The Journey of ATLAS

OHDSI Community Call
June 3, 2025 • 11 am ET

Upcoming Community Calls

Date	Topic
June 3	The Journey of ATLAS
June 10	ATLAS Deepdive: Data Sources and Vocabularies
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

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **Kyoung Jin Kim, Dachung Boo, Jimi Choi, Hyemin Yoon, Chai Young Jung, Seong Hee Ahn, Namki Hong, Beom-Jun Kim, Ji Seon Oh, and Seng Chan You** on the publication of **Comprehensive Evaluation of Treatment Patterns in Postmenopausal Patients with Osteoporosis without Fractures: Insights from Tertiary Care Institutions and Nationwide OMOP-CDM Data** in *Endocrinology and Metabolism*.

EnM ENDOCRINOLOGY AND METABOLISM

Original Article

Endocrinol Metab 2025 Forthcoming.
https://doi.org/10.3803/EnM.2024.2252
pISSN 2093-596X · eISSN 2093-5978

Comprehensive Evaluation of Treatment Patterns in Postmenopausal Patients with Osteoporosis without Fractures: Insights from Tertiary Care Institutions and Nationwide OMOP-CDM Data

Kyoung Jin Kim[1,*], Dachung Boo[2,3,*], Jimi Choi[1], Hyemin Yoon[4], Chai Young Jung[5], Seong Hee Ahn[6], Namki Hong[3,7], Beom-Jun Kim[8], Ji Seon Oh[4,9], Seng Chan You[2,3]

[1]Division of Endocrinology and Metabolism, Department of Internal Medicine, Korea University College of Medicine; [2]Department of Biomedical Systems Informatics, Yonsei University College of Medicine; [3]Institute for Innovation in Digital Healthcare, Yonsei University; [4]Big Data Research Center, Asan Institute of Life Science, Asan Medical Center, Seoul; [5]Biomedical Research Institute, Inha University Hospital; [6]Division of Endocrinology and Metabolism, Department of Internal Medicine, Inha University Hospital, Inha University College of Medicine, Incheon; [7]Department of Internal Medicine, Endocrine Research Institute, Severance Hospital, Yonsei University College of Medicine; [8]Division of Endocrinology & Metabolism, Department of Internal Medicine, Asan Medical Center, University of Ulsan College of Medicine; [9]Department of Information Medicine, Asan Medical Center, Seoul, Korea

Background: Osteoporosis is a global health concern. Despite emerging treatment options for this condition, limited data are available on hospital practices in South Korea. This study addresses the need for a hospital network database that reflects changes in routine clinical practice for osteoporosis in a timely manner.

Methods: We analyzed prescription patterns for anti-osteoporosis medications (AOMs) in postmenopausal women aged ≥50 years diagnosed with osteoporosis between 2012 and 2021 using data from Osteoporosis Analysis and Surveillance Initiative using Standardized data (OASIS) (four tertiary hospitals in South Korea) and a nationwide database from the Health Insurance Review and Assessment (HIRA) Service. AOMs were categorized into antiresorptive and anabolic agents, with a focus on secular changes in the use of oral bisphosphonates, denosumab, selective estrogen receptor modulators (SERMs), and anabolic agents.

Results: In the OASIS cohort, oral bisphosphonates were the most prescribed first-line AOM (49.0%), followed by denosumab (15.7%) and SERMs (18.0%). Denosumab use increased from 2% in 2016 to 40% in 2020, while oral bisphosphonate use declined from 69% in 2012 to 22% in 2021. The use of anabolic agents, including romosozumab and teriparatide, doubled to 6% after 2019. In the HIRA cohort, parenteral bisphosphonates were most common (54.3%), with significant denosumab use (17.3%).

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
Thursday	2 pm	Early-Stage Researchers
Thursday	7 pm	Dentistry
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

Latest Newsletter is Available

OHDSI

The Journey Newsletter (June 2025)

Our community continues to tackle evidence gaps in clinical guidelines through collaborative network studies. We recently shared progress and are gearing up for a deeper dive into ATLAS this June to keep the momentum going. In this edition, we highlight opportunities to share your work—including at the OHDSI Europe Symposium (July 7–9 in Hasselt, Belgium) and the Global Symposium (October 7–9 in East Brunswick, New Jersey). Plus, check out 16 new publications, meet our latest collaborator spotlight, and more.
#JoinTheJourney

Podcast: Guideline Studies, Europe Symposium Conversations, Global Showcase

In the June 2025 On The Journey podcast, Patrick Ryan and Craig Sachson discuss the continuing guideline-driven evidence studies and what next steps are to come. They look ahead to the upcoming OHDSI Europe Symposium (July 5-7), including a discussion around The European Health Data Space (EHDS), and they discuss the research possibilities for the Global Symposium (October 7-9); the deadline for #OHDSI2025 abstract submission is July 1. *(If video does not appear, please click 'View this email in your browser')*

Community Updates

Where Have We Been?

