park[1], Yul Hwangbo[1], HyoSoung Cha[2,3]

[1] Healthcare AI Team, National Cancer Center, Goyang, South Korea
[2] National Cancer Control Institute, National Cancer Center, Goyang, South Korea
[3] Graduate School of Cancer Science and Policy, National Cancer Center, Goyang, South Korea


#### Background

The number of cancer patients worldwide has been steadily increasing. Although the survival rate of cancer patients has improved due to advances in treatment technologies and preventive policies, the management of complications in cancer survivors remains a critical concern. Electrolyte abnormalities are common among cancer patients but are often overlooked. Mild electrolyte imbalances often do not present symptoms and are frequently detected through routine laboratory tests. However, if it severe, it can cause serious complications. In particular, hyponatremia is a common complication in cancer patients and requires early detection, because it is associated with increased long-term hospitalization and mortality. This study developed a previously unexplored model for predicting the occurrence of hyponatremia in cancer patients based on the Oncology Common Data Model (CDM).

#### Methods

**Data source:**
The oncology CDM (version 5.4), converted from data of 119,854 cancer patients who visited the National Cancer Center from January 2010 to December 2021, was used in the analysis.

**Cohort Selection Criteria:**
(1) patients who were not diagnosed with cancer before the onset of hyponatremia were excluded.
(2) Patients diagnosed with liver cancer or liver disease were excluded because hyponatremia could be induced due to decreased liver function.
(3) Patients who developed hyponatremia 365 days after cancer diagnosis were excluded.
(4) To develop a predictive model one day before the onset of hyponatremia, the observation period was defined from the date of cancer diagnosis to the day before the onset of hyponatremia.

Finally, 33,476 patients were included in the analysis.
- Hyponatremia group: 3,102 patients (9.3%)
- Normonatremia group: 30,374 patients (90.7%)




Figure 1. Time window diagram

**Outcome:**
Occurrence of hyponatremia (serum sodium ≤ 125 mmol/L).

**Model Development:**
We developed a stacking ensemble model using logistic regression(LR), random forest(RF), support vector machine(SVM), and light gradient boosting(LGB) as a base model, and extreme gradient boosting(XGB) as a meta model.

**Model Validation and Evaluation:**
The dataset was divided into training(80%) and test(20%) sets, and 5-fold cross-validation was applied to the training set. The performance of the model was evaluated with the base model and stacking ensemble using area under the receiver operating characteristic curve(AUROC), accuracy, sensitivity, specificity, positive predictive value(PPV), negative predictive value(NPV), and F1 scores. The Shapley Additive Explanations(SHAP) algorithm was used to visually express the effect of each variable on the prediction of hyponatremia.

#### Results

Most models demonstrated excellent performance in AUROC, accuracy, and F1 score, but relatively low sensitivity and Positive Predictive Value(PPV). (Table 1). To address these limitations, we employed a stacking ensemble model, which achieved consistently high performance across all evaluation metrics.

Table 1. Model Performance of Hyponatremia Prediction

| Models | AUROC | Accuracy | Sensitivity | Specificity | PPV | NPV | F1 |
|---|---|---|---|---|---|---|---|
| LR | 0.899 | 0.929 | 0.487 | 0.976 | 0.683 | 0.947 | 0.765 |
| RF | 0.899 | 0.919 | 0.184 | 0.996 | 0.837 | 0.920 | 0.629 |
| SVM | 0.899 | 0.925 | 0.295 | 0.992 | 0.787 | 0.930 | 0.695 |
| LGB | 0.899 | 0.937 | 0.506 | 0.982 | 0.752 | 0.950 | 0.785 |
| XGB | 0.930 | 0.938 | 0.511 | 0.984 | 0.768 | 0.950 | 0.614 |
| **Stack Ens** | **0.922** | **0.939** | **0.753** | **0.969** | **0.855** | **0.969** | **0.793** |



Figure 2. SHAP summary plot of the LightGBM model

The SHAP summary plot shows the relative importance of each clinical feature in predicting hyponatremia. (Figure 2). The higher the age and the lower the chloride level, the greater the likelihood of developing hyponatremia. This is consistent with previous research findings, as sodium and chloride levels are closely correlated. Cancer-specific information found that SEER stage 7 and M stage 1 were associated with a high incidence of hyponatremia, indicating a high probability of hyponatremia in patients with tumor metastasis.

