Welcome OHDSI Newcomers!

OHDSI Community Call

July 28, 2026 • 11 am ET

Upcoming Community Calls

Date	Topic
July 28	OHDSI Newcomer Introductions
Aug. 4	Clinical Definitions
Aug. 11	Oncology Research in OHDSI
Aug. 18	Workgroup Spotlight: Eyecare and Vision Research, Rehabilitation
Aug. 25	Asia-Pacific Regional Updates
Sept. 1	Vocabulary Refresh, Summer 2026 Version
Sept. 8	Workgroup Spotlight: Medical Imaging, TBA
Sept. 15	Global Symposium Preview

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **Josh Lemieux, Emily Pfaff, Richard Moffitt, Jay Nakashima, Raymond Duncan, Erica Galvez, John Giannini, Julie McMurry, Anita Walden, and Melissa Haendel** on the recent publication of **From clinics to discoveries: harnessing national EHR interoperability for research** in *JAMIA Open*.

JAMIA Open, 2026, 9(4), ooag137
https://doi.org/10.1093/jamiaopen/ooag137
Brief Communications

AMIA OXFORD

From clinics to discoveries: harnessing national EHR interoperability for research

Josh Lemieux, BA[1], Emily R. Pfaff, PhD[2], Richard A. Moffitt, PhD[3], Jay Nakashima, MBA[4], Raymond Duncan, MD[5], Erica Galvez, MS[6], John P. Giannini, PhD[7], Julie A. McMurry, MPH[1], Anita Walden, MS[1], Melissa Haendel, PhD*[,1]

[1]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599, United States
[2]Department of Medicine, University of North Carolina, Chapel Hill, NC 27599, United States
[3]Department of Hematology and Medical Oncology, Department of Biomedical Informatics, Emory University, Atlanta, GA 30032, United States
[4]eHealth Exchange, Denver, CO 80210, United States
[5]Enterprise Information Services, Cedars-Sinai Medical Center, Los Angeles, CA 90066, United States
[6]Manifest MedEx, Riverside, CA 92506, United States
[7]All of Us Research Program, National Institutes of Health, Bethesda, MD 20817, United States
*Corresponding author: Melissa Haendel, PhD, Department of Genetics, University of North Carolina, UNC-Chapel Hill, 120 Mason Farm Road, 5000 D, Genetic Medicine Building CB#7264, Chapel Hill, NC 27599-7264, United States (mhaendel@unc.edu)

Abstract

Objectives: **To assess whether existing nationwide Health Information Networks (HINs) and Health Information Exchanges (HIEs) can be leveraged for electronic health record (EHR) data acquisition for research—specifically, for the NIH's *All of Us* Research Program.**

Materials and Methods: **The *All of Us* Center for Linkage and Acquisition of Data (CLAD) collaborated with eHealth Exchange, the nation's largest HIN, to transmit participant-authorized queries to an HIE and a hospital system. Returned FHIR and C-CDA records were mapped to the OMOP Common Data Model and compared with existing *All of Us* EHR data.**

Results: **Newly obtained records provided complementary information, enhancing completeness and research utility.**

Discussion: **This pilot demonstrates the feasibility of using HINs/HIEs for patient-authorized data exchange for research. Barriers remain—including inconsistent technical capacity to transact authorizations and data quality variations.**

Conclusion: **HINs and HIEs can support secure, patient-authorized EHR exchange for research, offering a pathway to more comprehensive, participant-driven biomedical discovery.**

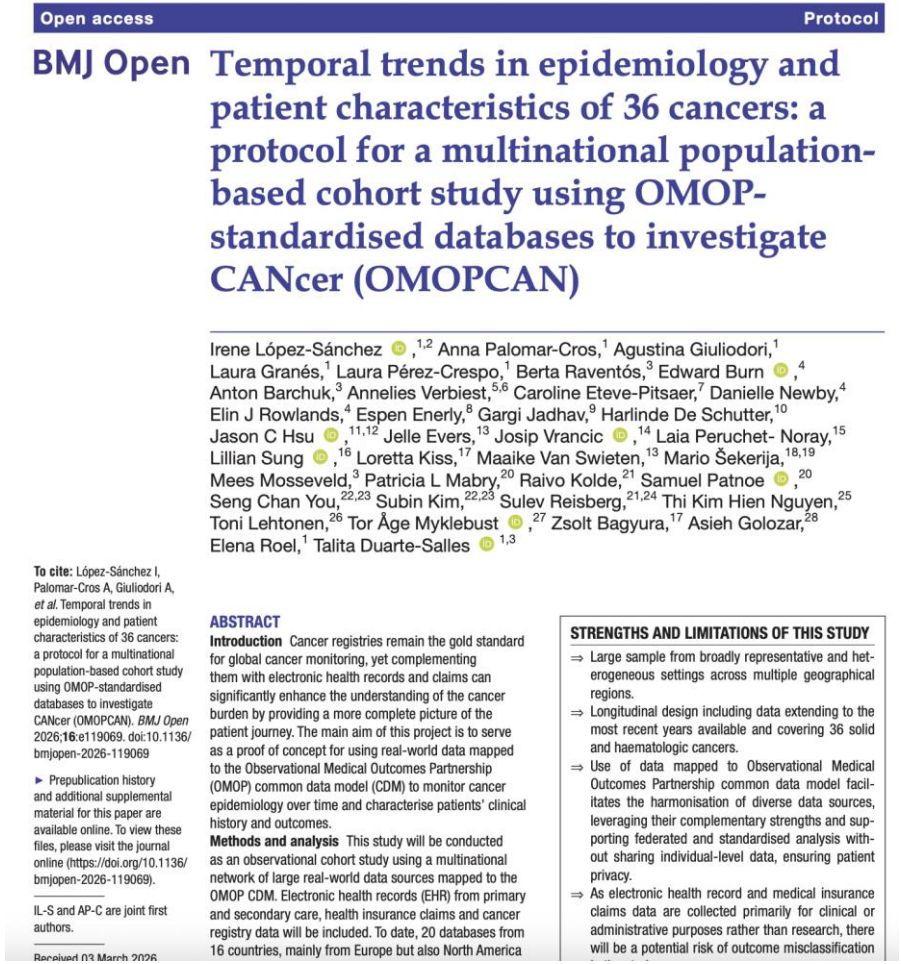
Key words: real-world data; health information exchange; EHR data; interoperability

Lay Summary

This pilot study explores how to best fill gaps in people's health records for use in research. Patients are often seen at many different clinics and hospitals, making it hard for researchers to see a complete picture of someone's health. This is a challenge for large programs like the National Institutes of Health's *All of Us* Research Program, which aims to collect comprehensive health data from over one million volunteer participants.
We evaluated whether nationwide health data-sharing networks used for clinical care data exchange could be adapted for research. We sent secure requests for medical records through one of the country's largest networks. The returned data were then compared with existing records supplied originally by the study's recruitment sites.
Our results show not only that the approach works, but also that we can acquire important missing information, including for participants who previously had no medical records in the system.
However, challenges remain, such as differences in data quality and technical limitations in sharing information. Overall, the study demonstrates that using existing health data networks, with patient consent, could significantly improve the quality and completeness of data used in medical research.