- Clinical guidelines don't just recommend treatments—they reveal where evidence is missing. In 2025, OHDSI launched an initiative to turn those gaps into research questions, driving real-world studies across the OHDSI Evidence Network. Ten study leads recently provided updates, discussed next steps and highlighted where others can collaborate.
- Four teams have won a prize competition aimed at integrating eye care and ocular imaging data into studies using large healthcare datasets in biomedical research. The selected teams participated in the $1-million challenge called Expand OHDSI Initiative for Eye Care and Ocular Imaging Challenge, hosted by the National Eye Institute (NEI), part of the National Institutes of Health.
- The OHDSI Maternal Health Fellowship was designed to empower early-stage clinical investigators to leverage emerging technologies for improved maternal and neonatal care while reducing morbidity and mortality. Members of the first cohort joined a May community call to highlight their presentations; you can see those videos at the bottom of this newsletter.
- Andromeda 1.0.0 has been released. This is a major update, where we switch the backend from SQLite to DuckDB for greatly improved performance and smaller file sizes.

Where Are We Now?

- ATLAS is a free, publicly available, web-based tool developed by the OHDSI community that facilitates the design and execution of analyses on standardized, patient-level, observational data in the OMOP CDM format. The ATLAS team will showcase the capabilities of the tool, while also facilitating a conversation to build the roadmap for the ATLAS future, throughout our June community calls.
- This is the final month to prepare your research for the 2025 Global Symposium Collaborator Showcase. Brief reports for poster presentations, oral talks, or software demonstration must be submitted by 8 PM ET on July 1, 2025.
- The #OHDSISocialShowcase is featuring research from the 2024 Asia-Pacific (APAC) Symposium this month on our LinkedIn, Twitter/X and Instagram feeds. Please make sure you are following OHDSI on our social channels to receive daily updates on the research presented by our community.

Europe Symposium (July 5-7, Belgium) Focuses on Joining the Network to Advance RWE

Real-world evidence is transforming healthcare research in Europe, especially as the upcoming European Health Data Space (EHDS) legislation paves the way for secure, interoperable, and trusted cross-border data sharing. The sixth European OHDSI Symposium – Join the Network: Advancing Real-World Evidence in Europe – will take place July 5-7, 2025, in Hasselt, Belgium; it will bring together researchers, policymakers, and industry leaders to explore the latest developments in OMOP-CDM, open-source tools, and collaborative projects that harness the power of networks to generate impactful evidence.

This symposium is a premier platform to connect, collaborate, and share insights on advancing real-world evidence in Europe. Attendees will gain valuable perspectives on National Nodes, large-scale European projects, and other groundbreaking initiatives shaping the future of health data research. The event will feature an engaging collaborator showcase, including posters, podium presentations, and interactive demonstrations of OHDSI's open-source tools, highlighting the power of community-driven research.

To further strengthen the network, the pre-symposium workshops and tutorials on July 5-6 offer hands-on learning, in-depth discussions, and one-on-one expert sessions—ideal for both newcomers and experienced OHDSI members looking to deepen their engagement.

Join us at the historic Old Prison in Hasselt for this pivotal event, and be part of the journey toward a more connected and data-driven future in healthcare research!

Register Me for the OHDSI Europe Symposium!

May Publications

Sarrat-González D, Escribà-Montagut X, Houghtaling J, González JR. dsOMOP: Bridging OMOP CDM and DataSHIELD for Secure Federated Analysis of Standardized Clinical Data. Bioinformatics. 2025 May 6:btaf286. doi: 10.1093/bioinformatics/btaf286. Epub ahead of print. PMID: 40327502.

Schwinn J, Sheikhalishahi S, Morhart M, Kaspar M, Hinske LC. A Federated Learning Model for the Prediction of Blood Transfusion in Intensive Care Units. Stud Health Technol Inform. 2025 May 15;327:227-228. doi: 10.3233/SHTI250311. PMID: 40380423.

Leis A, Mortier P, Amigo F, Bhargav M, Conde S, Ferrer M, Flygare O, Kizilaslan B, Latorre Moreno L, Mayer MA, Pérez Sola V, Portillo van Diest A, Ramírez-Anguita JM, Sanz F, Vilagut G, Alonso J, Mehlum L, Arensman E, Bjureberg J, Pastor M, Qin P. Machine Learning-Based Clinical Decision Support System for Suicide Risk Management: The PERMANENS Project. Stud Health Technol Inform. 2025 May 15;327:221-222. doi: 10.3233/SHTI250308. PMID: 40380420.

Jouned A, Düsseldorf H, Katsch F, Jafarpour M, Duftschmid G. Comparative Study of ETL Tools for Transforming Healthcare Data to the OMOP Common Data Model (CDM). Stud Health Technol Inform. 2025 May 15;327:1238-1239. doi: 10.3233/SHTI250590. PMID: 40380695.

Michel-Backofen A, Blasini R, Beck J, Marquardt K. Building a Research Infrastructure with REDCap and FHIR. Stud Health Technol Inform. 2025 May 15;327:763-764. doi: 10.3233/SHTI250457. PMID: 40380566.