#### Conclusions

In this study, we developed a predictive model to forecast the occurrence of hyponatremia one day in advance in cancer patients using the oncology CDM. Laboratory tests, age, and progression of cancer were identified as major predictors. A review of recent sodium test figures and test dates before hyponatremia confirmed a sharp drop in sodium levels above 15 mmol/L without a sodium test for more than 100 days in some patients. Based on the major predictors identified in this study, providing baseline data for the early detection of hyponatremia enables medical staff to anticipate and prevent its occurrence in advance. Furthermore, patients requiring sodium monitoring can be identified proactively, thereby reducing the risk of overlooking patients for extended periods. However, as this study was based on single-center data, external verification will be performed through multicenter federated learning in the future to reflect the characteristics of various patients. We are currently exploring various federated learning platforms and are considering FeederNet, which is actively utilized in large hospitals.


Contact: leeyeji14@ncc.re.kr



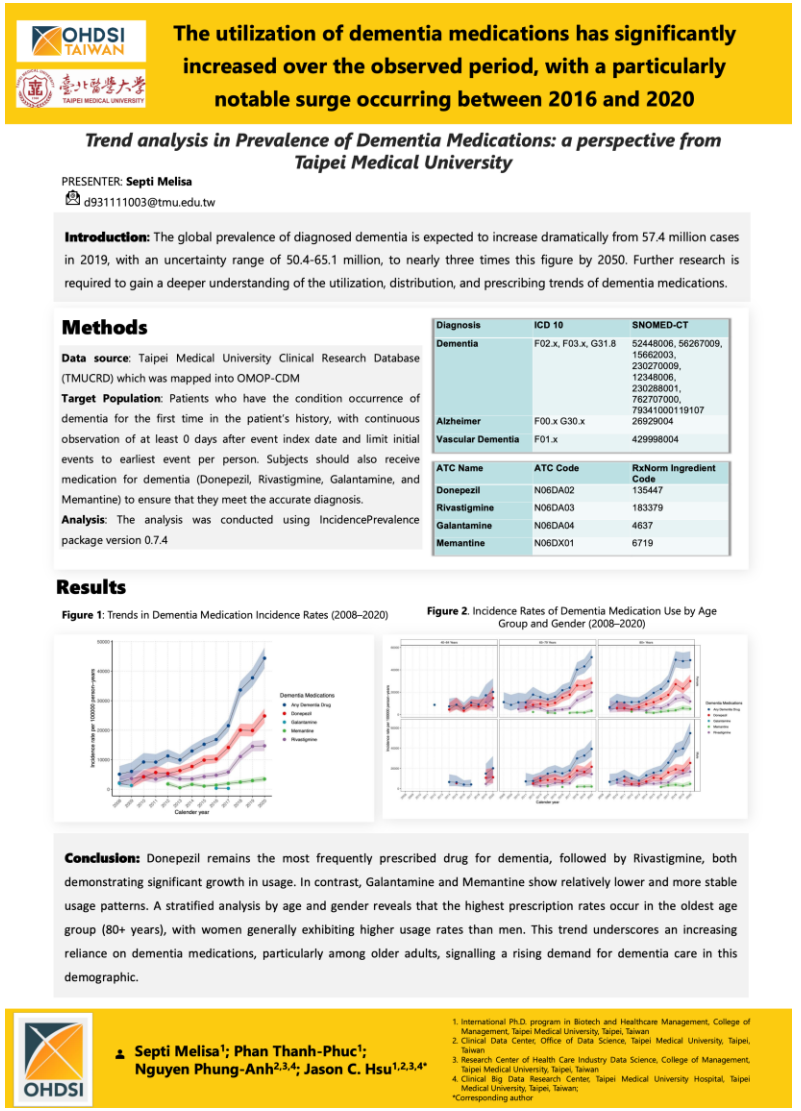
This study was supported by the Bio-Industrial Technology Development Program (20014841) funded By the Ministry of Trade, Industry & Energy (MOTIE, Korea).