OHDSI Shoutouts!

Congratulations to the team of Irene López-Sánchez, Anna Palomar-Cros, Agustina Giuliodori, Laura Granés, Laura Pérez-Crespo, Berta Raventós, Edward Burn, Anton Barchuk, Annelies Verbiest, Caroline Eteve-Pitsaer, Danielle Newby, Elin Rowlands, Espen Enerly, Gargi Jadhav, Harlinde De Schutter, Jason Hsu, Jelle Evers, Josip Vrancic, Laia Peruchet-Noray, Lillian Sung, Loretta Kiss, Maaike Van Swieten, Mario Šekerija, Mees Mosseveld, Patricia L Mabry, Raivo Kolde, Samuel Patnoe, Seng Chan You, Subin Kim, Sulev Reisberg, Thi Kim Hien Nguyen, Toni Lehtonen, Tor Åge Myklebust, Zsolt Bagyura, Asieh Golozar, Elena Roel, and Talita Duarte-Salles on the recent publication of **Temporal trends in epidemiology and patient characteristics of 36 cancers: a protocol for a multinational population-based cohort study using OMOP-standardised databases to investigate CANcer (OMOPCAN)** in *BMJ Open*.



Open access Protocol

BMJ Open

Temporal trends in epidemiology and patient characteristics of 36 cancers: a protocol for a multinational population-based cohort study using OMOP-standardised databases to investigate CANcer (OMOPCAN)

Irene López-Sánchez [1,2] Anna Palomar-Cros,[1] Agustina Giuliodori,[1] Laura Granés,[1] Laura Pérez-Crespo,[1] Berta Raventós,[3] Edward Burn [4] Anton Barchuk,[3] Annelies Verbiest,[5,6] Caroline Eteve-Pitsaer,[7] Danielle Newby,[4] Elin J Rowlands,[4] Espen Enerly,[8] Gargi Jadhav,[9] Harlinde De Schutter,[10] Jason C Hsu [11,12] Jelle Evers,[13] Josip Vrancic [14] Laia Peruchet- Noray,[15] Lillian Sung [16] Loretta Kiss,[17] Maaike Van Swieten,[13] Mario Šekerija,[18,19] Mees Mosseveld,[3] Patricia L Mabry,[20] Raivo Kolde,[21] Samuel Patnoe [20] Seng Chan You,[22,23] Subin Kim,[22,23] Sulev Reisberg,[21,24] Thi Kim Hien Nguyen,[25] Toni Lehtonen,[26] Tor Åge Myklebust [27] Zsolt Bagyura,[17] Asieh Golozar,[28] Elena Roel,[1] Talita Duarte-Salles [1,3]

To cite: López-Sánchez I, Palomar-Cros A, Giuliodori A, *et al*. Temporal trends in epidemiology and patient characteristics of 36 cancers: a protocol for a multinational population-based cohort study using OMOP-standardised databases to investigate CANcer (OMOPCAN). *BMJ Open* 2026;**16**:e119069. doi:10.1136/bmjopen-2026-119069

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2026-119069).

IL-S and AP-C are joint first authors.

Received 03 March 2026

ABSTRACT

Introduction Cancer registries remain the gold standard for global cancer monitoring, yet complementing them with electronic health records and claims can significantly enhance the understanding of the cancer burden by providing a more complete picture of the patient journey. The main aim of this project is to serve as a proof of concept for using real-world data mapped to the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) to monitor cancer epidemiology over time and characterise patients' clinical history and outcomes.

Methods and analysis This study will be conducted as an observational cohort study using a multinational network of large real-world data sources mapped to the OMOP CDM. Electronic health records (EHR) from primary and secondary care, health insurance claims and cancer registry data will be included. To date, 20 databases from 16 countries, mainly from Europe but also North America

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Large sample from broadly representative and heterogeneous settings across multiple geographical regions.

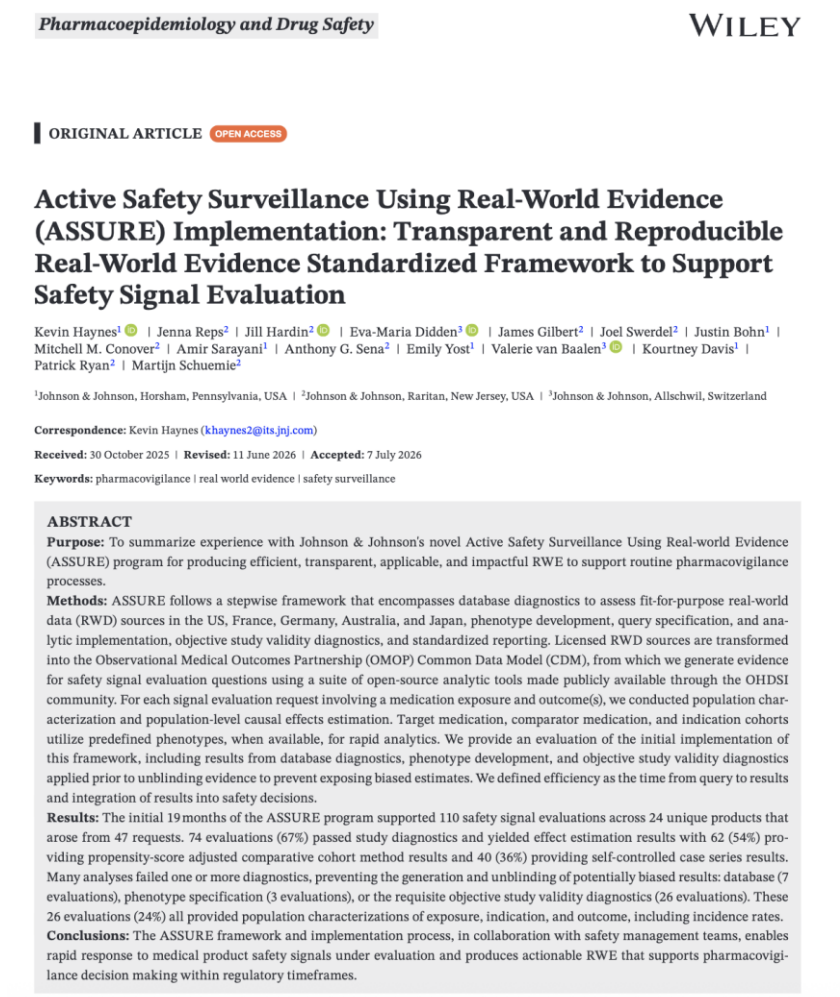
⇒ Longitudinal design including data extending to the most recent years available and covering 36 solid and haematologic cancers.