Jayathissa P, Rohatsch L, Sauermann S, Hussein R. OMOP-on-FHIR: Integrating the Clinical Data Through FHIR Bundle to OMOP CDM. Stud Health Technol Inform. 2025 May 15;327:667-671. doi: 10.3233/SHTI250432. PMID: 40380541.

Sheikhalishahi S, Schwinn J, Morhart M, Kaspar M, Hinske LC. Federated Learning for Predictive Analytics in Weaning from Mechanical Ventilation. Stud Health Technol Inform. 2025 May 15;327:613-614. doi: 10.3233/SHTI250418. PMID: 40380528.

Jung H, Kim S, Yoo S. Conversion of Nursing Statements into the OMOP Common Data Model. Stud Health Technol Inform. 2025 May 15;327:611-612. doi: 10.3233/SHTI250417. PMID: 40380527.

Bachir S, Vengadeswaran A, Storf H, Kadioglu D. Metadata-Driven Approach to Generalisation of Transformations in ETL Processes. Stud Health Technol Inform. 2025 May 15;327:1438-1442. doi: 10.3233/SHTI250640. PMID: 40380743.

Spotlight: Liesbet Peeters

"What makes OHDSI unique is the shared belief that better, more equitable healthcare is possible—and that we each have a role to play in making that happen. People contribute not because they have to, but because they want to. That entrepreneurial mindset, that sense of collective ambition, is what makes the community so powerful. It's also why I feel so at home in it."

- Liesbet Peeters

ohdsi.org/spotlight-liesbet-peeters

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*				

Save The Date!

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.

#OHDSISocialShowcase This Week

Monday

Incidence, prevalence and treatment pattern of Parkinson disease from Taipei Medical University: an integration of open-software analytic tools

(Phan Thanh-Phuc, Jack Janetzki, Nguyen Phung-Anh, Nicole Pratt, Jason C. Hsu)

OHDSI TAIWAN

Incidence, prevalence and treatment pattern of Parkinson disease from Taipei Medical University: an integration of open-software analytic tools

Open-source analytical tools developed by the OHDSI community enable the determination and analysis of incidence and prevalence rates of Parkinson's disease

Background: We aim to use the open-source R package, IncidencePrevalence and TreatmentPatterns, to assess the incidence and prevalence of Parkinson disease, as well as explore treatment patterns of anti-Parkinson drugs.

Method: Patients who have the condition occurrence of Parkinson and receive medication for Parkinson (Amantadine, Benzatropine, Bromocriptine, Cabergoline, Levodopa, Entacapone, Opicapone, Pramipexole, Rasagiline, Rotigotine, Safinamide).

Results

Figure 1. Age and gender distribution

Figure 2. The incidence and Prevalence of Parkinson disease and its subtypes

Figure 3. The Treatment pattern of Parkinson medications

Conclusion: The findings serve as a valuable reference for local government initiatives in optimizing healthcare resource allocation for Parkinson's therapies. The next step of the study is to establish international collaboration and include mortality rates in the analysis. We replicated analytic code at: https://darwin-eu.github.io/IncidencePrevalence/ & https://darwin-eu.github.io/TreatmentPatterns/

Phan Thanh-Phuc[1], Jack Janetzki[2], Nguyen Phung-Anh[3,4,5], Nicole Pratt[2], Jason C. Hsu[1,3,4,5*]

1. International Ph.D. program in Biotech and Healthcare Management, College of Management, Taipei Medical University, Taipei, Taiwan
2. Quality use of Medicines and Pharmacy Research Center, University of South Australia, Australia
3. Clinical Data Center, Office of Data Science, Taipei Medical University, Taipei, Taiwan
4. Research Center of Health Care Industry Data Science, College of Management, Taipei Medical University, Taipei, Taiwan
5. Clinical Big Data Research Center, Taipei Medical University Hospital, Taipei Medical University, Taipei, Taiwan,

#OHDSISocialShowcase This Week

Tuesday

Enhancing Infectious Disease Data Integration and management through OMOP-CDM in South Korea

(Min Ho An, Seok Kim, ByungJin Choi, Sooyoung Yoo, Rae Woong Park, Ji Seon Oh)

Enhancing Infectious Disease Data Integration and management through OMOP-CDM in South Korea

Min Ho An, MD[1,2*], Seok Kim, M.P.H[3*], ByungJin Choi, MD[1,2], Sooyoung Yoo, Ph.D[3], Rae Woong Park, MD, Ph.D[1], Ji Seon Oh, MD, Ph.D[4]

[1]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Republic of Korea
[2]Department of Medical Sciences, Graduate School of Ajou University, Suwon, Republic of Korea
[3]Office of eHealth Research and Business, Seoul National University Bundang Hospital, Republic of Korea
[4]Information Medicine, Big Data Research Center, Asan Medical Center, Seoul, Republic of Korea
*These authors are equally contributed to this work