# #OHDSISocialShowcase This Week

## Tuesday

## Trend analysis in Prevalence of Dementia Medications: a perspective from Taipei Medical University

(Septi Melisa, Phan Thanh-Phuc, Nguyen Phung-Anh, Jason C. Hsu)



**The utilization of dementia medications has significantly increased over the observed period, with a particularly notable surge occurring between 2016 and 2020**

*Trend analysis in Prevalence of Dementia Medications: a perspective from Taipei Medical University*

PRESENTER: **Septi Melisa**
d931111003@tmu.edu.tw

**Introduction:** The global prevalence of diagnosed dementia is expected to increase dramatically from 57.4 million cases in 2019, with an uncertainty range of 50.4-65.1 million, to nearly three times this figure by 2050. Further research is required to gain a deeper understanding of the utilization, distribution, and prescribing trends of dementia medications.

**Methods**

**Data source**: Taipei Medical University Clinical Research Database (TMUCRD) which was mapped into OMOP-CDM

**Target Population**: Patients who have the condition occurrence of dementia for the first time in the patient's history, with continuous observation of at least 0 days after event index date and limit initial events to earliest event per person. Subjects should also receive medication for dementia (Donepezil, Rivastigmine, Galantamine, and Memantine) to ensure that they meet the accurate diagnosis.

**Analysis**: The analysis was conducted using IncidencePrevalence package version 0.7.4

| Diagnosis | ICD 10 | SNOMED-CT |
|---|---|---|
| Dementia | F02.x, F03.x, G31.8 | 52448006, 56267009, 15662003, 230270009, 12348006, 230288001, 762707000, 79341000119107 |
| Alzheimer | F00.x G30.x | 26929004 |
| Vascular Dementia | F01.x | 429998004 |

| ATC Name | ATC Code | RxNorm Ingredient Code |
|---|---|---|
| Donepezil | N06DA02 | 135447 |
| Rivastigmine | N06DA03 | 183379 |
| Galantamine | N06DA04 | 4637 |
| Memantine | N06DX01 | 6719 |

**Results**

**Figure 1**: Trends in Dementia Medication Incidence Rates (2008–2020)

**Figure 2**. Incidence Rates of Dementia Medication Use by Age Group and Gender (2008–2020)

**Conclusion:** Donepezil remains the most frequently prescribed drug for dementia, followed by Rivastigmine, both demonstrating significant growth in usage. In contrast, Galantamine and Memantine show relatively lower and more stable usage patterns. A stratified analysis by age and gender reveals that the highest prescription rates occur in the oldest age group (80+ years), with women generally exhibiting higher usage rates than men. This trend underscores an increasing reliance on dementia medications, particularly among older adults, signalling a rising demand for dementia care in this demographic.

Septi Melisa[1]; Phan Thanh-Phuc[1]; Nguyen Phung-Anh[2,3,4]; Jason C. Hsu[1,2,3,4*]

1. International Ph.D. program in Biotech and Healthcare Management, College of Management, Taipei Medical University, Taipei, Taiwan
2. Clinical Data Center, Office of Data Science, Taipei Medical University, Taipei, Taiwan
3. Research Center of Health Care Industry Data Science, College of Management, Taipei Medical University, Taipei, Taiwan
4. Clinical Big Data Research Center, Taipei Medical University Hospital, Taipei Medical University, Taipei, Taiwan;
*Corresponding author

# #OHDSISocialShowcase This Week

## Wednesday

## Predicting outcome in emergency room patients with Suspected Gastrointestinal Infection using OMOP-CDM

(So Hee Lee, Byungjin Choi, Min Ho An, Junhyuk Chang, Harrin Kim, Rae Woong Park)