⇒ Use of data mapped to Observational Medical Outcomes Partnership common data model facilitates the harmonisation of diverse data sources, leveraging their complementary strengths and supporting federated and standardised analysis without sharing individual-level data, ensuring patient privacy.

⇒ As electronic health record and medical insurance claims data are collected primarily for clinical or administrative purposes rather than research, there will be a potential risk of outcome misclassification

OHDSI Shoutouts!

Congratulations to the team of **Kevin Haynes, Jenna Reps, Jill Hardin, Eva-Maria Didden, James Gilbert, Joel Swerdel, Justin Bohn, Mitchell Conover, Amir Sarayani, Anthony Sena, Emily Yost, Valerie van Baalen, Kourtney Davis, Patrick Ryan, and Martijn Schuemie** on the recent publication of **Active Safety Surveillance Using Real-World Evidence (ASSURE) Implementation: Transparent and Reproducible Real-World Evidence Standardized Framework to Support Safety Signal Evaluation** in *Pharmacoepidemiology & Drug Safety*.



Pharmacoepidemiology and Drug Safety WILEY

ORIGINAL ARTICLE OPEN ACCESS

Active Safety Surveillance Using Real-World Evidence (ASSURE) Implementation: Transparent and Reproducible Real-World Evidence Standardized Framework to Support Safety Signal Evaluation

Kevin Haynes[1] | Jenna Reps[2] | Jill Hardin[2] | Eva-Maria Didden[3] | James Gilbert[2] | Joel Swerdel[2] | Justin Bohn[1] | Mitchell M. Conover[2] | Amir Sarayani[1] | Anthony G. Sena[2] | Emily Yost[1] | Valerie van Baalen[3] | Kourtney Davis[1] | Patrick Ryan[2] | Martijn Schuemie[2]

[1]Johnson & Johnson, Horsham, Pennsylvania, USA | [2]Johnson & Johnson, Raritan, New Jersey, USA | [3]Johnson & Johnson, Allschwil, Switzerland

Correspondence: Kevin Haynes (khaynes2@its.jnj.com)

Received: 30 October 2025 | **Revised:** 11 June 2026 | **Accepted:** 7 July 2026

Keywords: pharmacovigilance | real world evidence | safety surveillance

ABSTRACT

Purpose: To summarize experience with Johnson & Johnson's novel Active Safety Surveillance Using Real-world Evidence (ASSURE) program for producing efficient, transparent, applicable, and impactful RWE to support routine pharmacovigilance processes.

Methods: ASSURE follows a stepwise framework that encompasses database diagnostics to assess fit-for-purpose real-world data (RWD) sources in the US, France, Germany, Australia, and Japan, phenotype development, query specification, and analytic implementation, objective study validity diagnostics, and standardized reporting. Licensed RWD sources are transformed into the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM), from which we generate evidence for safety signal evaluation questions using a suite of open-source analytic tools made publicly available through the OHDSI community. For each signal evaluation request involving a medication exposure and outcome(s), we conducted population characterization and population-level causal effects estimation. Target medication, comparator medication, and indication cohorts utilize predefined phenotypes, when available, for rapid analytics. We provide an evaluation of the initial implementation of this framework, including results from database diagnostics, phenotype development, and objective study validity diagnostics applied prior to unblinding evidence to prevent exposing biased estimates. We defined efficiency as the time from query to results and integration of results into safety decisions.

Results: The initial 19 months of the ASSURE program supported 110 safety signal evaluations across 24 unique products that arose from 47 requests. 74 evaluations (67%) passed study diagnostics and yielded effect estimation results with 62 (54%) providing propensity-score adjusted comparative cohort method results and 40 (36%) providing self-controlled case series results. Many analyses failed one or more diagnostics, preventing the generation and unblinding of potentially biased results: database (7 evaluations), phenotype specification (3 evaluations), or the requisite objective study validity diagnostics (26 evaluations). These 26 evaluations (24%) all provided population characterizations of exposure, indication, and outcome, including incidence rates.

Conclusions: The ASSURE framework and implementation process, in collaboration with safety management teams, enables rapid response to medical product safety signals under evaluation and produces actionable RWE that supports pharmacovigilance decision making within regulatory timeframes.

OHDSI Shoutouts!

Congratulations to the team of **Obinwa Ozonze, Ogechukwu Okonor, Uzondu Dike, and Taiwo Adedeji** on the recent publication of **Enabling interoperability across disparate health data sources using Large Language Models** in *PLOS Digital Health*.

PLOS Digital Health

RESEARCH ARTICLE

Enabling interoperability across disparate health data sources using Large Language Models

Obinwa Ozonze[1], Ogechukwu Okonor[2], Uzondu Dike[3], Taiwo Adedeji[1]*

1 School of Computing, University of Portsmouth, Portsmouth, Hampshire, United Kingdom, **2** Department of Computing & Emerging Technologies, Ravensbourne University London, London, United Kingdom, **3** Insight and Analytics, Emcor Group, Eastleigh, Hampshire, United Kingdom