Background

- The Platform for Harmonizing and Accessing Data in Real-time on Infectious Disease Surveillance (PHAROS) was initiated to address challenges in data integration and management.
- PHAROS focuses on developing an integrated infectious disease data management system based on the OMOP-CDM in Korea, with the goal of enhancing real-time clinical information collection and improving treatment and disease management strategies.
- To support this, data encompassing microbial test results, infectious disease consultation notes, vaccination-related information, emergency room data, and legal infectious disease reports, were utilized, aimed at improving accessibility and ensuring clear representation of information.
- The codes within infectious disease consultation notes, vaccination-related information, emergency room data, and legal infectious disease reports are newly mapped and integrated as CDM records.
- Moreover, to address challenge of identifying detailed culture information, we developed new Extract Transform Load (ETL) method that suits to specifically store data drawn from specimen culture.
- While this model maintains the relationship between microbial tests and drug resistance, it captures various aspects of culture information without requiring additional data tables, thus improving the comprehensiveness and utility of information from specimen culture.

Methods

- In this study, OMOP-CDM was utilized to include infection-related clinical data. We used CDM version 5.4 without any additional columns. Infectious disease department consultation notes are integrated into the CDM's Note domain using specific concept ids, with consultation request recorded in the observation table.
- Additionally, vaccination-related reports are thoroughly documented in the drug domain, with dose information recorded in the observation table for detailed tracking.
- Primary symptom information from the National Emergency Department Information System (NEDIS) system is integrated by mapping chief complaints to SNOMED-CT and inserting them into the condition table or the observation table if no suitable mapping exists.
- We also utilized patient travel history from legal communicable disease reports. Particularly, Microbial test results were stored across three tables: specimens were stored in the specimen table, cultured microorganisms and antibiotic susceptibility results were stored in the measurement table, and the type of microorganism identified were stored in the observation table.
- These tables were designed to be linked using connection keys, facilitating the proper extraction of necessary data for various purposes.

Conclusions

- This study addresses infectious disease data integration challenges using the OMOP-CDM framework, standardizing clinical data for better accessibility and comprehensiveness. The new ETL method stores detailed culture information without extra tables, preserving key relationships between microbial tests and drug resistance. This approach may enhance research, supports rapid outbreak response, and improves disease management

Acknowledgement

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

Contact: minho.an23@gmail.com

Results

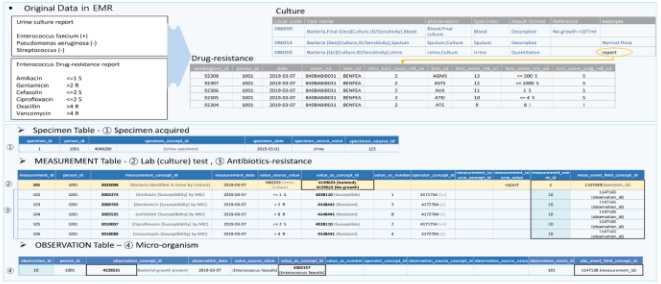


Figure 1. Culture Modeling Table Schema

Type	Number of Patients	Number of Data	Detailed Data Items
Person	239,310	239,310	De-identified ID, gender, birthdate, etc.
Visit Occurrence	239,108	5,928,625	Visit start/end time, visit type (outpatient, inpatient, etc.)
Condition Occurrence	238,859	84,261,109	Diagnosis code, diagnosis date
Drug Exposure	238,769	9,565,547	Drug code, prescription date, drug quantity, etc.
Procedure	238,707	159,246,096	Procedure code, procedure date
Measurement	227,579	58,070,197	Measurement code, result, unit (continuous, categorical, text, etc.)
Device	224,278	9,867,621	Medical device code, order date, amount
Death	222,246	14,820,106	Death date, cause of death
Observation	201,959	78,984,878	Other clinical information, observation date
Specimen	6,211	6,211	Specimen code, collection date, quantity, unit

Table 1. Converted Data Summary in Ajou University Hospital

- A total of 560 codes for infection types, testing procedures, antimicrobial sensitivity, and travel history were mapped. Additionally, the National Emergency Department Information System (NEDIS) was mapped to include 1,114 codes for major symptoms and issues.
- A total of 2,226 codes were mapped for legal infectious diseases. Furthermore, how infection-specific data such as microbial tests and antibiotic susceptibility results are stored in the CDM is illustrated in Figure 1.
- The information for specimen acquisition is recorded in the specimen table with the corresponding specimen concept ID (①).
- The results of laboratory (culture) tests are documented in the measurement table in "value_as_concept_id," indicating the existence of microorganisms by "isolated" or "no growth" (4139623) and linked to the specimen table through the "measurement_event_id" and "meas_event_field_concept_id" to trace the source (②).
- Additionally, antibiotic susceptibility data (③) is loaded into the measurement table. The differentiation from laboratory (culture) tests is achieved by using "meas_event_field_concept_id" with the related field as "observation_id".
- Lastly, the type of identified microorganism is recorded in the observation table (④), with the presence identified by observation_concept_id, and the name of the microorganism designated in "value_as_concept_id". This data is linked through the field "observation_id" matched with "measurement_event_id" in the measurement table.