# Predicting outcome in emergency room patients with Suspected Gastrointestinal Infection using OMOP-CDM

So Hee Lee[1], Byungjin Choi, MD[2], Min Ho An, MD[2], Junhyuk Chang[1], Harrin Kim[1], Sujin Gan, RN[1], Rae Woong Park, MD, Ph.D [1,2]
[1]Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, Republic of Korea
[2]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Republic of Korea

## Background & Objectives

- Gastroenteritis is among the most frequently diagnosed conditions in emergency departments across the country
- Symptoms of gastroenteritis often begin mildly, with fever, diarrhea, abdominal pain, and vomiting. However, without prompt treatment, these symptoms can worsen, leading to more severe conditions that affect individuals of all ages
- In this study, we aim to predict 7-day 1) ER revisits, 2) ICU admissions, and 3) mortality, to assess the severity of gastroenteritis
- By identifying patients in high risk for disease prognosis, it enables quick triage and timely treatment, ultimately improving outcomes and reducing transmission

## Methods

***Data sources***

- Ajou University School of Medicine (AUSOM) database
  - Electronic health records (1994.01 ~ 2024.02)
  - OMOP-CDM v5.3.4

***Study population***

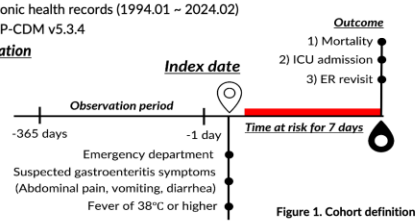



Figure 1. Cohort definition

***Model development and evaluation***

- Machine learning Algorithms:
  - Gradient boosting model (GBM) & Least absolute shrinkage and selection operator (LASSO)
- Covariates
  - Demographics, condition, drug, measurement and visit
  - Time frames : Long-term (-365 days) and Short term (-1 day) prior to the index date
  - The short-term period was chosen to capture recent conditions crucial for predicting the severity and treatment response of the patient
- Data split: Split into the train (75%) and test set (25%) in 3-fold cross validation
- Model performance
  - The area under the receiver operating curve (AUROC)
  - Youden index to determine threshold for high and low risk groups

***Survival Analysis for ICU admission***

- To assess the association between the risk of ICU admission and the incidence of 7-day mortality through the survival analyses
- Cox proportional model: for calculating the hazard ratio
- Kaplan-Meier method: for plotting the survival curve

## Results

- GBM outperformed LASSO in AUROC for predicting ER revisits, ICU admissions, and mortality (0.758 vs. 0.679, 0.964 vs. 0.947, and 0.990 vs. 0.980, respectively)

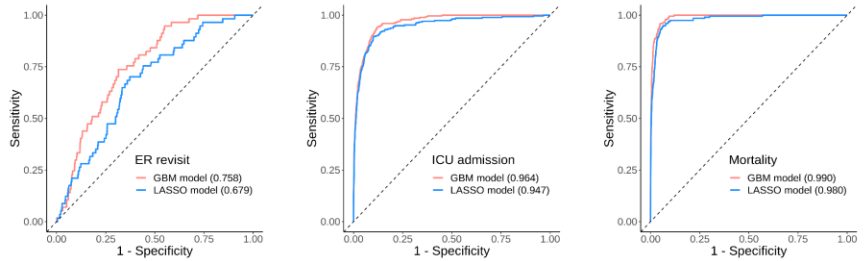



Figure 2. Model Performance for severity of gastroenteritis; (a) Prediction for ER revisit, (b) Prediction for ICU admission, (c) Prediction for mortality

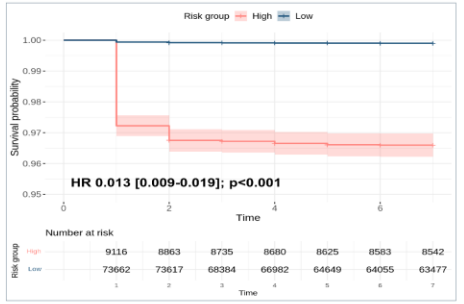



Figure 3. Kaplan-Meier survival analysis of Gradient boosting machine for ICU admission