* taiwo.adedeji@port.ac.uk

Abstract

The heterogeneity of health data systems remains a major barrier to interoperability and data integration, and a challenge for clinical decision support and AI-driven research. The coexistence of multiple Common Data Models (CDMs) complicates integration, requiring complex manual or rule-based schema mapping, which is labor-intensive, error-prone, and difficult to scale. We evaluated three leading Large Language Models (LLMs): ChatGPT, Gemini, and DeepSeek, for their ability to match data elements from four widely used CDMs (PCORnet, OMOP, Sentinel, and i2b2) to FHIR R4 via web-based chat interfaces. Reference mappings from the HL7 Common Data Models Harmonization Implementation Guide were used as the benchmark. Standardized prompts were submitted across 36 total runs (3 LLMs × 4 schemas × 3 repetitions), and accuracy was measured using F1-scores with 95% confidence intervals, alongside strict (AND) and relaxed (OR) match criteria. DeepSeek obtained the highest numerical mean F1-scores (0.78-0.99) with strong consistency (AND Match = 0.87). Gemini showed good coverage (OR Match = 0.96) but greater variability (AND Match = 0.64), while ChatGPT achieved moderate accuracy with lower consistency (AND Match = 0.51). A two-way ANOVA indicated no statistically significant differences by LLM type ($p = 0.138$), CDM schema ($p = 0.194$), or their interaction ($p = 0.451$). Performance varied by domain, with Demographics and Condition mapped most accurately, while Observation showed the highest error rates (44% of total errors). Out of 2,583 total matchings, 314 errors were observed and categorized into four failure patterns: semantic confusion (42%), structural misalignment (28%), ambiguous terminology (18%), and terminology variance (12%). This study demonstrates the feasibility of using LLMs to automate health data integration as a complementary approach to traditional ontology-based methods, particularly for well-structured, context-rich CDMs. However, safe deployment requires human-in-the-loop validation, governance, prompt optimization, metadata integration, and validation to ensure clinical and research reliability.

Check for updates

OPEN ACCESS

Citation: Ozonze O, Okonor O, Dike U, Adedeji T (2026) Enabling interoperability across disparate health data sources using Large Language Models. PLOS Digit Health 5(7): e0001511. https://doi.org/10.1371/journal.pdig.0001511

Editor: Hadi Amiri, UMass Lowell: University of Massachusetts Lowell, UNITED STATES OF AMERICA

Received: July 1, 2025

Accepted: June 5, 2026

Published: July 23, 2026

Data availability statement: All research materials are publicly available in a GitHub repository (https://github.com/taiwoadedeji/Automate_Health_Data_Integration) and are accompanied by a Zenodo DOI (https://doi.org/10.5281/zenodo.18407899) upon

OHDSI Shoutouts!

Congratulations to the team of **Dong Han Yu, Sun Mi Lim, Hyun Chul Shin, Moosung Kim, Seng Chan You, Rae Woong Park, and Chungsoo Kim** on the recent publication of **Data Resource Profile: Health Insurance Review and Assessment Service Korean nationwide claims OMOP CDM (2015–24) database, HIRA K-OMOP** in the *International Journal of Epidemiology*.

International Journal of Epidemiology, 2026, **55(4)**, dyag131
https://doi.org/10.1093/ije/dyag131
Data Resource Profile

IEA International Epidemiological Association — OXFORD

Data Resource Profile: Health Insurance Review and Assessment Service Korean nationwide claims OMOP CDM (2015–24) database, HIRA K-OMOP

Dong Han Yu[1], Sun Mi Lim[1], Hyun Chul Shin[1], Moosung Kim[1], Seng Chan You[2], Rae Woong Park[3], Chungsoo Kim[3,4,*]

[1]Big Data Department, Health Insurance Review and Assessment Service, Wonju, 26465, Republic of Korea
[2]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, 03722, Republic of Korea
[3]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, 16499, Republic of Korea
[4]Center for Outcomes Research and Evaluation, Yale New Haven Hospital, New Haven, Connecticut, 06510, United States
*Corresponding author: Department of Biomedical Informatics, Ajou University School of Medicine, 164 World cup-ro, Hongjae Bd, 5th fl, Suwon 16499, Republic of Korea. E-mail: ted9219@ajou.ac.kr

Keywords claims, OMOP, common data model, data standardization, observational study

Key Features

- HIRA K-OMOP is a nationwide Observational Medical Outcome Partnership Common Data Model resource derived from South Korea's National Health Insurance claims.
- It covers the entire insured population from 2015 to 2024 (56 416 773 persons), enabling population-scale health-services and pharmacoepidemiology research.
- The database captures diagnoses, procedures, costs, and medication data, with prescribing and pharmacy dispensing recorded as separate events to improve drug-exposure validity.
- Access is provided via a privacy-preserving secure research environment under Health Insurance Review and Assessment Service (HIRA) governance and requires approval from the Institutional Review Board and authorization from HIRA.

OHDSI Shoutouts!

Congratulations to the team of **Haeun Lee, Benjamin Martin, Andreea Creanga, Evan Minty, Asieh Golozar, Cindy Cai, Cynthia Sung, Sean O'Reilly, Khyzer Aziz, and Paul Nagy** on the recent publication of **Assessment of real-world evidence research competencies in federated research networks: a maternal health fellowship program evaluation** in *JAMIA Open*.

JAMIA Open, 2026, 9(4), ooag128
https://doi.org/10.1093/jamiaopen/ooag128
Research and Applications

AMIA OXFORD

Assessment of real-world evidence research competencies in federated research networks: a maternal health fellowship program evaluation

Haeun Lee, MS*[,1], Benjamin Martin, PhD[1], Andreea A. Creanga, MD, PhD[2], Evan Minty, MD, MSc[3], Asieh Golozar, MD, PhD[1,4], Cindy X. Cai, MD[1,5], Cynthia Sung, PhD[6], Sean O'Rilley, BS[1], Khyzer Aziz, MD[1,7], Paul Nagy, PhD[1]

[1]Biomedical Informatics and Data Science, Johns Hopkins School of Medicine, Johns Hopkins University, Baltimore, MD, 21205, United States
[2]Department of Gynecology and Obstetrics, The Johns Hopkins University School of Medicine, Baltimore, MD, 21287, United States
[3]Faculty of Medicine, O'Brien Institute for Public Health, University of Calgary, Calgary, Alberta, 2500, Canada
[4]Nemesis Health, New York, NY, 10009, United States
[5]Wilmer Eye Institute, Johns Hopkins School of Medicine, Johns Hopkins University, Baltimore, MD, United States
[6]Center of Regulatory Excellence, Duke-NUS Medical School, Singapore
[7]Department of Neonatology, The Johns Hopkins University School of Medicine, Baltimore, MD, United States
*Corresponding author: Haeun Lee, MS, Biomedical Informatics and Data Science, Johns Hopkins School of Medicine, Johns Hopkins University, 2024 E. Monument Street, Baltimore, MD, 21205, United States (hlee292@jh.edu)