#OHDSISocialShowcase This Week

Wednesday

Enabling i2b2 on OMOP CDM Cohort Data semi-automatically by using Atlas and SQLMesh

(Natpatchara Pongjirapat, Natthawut Adulyanukosol)

OHDSI

Siriraj Informatics and Data Innovation Center SiData

Enabling i2b2 on OMOP CDM Cohort Data semi-automatically by using Atlas and SQLMesh

Presenter: Natpatchara Pongjirapat Email: natpatchara.pon@mahidol.edu

Background

Informatics for Integrating Biology and the Bedside (i2b2) is an open-source software platform widely used in clinical research for querying and analyzing de-identified patient data. In December 2023, i2b2 version 1.8.0 expanded its support for the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM), a standardized format for observational health data.

Our institution utilizes i2b2 for research feasibility assessments of our Electronic Health Record (EHR) data, which is formatted in OMOP CDM. We have also developed a data pipeline that restricts data access to specific cohorts. This approach enables a semi-automated workflow for creating OMOP data cohorts and integrating them into i2b2.

Method

Our framework utilizes the OMOP Common Data Model (CDM) for initial analysis. To begin, we define and construct a cohort. This cohort consists of three columns: patient number, cohort start time, and cohort end time. Researchers can define cohorts within Atlas. Once the query is obtained, the code is executed to establish the cohort within our database.

Following cohort creation, we develop an i2b2-compatible view. We achieve this by utilizing i2b2-on-OMOP and the ENACT Ontology V4.1. To ensure adaptability for future data model changes, we implement SQLMesh to generate corresponding data models.

```
MODEL (
  name i2b2.condition_view,
  kind VIEW
);

SELECT
  visit_occurrence_id AS ENCOUNTER_NUM,
  PERSON_ID AS PATIENT_NUM,
  condition_concept_id::VARCHAR(50) AS CONCEPT_CD,
  COALESCE(provider_id::VARCHAR(50), '@'::VARCHAR) AS provider_id,
  condition_start_datetime AS START_DATE,
  condition_end_datetime AS END_DATE,
  '@'::VARCHAR(100) AS modifier_cd,
  1 AS INSTANCE_NUM,
  NULL::VARCHAR(50) AS valtype_cd,
  NULL::VARCHAR(50) AS location_cd,
  NULL::VARCHAR(255) AS observ_char,
  NULL::VARCHAR(50) AS observ_num,
  NULL::VARCHAR(50) AS unitsfrg_cd,
  NULL::VARCHAR(50) AS units_cd,
  NULL::DOUBLE PRECISION AS CONFIDENCE_NUM,
  NULL::VARCHAR(50) AS SOURCESYSTEM_CD,
  NULL::TIMESTAMP AS UPDATE_DATE,
  NULL::TIMESTAMP AS DOWNLOAD_DATE,
  NULL::TIMESTAMP AS IMPORT_DATE,
  NULL AS UNKNOWATION_BLOB,
  NULL::INT AS upload_id,
  NULL::INT AS quantity_num,
  condition_SOURCE_concept_id AS SOURCE_CONCEPT_ID,
  condition_source_value AS SOURCE_VALUE,
  'CONDITION' AS DOMAIN_ID
FROM
  omop.CONDITION_OCCURRENCE
WHERE
  PERSON_ID IN
  (SELECT PERSON_ID FROM omop.dm_cohort;
```

Figure 1: SQLmesh code example for create i2b2 data models.

SQLMesh is an open-source data transformation framework that enables version control and automated schema evolution for SQL-based data models. This facilitates the creation and updating of cohort-specific schemas, allowing for flexible handling of different cohorts and streamlining the process of integrating new data models into our i2b2-compatible views.

Additionally, we automate the creation and population of i2b2-compatible schemas using a Python script.

Result

Our methodology leverages i2b2 for user-friendly data querying and Atlas, in conjunction with Python and SQLMesh, for cohort generation. In our experiments, Atlas demonstrated its versatility as both a cohort definition tool and a SQL code generator. When combined with SQLMesh's ability to manage multiple data environments, this proved effective in filtering OMOP data into distinct patient groups.

Figure 2: Overview of our proposed framework

By isolating each cohort into separate environments, i2b2 can be configured to access cohorts on an as-needed basis. This approach ensures controlled data access while minimizing the overhead associated with data provision. The integration of Python further enhances the reproducibility of the process, reducing costs for future cohort generation.

Figure 3: Example cohort comparing query performed on our de-identified data and on subset of our data. left: the query performed on the whole de-identified data, right: the query performed on the demo cohort.

The diabetes cohort, serving as a case study for our framework, demonstrates the potential of this approach. The varying query results validate the effectiveness of our method in separating data based on the defined cohorts.

Despite its promising potential, the proposed framework is still in its early stages, and further testing is needed to evaluate its advantages.