- Based on the predicted results (Youden index: 0.005) by GBM prediction model in ICU admission, the low-risk group had a significantly lower hazard ratio for 7-day mortality (HR 0.013, 95% CI [0.009–0.019], p < 0.001)

## Conclusions

- We developed CDM-based prediction models to assess the severity of gastroenteritis outcomes, including ER revisits, ICU admissions, and mortality
- These models demonstrated moderate accuracy in predicting severity and distinguishing high-risk patients, aiding in timely treatment and improving patient outcomes

## Acknowledgements

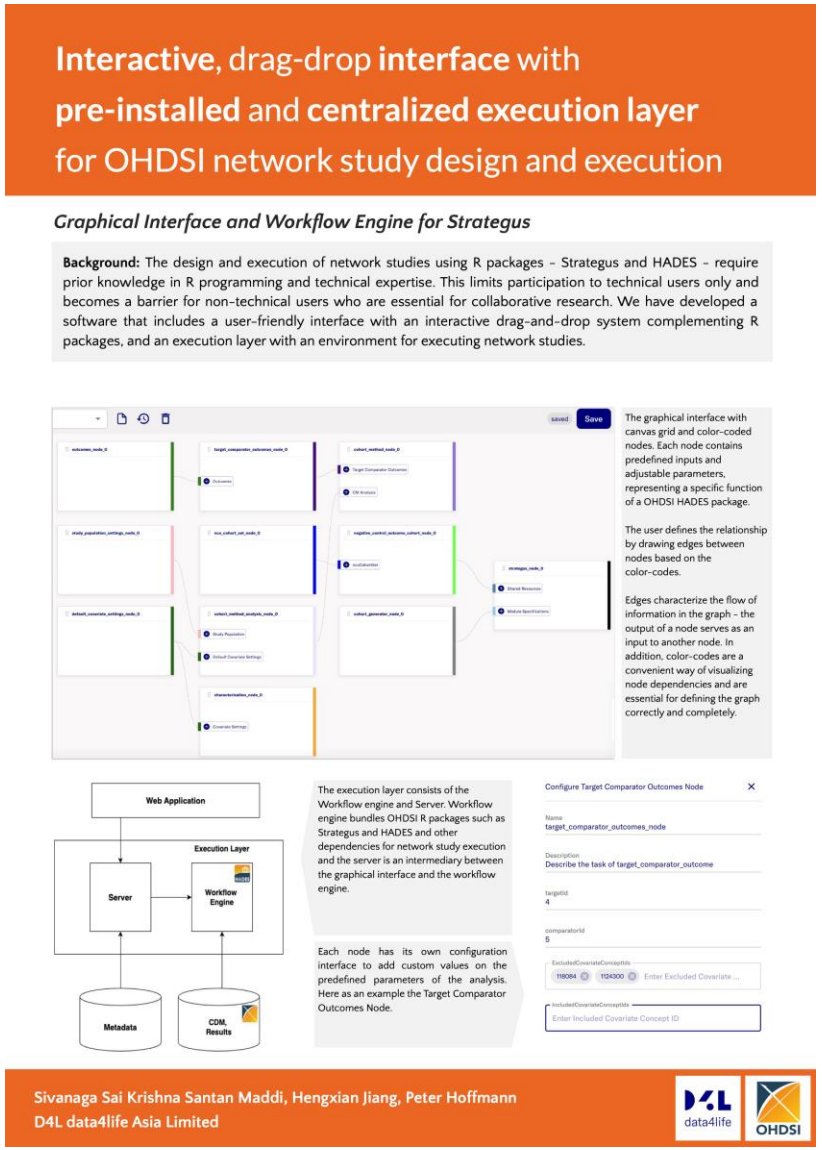
This research was funded by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HR16C0001) and this research was supported by a Government-wide R&D Fund project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