Abstract

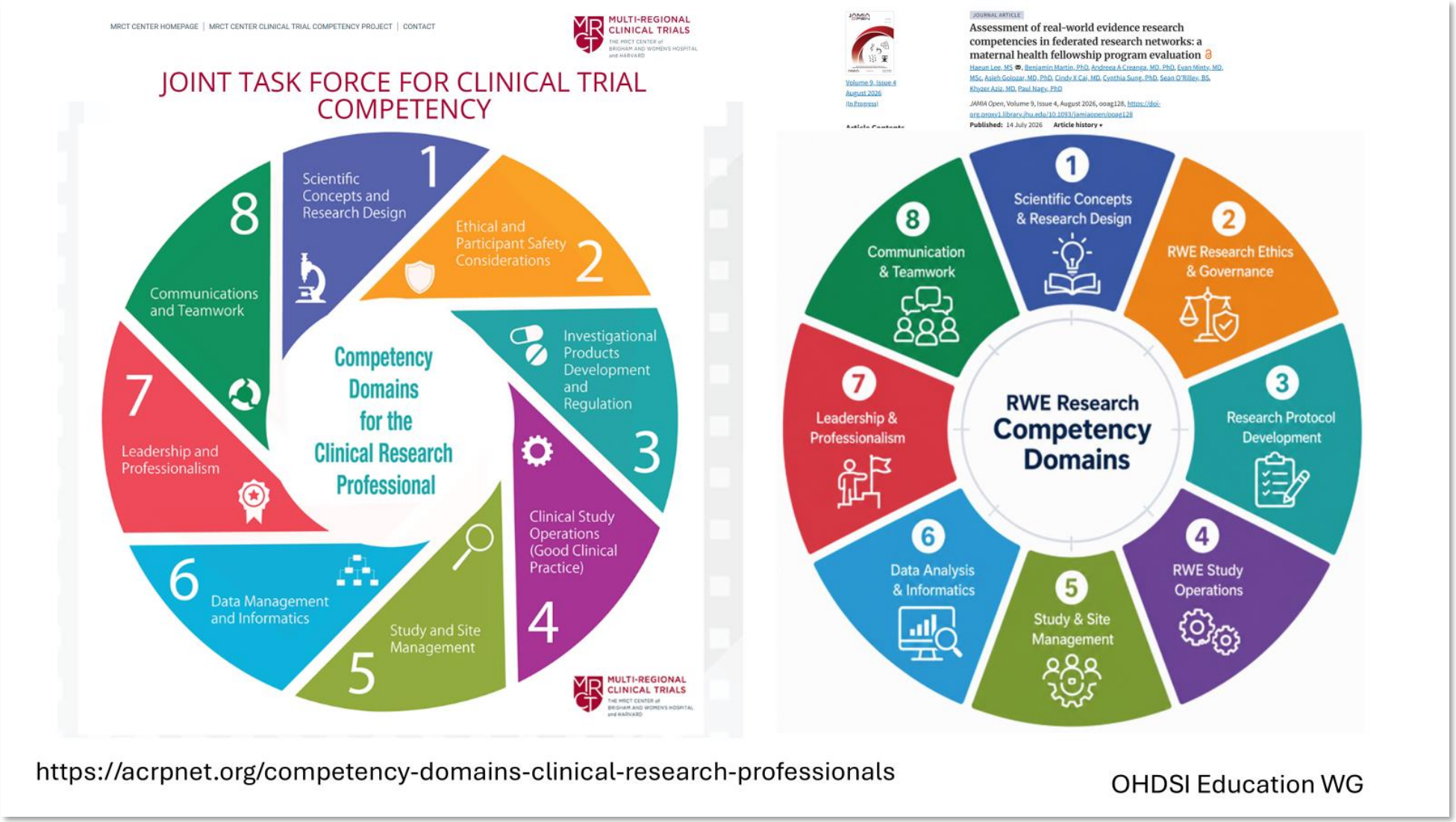
Background and Significance: **Generating real-world evidence (RWE) approaches across federated evidence networks based on real-world data requires learning new methodological skills. However, standardized frameworks for defining and assessing competencies for large-scale RWE research remain limited. This study aims to develop an assessment framework for RWE studies and evaluate the competencies of RWE research professionals.**

Methods: **We adapted the Joint Task Force (JTF) Core Competency Framework for RWE research using the Observational Medical Outcomes Partnership (OMOP) Common Data Model. Through iterative expert review, 4 new domains were created, including Ethics and Governance, Protocol Development, Study Operations, and Data Analysis and Informatics, and 4 original JTF domains were adapted. Participants completed pre- and post-training self-assessments on a 10-point scale. Paired *t*-tests with effect sizes were used to evaluate changes in competency scores.**

Results: **Seventeen participants from academic, data, and clinical roles completed the assessment. Mean self-reported competency scores increased by 2.5 points across the 8 competency domains, with statistically significant changes observed in all domains. The largest gain was observed in Leadership and Professionalism (mean increase 3.66; 95% CI, 2.69-4.62).**

Discussion: **The RWE research competency assessment tool captured changes in self-reported skill proficiency, perceived readiness to conduct RWE studies, and methodological knowledge gaps. These findings suggest the potential to identify competency gaps, tailor educational interventions, and inform quality benchmarks for RWE research training programs.**

OHDSI Shoutouts!

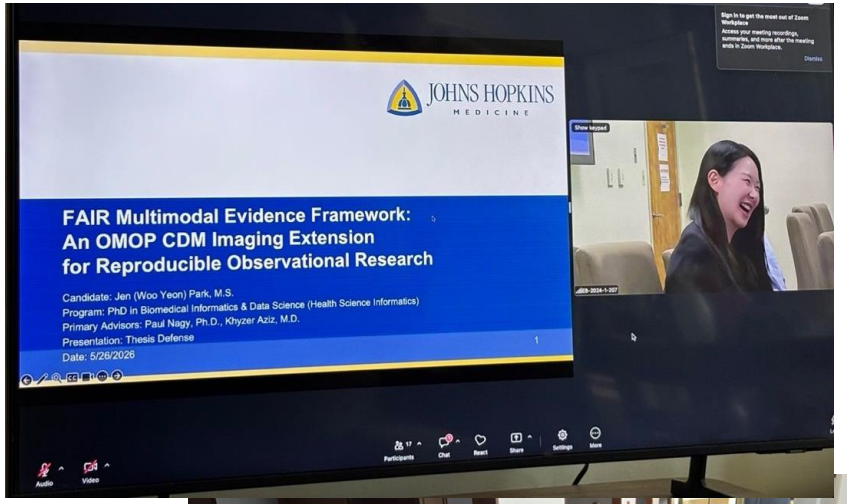


https://acrpnet.org/competency-domains-clinical-research-professionals

OHDSI Education WG

OHDSI Shoutouts!