Our experiments also revealed certain challenges, particularly concerning the adaptation of the OMOP ENACT Ontology to our data.

All code used will be released in the future in our github.
To visit scan the qr code or go to https://github.com/sidataplus

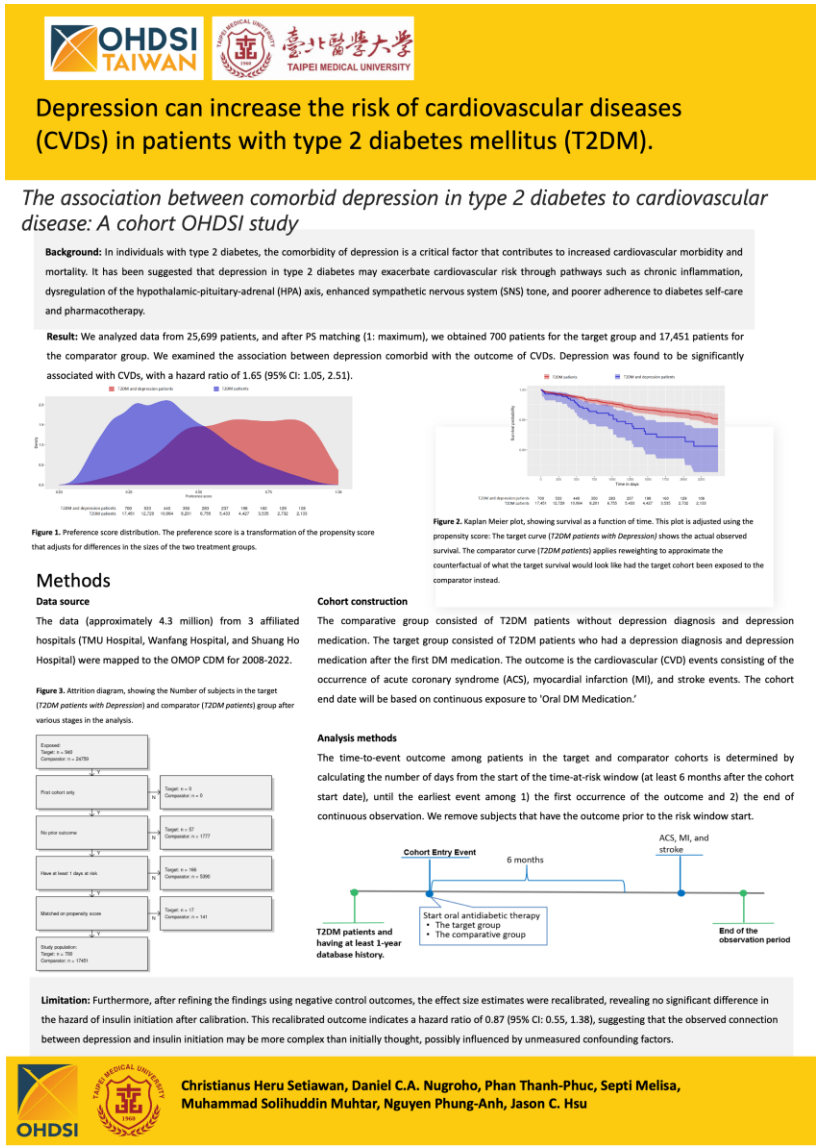
Natpatchara Pongjirapat, and Natthawut Adulyanukosol

#OHDSISocialShowcase This Week

Thursday

The association between comorbid depression in type 2 diabetes to cardiovascular disease: A cohort OHDSI study

(Christianus Heru Setiawan, Phan Thanh-Phuc, Septi Melisa, Muhammad Solihuddin Muhtar, Nguyen Phung-Anh, Jason C. Hsu)



OHDSI TAIWAN | TAIPEI MEDICAL UNIVERSITY

Depression can increase the risk of cardiovascular diseases (CVDs) in patients with type 2 diabetes mellitus (T2DM).

The association between comorbid depression in type 2 diabetes to cardiovascular disease: A cohort OHDSI study

Background: In individuals with type 2 diabetes, the comorbidity of depression is a critical factor that contributes to increased cardiovascular morbidity and mortality. It has been suggested that depression in type 2 diabetes may exacerbate cardiovascular risk through pathways such as chronic inflammation, dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis, enhanced sympathetic nervous system (SNS) tone, and poorer adherence to diabetes self-care and pharmacotherapy.

Result: We analyzed data from 25,699 patients, and after PS matching (1: maximum), we obtained 700 patients for the target group and 17,451 patients for the comparator group. We examined the association between depression comorbid with the outcome of CVDs. Depression was found to be significantly associated with CVDs, with a hazard ratio of 1.65 (95% CI: 1.05, 2.51).

Figure 1. Preference score distribution. The preference score is a transformation of the propensity score that adjusts for differences in the sizes of the two treatment groups.