# #OHDSISocialShowcase This Week

## Thursday

## A Graphical Interface and Workflow Engine for OHDSI Network Study design and execution

(Sivanaga Sai Krishna Santan Maddi, Hengxian Jiang, Peter Hoffmann)

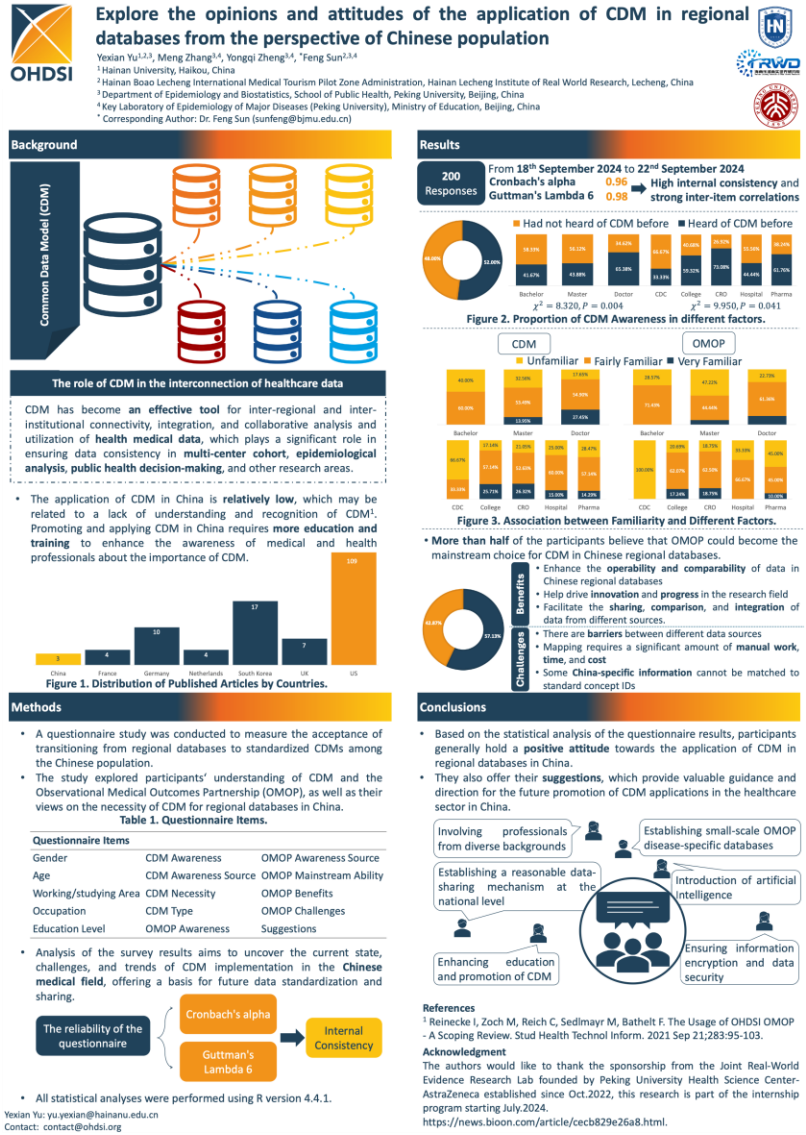


# #OHDSISocialShowcase This Week

## Friday

**Explore the opinions and attitudes of the application of common data models in regional databases from the perspective of Chinese people**

(Yexian Yu, Meng Zhang, Yongqi Zheng, Feng Sun)



Explore the opinions and attitudes of the application of CDM in regional databases from the perspective of Chinese population

Yexian Yu[1,2,3], Meng Zhang[3,4], Yongqi Zheng[3,4], *Feng Sun[2,3,4]
[1] Hainan University, Haikou, China
[2] Hainan Boao Lecheng International Medical Tourism Pilot Zone Administration, Hainan Lecheng Institute of Real World Research, Lecheng, China
[3] Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, China
[4] Key Laboratory of Epidemiology of Major Diseases (Peking University), Ministry of Education, Beijing, China
* Corresponding Author: Dr. Feng Sun (sunfeng@bjmu.edu.cn)

Background

Common Data Model (CDM)

The role of CDM in the interconnection of healthcare data

CDM has become **an effective tool** for inter-regional and inter-institutional connectivity, integration, and collaborative analysis and utilization of **health medical data**, which plays a significant role in ensuring data consistency in **multi-center cohort**, **epidemiological analysis**, **public health decision-making**, and other research areas.