Congratulations to **Jen Wooyean Park** on completing her Ph.D. recently at Johns Hopkins University.



Her dissertation was entitled *FAIR Multimodal Evidence Framework: An OMOP Common Data Model Imaging Extension for Reproducible Observational Research.*

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

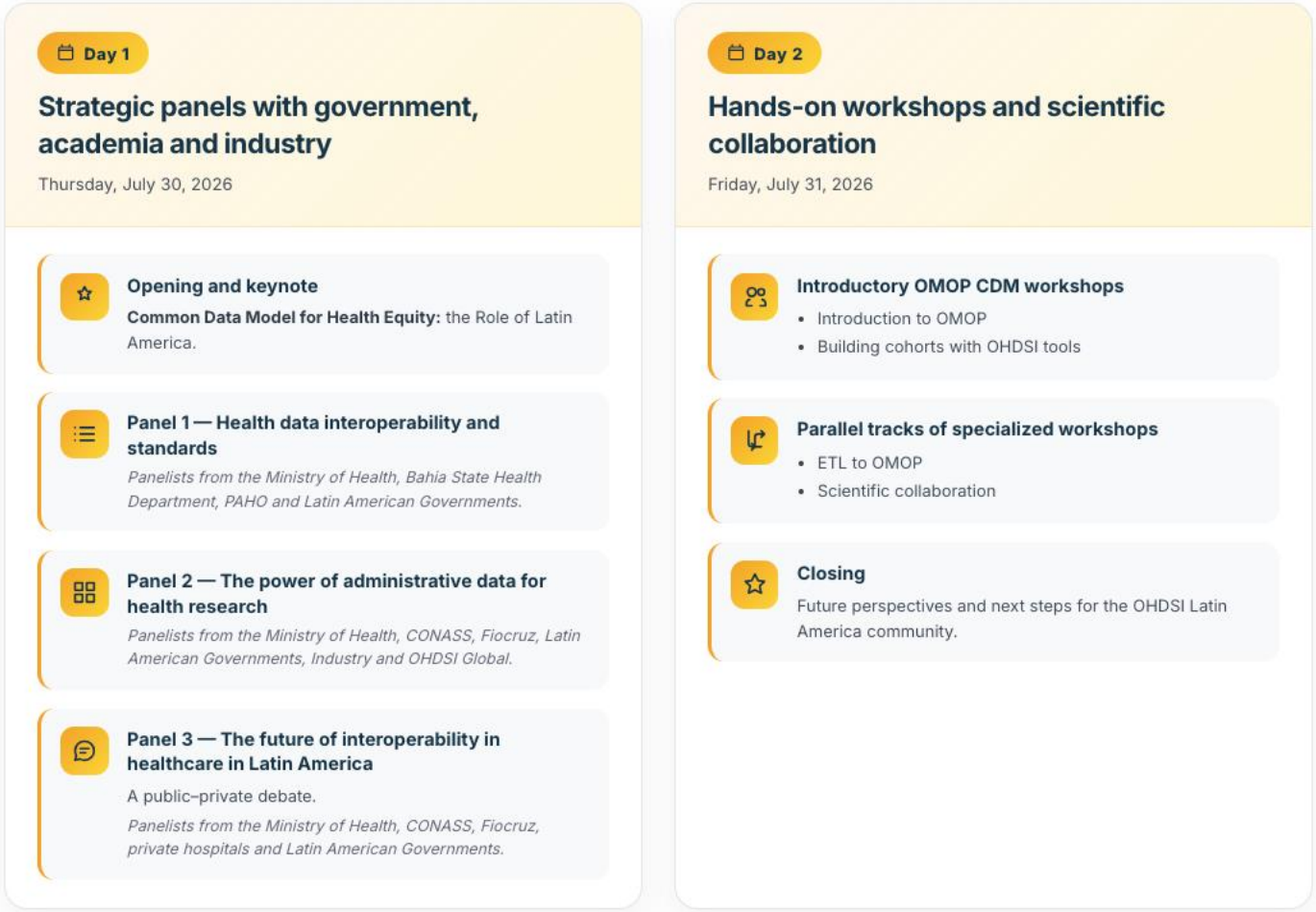
Where Are We Going?

Upcoming Workgroup Calls

Date	Time (ET)	Meeting
Tuesday	12 pm	CDM Vocabulary
Wednesday	10 am	Women of OHDSI
Thursday	6:30 am	India Community Call
Thursday	9 am	Phenotype Development and Evaluation
Thursday	10 am	GIS-Geographic Information System
Thursday	10 am	Africa Chapter
Thursday	11 am	Perinatal & Reproductive Health
Friday	11:30 am	Steering
Monday	10 am	Health Systems Interest Group
Tuesday	10 am	CDM Survey

First Latin America Symposium – July 30-31

Registration is open for the first OHDSI Latin America Symposium, taking place July 30-31 in Salvador, Brazil.



ohdsilatam.org

2026 OHDSI Global Symposium

Registration is OPEN for the **2026 OHDSI Global Symposium**, which will be held Oct. 20-22 in New Brunswick, N.J., USA.

Oct. 20: Tutorials
Oct. 21: Plenaries, Showcase
Oct. 22: Workgroup Activities

ohdsi.org/OHDSI2026

2026 Symposium Tutorials – Session 1

- An Introduction to the Journey from Data to Evidence Using OHDSI
- An Introduction to ATLAS
- Bringing FAIR to Imaging Research with the Medical Imaging OMOP Extension
- Complex Phenotyping at Scale with and without LLMs Using PhenotypeR
- OHDSI Leadership Storytelling Workshop
- Mastering OMOP: Transforming EHR Data with Practical Strategies, Best Practices, and OHDSI Integration

2026 Symposium Tutorials – Session 2

- Building and Using the OHDSI Evidence Network: From Data Partner to Federated Study Execution
- From Multi-Modal Data to Real-World Evidence: Hands-on with the Data2Evidence Platform for OMOP Data Curation and Analytics
- Integrating Geospatial Data Into OMOP CDM
- Introduction to OHDSI Phenotype Development & Evaluation
- OHDSI Standardized Vocabularies on FHIR: A Deep Dive Using the Echidna Terminology Server
- Using OMOP Model in Registry Context & Clinical Trials Standardization Context: Conventions, Past Use Cases, SDTM & Regulatory Consideration, Challenges