Figure 2. Kaplan Meier plot, showing survival as a function of time. This plot is adjusted using the propensity score: The target curve (*T2DM patients with Depression*) shows the actual observed survival. The comparator curve (*T2DM patients*) applies reweighting to approximate the counterfactual of what the target survival would look like had the target cohort been exposed to the comparator instead.

Methods

Data source

The data (approximately 4.3 million) from 3 affiliated hospitals (TMU Hospital, Wanfang Hospital, and Shuang Ho Hospital) were mapped to the OMOP CDM for 2008-2022.

Figure 3. Attrition diagram, showing the Number of subjects in the target (*T2DM patients with Depression*) and comparator (*T2DM patients*) group after various stages in the analysis.

Cohort construction

The comparative group consisted of T2DM patients without depression diagnosis and depression medication. The target group consisted of T2DM patients who had a depression diagnosis and depression medication after the first DM medication. The outcome is the cardiovascular (CVD) events consisting of the occurrence of acute coronary syndrome (ACS), myocardial infarction (MI), and stroke events. The cohort end date will be based on continuous exposure to 'Oral DM Medication.'

Analysis methods

The time-to-event outcome among patients in the target and comparator cohorts is determined by calculating the number of days from the start of the time-at-risk window (at least 6 months after the cohort start date), until the earliest event among 1) the first occurrence of the outcome and 2) the end of continuous observation. We remove subjects that have the outcome prior to the risk window start.

Limitation: Furthermore, after refining the findings using negative control outcomes, the effect size estimates were recalibrated, revealing no significant difference in the hazard of insulin initiation after calibration. This recalibrated outcome indicates a hazard ratio of 0.87 (95% CI: 0.55, 1.38), suggesting that the observed connection between depression and insulin initiation may be more complex than initially thought, possibly influenced by unmeasured confounding factors.

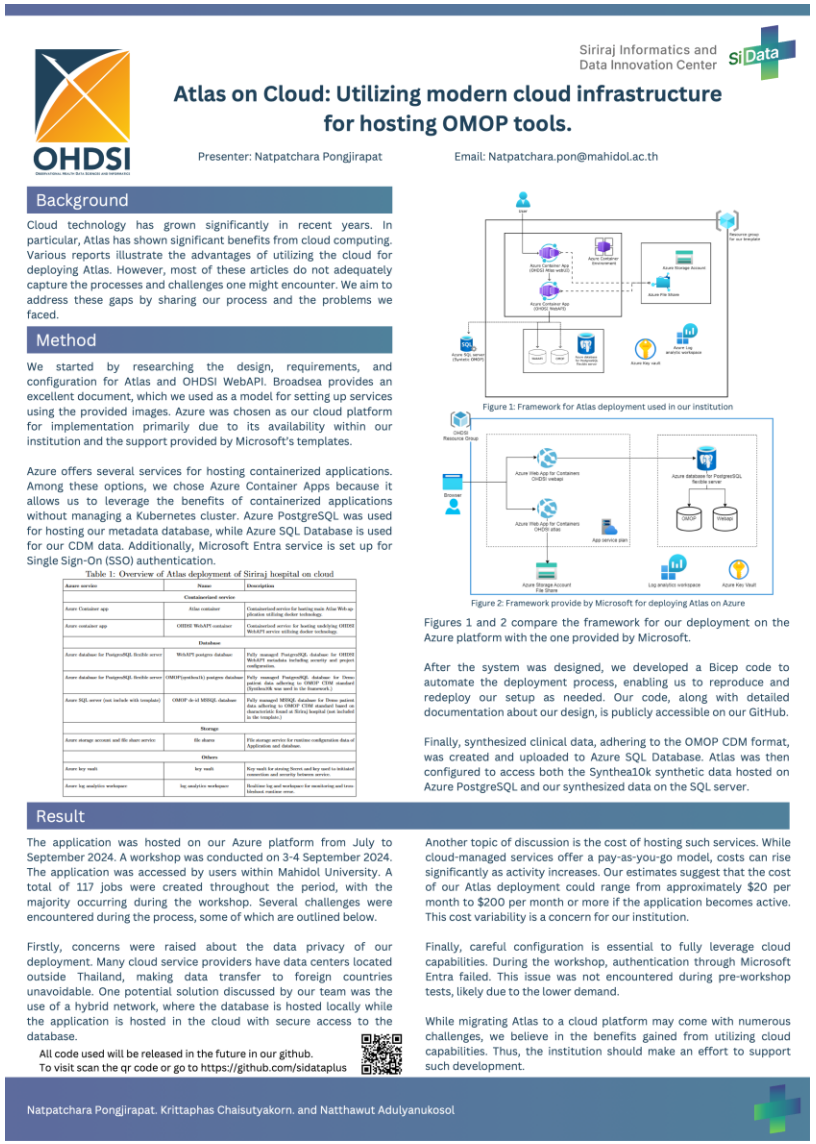
OHDSI | Christianus Heru Setiawan, Daniel C.A. Nugroho, Phan Thanh-Phuc, Septi Melisa, Muhammad Solihuddin Muhtar, Nguyen Phung-Anh, Jason C. Hsu

#OHDSISocialShowcase This Week

Friday

Atlas on Cloud: Utilizing modern cloud infrastructure for hosting OMOP tools

(Natpatchara Pongjirapat, Natthawut Adulyanukosol, and Krittaphas Chaisutyakorn)



Siriraj Informatics and Data Innovation Center

Atlas on Cloud: Utilizing modern cloud infrastructure for hosting OMOP tools.