- The application of CDM in China is **relatively low**, which may be related to a lack of understanding and recognition of CDM[1]. Promoting and applying CDM in China requires **more education and training** to enhance the awareness of medical and health professionals about the importance of CDM.

Figure 1. Distribution of Published Articles by Countries.

Methods

- A questionnaire study was conducted to measure the acceptance of transitioning from regional databases to standardized CDMs among the Chinese population.
- The study explored participants' understanding of CDM and the Observational Medical Outcomes Partnership (OMOP), as well as their views on the necessity of CDM for regional databases in China.

**Table 1. Questionnaire Items.**

| Questionnaire Items | | |
|---|---|---|
| Gender | CDM Awareness | OMOP Awareness Source |
| Age | CDM Awareness Source | OMOP Mainstream Ability |
| Working/studying Area | CDM Necessity | OMOP Benefits |
| Occupation | CDM Type | OMOP Challenges |
| Education Level | OMOP Awareness | Suggestions |

- Analysis of the survey results aims to uncover the current state, challenges, and trends of CDM implementation in the **Chinese medical field**, offering a basis for future data standardization and sharing.

The reliability of the questionnaire → Cronbach's alpha, Guttman's Lambda 6 → Internal Consistency

- All statistical analyses were performed using R version 4.4.1.

Yexian Yu: yu.yexian@hainanu.edu.cn
Contact: contact@ohdsi.org

Results

200 Responses

From 18th September 2024 to 22nd September 2024
Cronbach's alpha 0.96
Guttman's Lambda 6 0.98
→ High internal consistency and strong inter-item correlations

Figure 2. Proportion of CDM Awareness in different factors.

Figure 3. Association between Familiarity and Different Factors.

- **More than half** of the participants believe that OMOP could become the mainstream choice for CDM in Chinese regional databases.

Benefits
- Enhance the **operability and comparability** of data in Chinese regional databases
- Help drive **innovation** and **progress** in the research field
- Facilitate the **sharing**, **comparison**, and **integration** of data from different sources.

Challenges
- There are **barriers** between different data sources
- Mapping requires a significant amount of **manual work**, **time**, and **cost**
- Some **China-specific information** cannot be matched to standard concept IDs

Conclusions

- Based on the statistical analysis of the questionnaire results, participants generally hold a **positive attitude** towards the application of CDM in regional databases in China.
- They also offer their **suggestions**, which provide valuable guidance and direction for the future promotion of CDM applications in the healthcare sector in China.

Involving professionals from diverse backgrounds
Establishing small-scale OMOP disease-specific databases
Establishing a reasonable data-sharing mechanism at the national level
Introduction of artificial intelligence
Enhancing education and promotion of CDM
Ensuring information encryption and data security

**References**
[1] Reinecke I, Zoch M, Reich C, Sedlmayr M, Bathelt F. The Usage of OHDSI OMOP - A Scoping Review. Stud Health Technol Inform. 2021 Sep 21;283:95-103.

**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.bioon.com/article/cecb829e26a8.html.

# 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?

# June 17: ATLAS Deepdive

## Cohorts and Concept Sets


## Christopher Knoll

Director, Observational Health Data Analytics
Janssen Research and Development



## Richard Boyce

Associate Professor, Department of Biomedical Informatics
University of Pittsburgh


**Join us throughout June to help create the roadmap for ATLAS!**

# Week 3 ATLAS Survey:

## Cohorts and Concept Sets

**These weekly surveys will help us build future versions of ATLAS!**

**We are asking for input throughout our global community!**


# 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