2026 Global Symposium Agenda

Start	End	Topic	Presenter/Lead
8:00 am	8:30 am	State of the Community	George Hripcsak
8:30 am	9:15 am	OHDSI Year In Review	Early-Stage Researcher WG
9:15 am	10:00 am	Collaborator Showcase: Posters and Demos (Session 1)	
10:00 am	11:00 am	Plenary 1: Federated Learning Meets Negative Control Calibration: Toward Reliable Multi-Site Evidence Generation	Yong Chen
11:00 am	12:00 pm	Plenary 2: Beyond the Defaults: How the OHDSI Community is Adapting, Extending, and Reimagining Its Tools	Scott Duvall
12:00 pm	1:00 pm	Network & Lunch	
1:00 pm	2:00 pm	Plenary 3: The role of national initiatives in supporting sustainability, collaboration, and growth of OHDSI	Ed Burn
2:00 pm	2:45 pm	Collaborator Showcase: Lightning Talks (Session 1)	5 presenters
2:45 pm	3:30 pm	Collaborator Showcase: Posters and Demos (Session 2)	
3:30 pm	4:15 pm	Collaborator Showcase: Posters and Demos (Session 3)	
4:15 pm	5:00 pm	Collaborator Showcase: Lightning Talks (Session 2)	5 presenters
5:00 pm	6:00 pm	Titan Awards, Closing	Patrick Ryan
6:00 pm	8:30 pm	Dinner on your own	
8:30 pm	11:30 pm	OHDSI Jam Session	Martijn Schuemie

2026 Symposium Workgroup Activities

Session 1 (8 am – 10 am): Eyecare and Vision Research, GIS – Geographic Information Systems, Early-Stage Researchers, Industry, Tidy R Programming with OMOP, HADES Hackathon, Generative AI and Foundation Models, Phenotype Development and Evaluation, Oncology, Health Equity, Vocabularies, APAC

Session 2 (10:30 am – 12:30 pm): Perinatal and Reproductive Health, Waveform, Medical Imaging, Industry, Tidy R Programming with OMOP, HADES Hackathon, Generative AI and Foundation Models, Phenotype Development and Evaluation, Oncology, Health Equity, Vocabularies, Rare Disease

Session 3 (1:30 pm – 3:30 pm): Dentistry, GIS & Waveform Cross-Pollination Meeting, Evidence Network, Women of OHDSI, Psychiatry, HADES Hackathon, Natural Language Processing, Health Economics & Value Assessment, ATLAS/WebAPI, CDM Survey, Surgery & Perioperative Medicine

Session 4 (3:30 pm – 5 pm): Workgroup Summary Session

2026 OHDSI Symposium Jam Session

Whether you are an amateur hobbyist, a seasoned pro, an instrumentalist, or a vocalist, our very first OHDSI Jam Session is open to everyone. We'll do some free-form improvisation, but we will also vote on a few songs ahead of time so everyone has a chance to practice.

The Gear: The organizers will provide a backline with amplifiers, drums, and microphones.
Your Instruments: We encourage you to bring your own gear! If you're open to letting fellow community members plug in or play your instrument, please let us know.
The Vibe: Low pressure, high energy, and entirely collaborative.
Want to join?

Let us know if you're thinking about participating so we can coordinate the song list and gear!

APAC Symposium: Nov. 13-15

The 2026 APAC Symposium will be held Nov. 13-15 at Yonsei University in Seoul, South Korea.

Nov. 13: Main Conference
Nov. 14: Tutorials, Datathon
Nov. 15: Datathon Team Activities

Registration will open soon, and the showcase submission deadline is August 17 at 8 pm ET.

ohdsi.org/APAC2026

Atlas 3 Demo Site

- Presented by Peter Hoffmann at 7/7 OHDSI Community Call
- We are asking for community feedback
- Published to https://atlas-preview.ohdsi.org/atlas/
  - Db Authenticaton enabled: use ohdsi/ohdsi
  - Will enable anonymous access in a future update
- Discussions at https://github.com/OHDSI/Atlas3/discussions

Wanted: HADES Developers

Help wanted - developers for HADES shiny viewer

Developers

jreps 5d

We have a set of R packages for developing modular shiny apps to explore the standardized results returned via HADES packages.

- OhdsiShinyAppBuilder - this builds the shiny app and defines what modules to include (e.g., cohort exporer, characterization, estimation, …)
- OhdsiShinyModules - the main shiny code for the key HADES packages are here
- OhdsiReportGenerator - this is a set of SQL that pulls out friendly data.framas with the results

However, there is a lot of potential for improvements in these packages including (but not limited to):

- adding bookmarking to the URL for easy sharing (Chapter 11 Bookmarking | Mastering Shiny)
- improving display time (improving data extraction speed)
- improving the UI

If you are a developer with knowledge or interest in R or shiny or web development/design and UI and want to help out with HADES development please post below or message me.