Presenter: Natpatchara Pongjirapat Email: Natpatchara.pon@mahidol.ac.th

Background

Cloud technology has grown significantly in recent years. In particular, Atlas has shown significant benefits from cloud computing. Various reports illustrate the advantages of utilizing the cloud for deploying Atlas. However, most of these articles do not adequately capture the processes and challenges one might encounter. We aim to address these gaps by sharing our process and the problems we faced.

Method

We started by researching the design, requirements, and configuration for Atlas and OHDSI WebAPI. Broadsea provides an excellent document, which we used as a model for setting up services using the provided images. Azure was chosen as our cloud platform for implementation primarily due to its availability within our institution and the support provided by Microsoft's templates.

Azure offers several services for hosting containerized applications. Among these options, we chose Azure Container Apps because it allows us to leverage the benefits of containerized applications without managing a Kubernetes cluster. Azure PostgreSQL was used for hosting our metadata database, while Azure SQL Database is used for our CDM data. Additionally, Microsoft Entra service is set up for Single Sign-On (SSO) authentication.

Table 1: Overview of Atlas deployment of Siriraj hospital on cloud

Azure service	Name	Description
Containerized service		
Azure Container app	Atlas container	Containerized service for hosting main Atlas Web application utilizing docker technology.
Azure container app	OHDSI WebAPI container	Containerized service for hosting underlying OHDSI WebAPI service utilizing docker technology.
Database		
Azure database for PostgreSQL flexible server	WebAPI postgres database	Fully managed PostgreSQL database for OHDSI WebAPI that includes including security and project configurations.
Azure database for PostgreSQL flexible server	OMOP(synthea10k) postgres database	Fully managed PostgreSQL database for Demo patient data adhering to OMOP CDM standard (Synthea10k was used in the framework.)
Azure SQL server (not include with template)	OMOP de-id MSSQL database	Fully managed MSSQL database for Demo patient data adhering to OMOP CDM standard based on characteristic found at Siriraj hospital (not included in the template.)
Storage		
Azure storage account and file share service	file shares	File storage service for runtime configuration data of Application and database.
Others		
Azure key vault	key vault	Key vault for storing Secret and key used to initiated connection and security between service.
Azure log analytics workspace	log analytics workspace	Realtime log and workspace for monitoring and troubleshoot runtime error.

Figure 1: Framework for Atlas deployment used in our institution

Figure 2: Framework provide by Microsoft for deploying Atlas on Azure

Figures 1 and 2 compare the framework for our deployment on the Azure platform with the one provided by Microsoft.

After the system was designed, we developed a Bicep code to automate the deployment process, enabling us to reproduce and redeploy our setup as needed. Our code, along with detailed documentation about our design, is publicly accessible on our GitHub.

Finally, synthesized clinical data, adhering to the OMOP CDM format, was created and uploaded to Azure SQL Database. Atlas was then configured to access both the Synthea10k synthetic data hosted on Azure PostgreSQL and our synthesized data on the SQL server.

Result

The application was hosted on our Azure platform from July to September 2024. A workshop was conducted on 3-4 September 2024. The application was accessed by users within Mahidol University. A total of 117 jobs were created throughout the period, with the majority occurring during the workshop. Several challenges were encountered during the process, some of which are outlined below.

Firstly, concerns were raised about the data privacy of our deployment. Many cloud service providers have data centers located outside Thailand, making data transfer to foreign countries unavoidable. One potential solution discussed by our team was the use of a hybrid network, where the database is hosted locally while the application is hosted in the cloud with secure access to the database.

All code used will be released in the future in our github.
To visit scan the qr code or go to https://github.com/sidataplus

Another topic of discussion is the cost of hosting such services. While cloud-managed services offer a pay-as-you-go model, costs can rise significantly as activity increases. Our estimates suggest that the cost of our Atlas deployment could range from approximately $20 per month to $200 per month or more if the application becomes active. This cost variability is a concern for our institution.

Finally, careful configuration is essential to fully leverage cloud capabilities. During the workshop, authentication through Microsoft Entra failed. This issue was not encountered during pre-workshop tests, likely due to the lower demand.

While migrating Atlas to a cloud platform may come with numerous challenges, we believe in the benefits gained from utilizing cloud capabilities. Thus, the institution should make an effort to support such development.

Natpatchara Pongjirapat. Krittaphas Chaisutyakorn. and Natthawut Adulyanukosol

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 3: The Journey of ATLAS

Christopher Knoll

Director, Observational Health Data Analytics
Janssen Research and Development

Join us throughout June 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