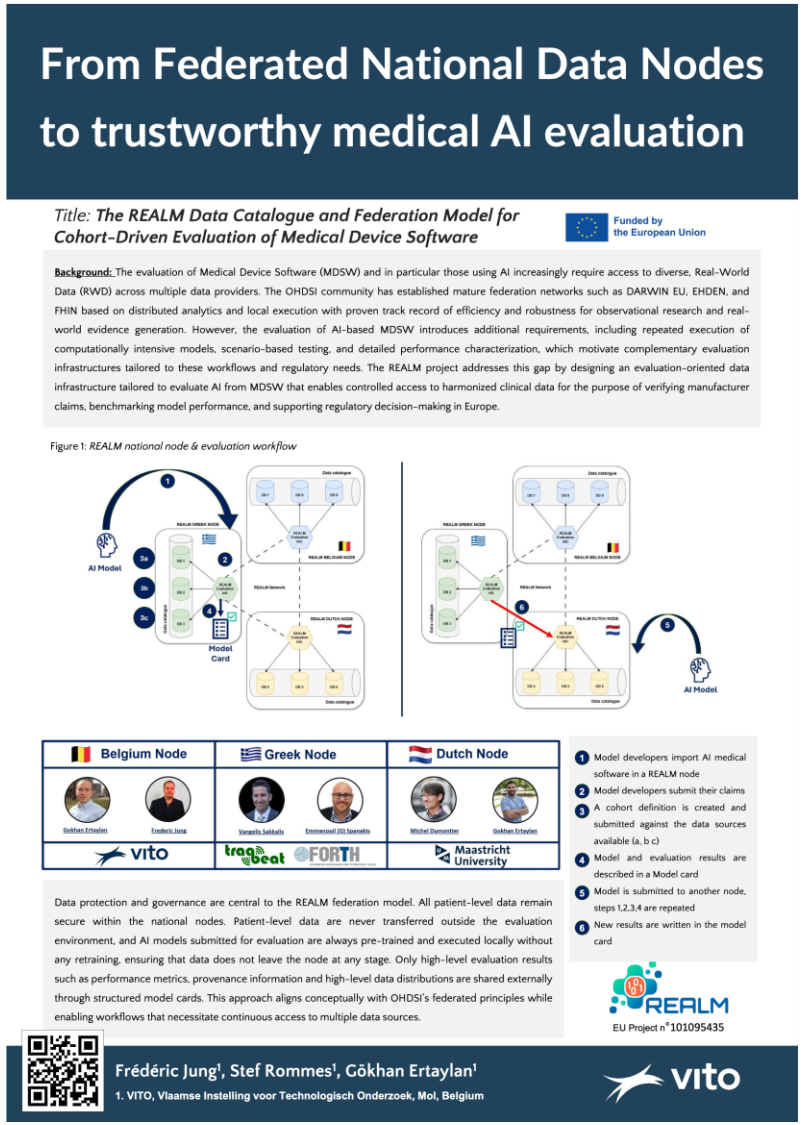
I'm planning to set up regular meetings for this work group.

#OHDSISocialShowcase This Week

Monday

The REALM Data Catalogue and Federation Model for Cohort-Driven Evaluation of Medical Device Software

(Frederic Jung, Stef Rommes, Gökhan Ertaylan)

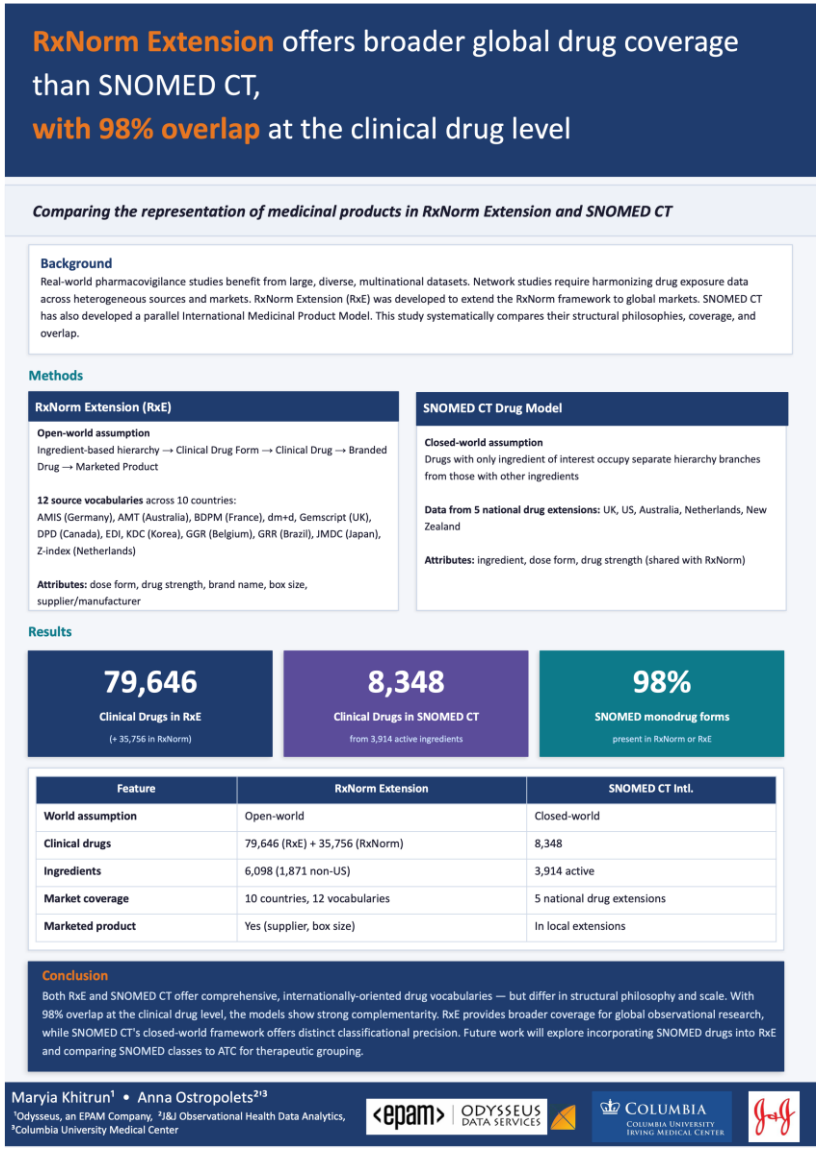


#OHDSISocialShowcase This Week

Tuesday

Comparing the representation of medicinal products in RxNorm Extension and SNOMED CT

(Maryia Khitrun, Anna Ostropolets)



RxNorm Extension offers broader global drug coverage than SNOMED CT,
with 98% overlap at the clinical drug level

Comparing the representation of medicinal products in RxNorm Extension and SNOMED CT

Background
Real-world pharmacovigilance studies benefit from large, diverse, multinational datasets. Network studies require harmonizing drug exposure data across heterogeneous sources and markets. RxNorm Extension (RxE) was developed to extend the RxNorm framework to global markets. SNOMED CT has also developed a parallel International Medicinal Product Model. This study systematically compares their structural philosophies, coverage, and overlap.

Methods

RxNorm Extension (RxE)

Open-world assumption
Ingredient-based hierarchy → Clinical Drug Form → Clinical Drug → Branded Drug → Marketed Product

12 source vocabularies across 10 countries:
AMIS (Germany), AMT (Australia), BDPM (France), dm+d, Gemscript (UK), DPD (Canada), EDI, KDC (Korea), GGR (Belgium), GRR (Brazil), JMDC (Japan), Z-index (Netherlands)

Attributes: dose form, drug strength, brand name, box size, supplier/manufacturer

SNOMED CT Drug Model

Closed-world assumption
Drugs with only ingredient of interest occupy separate hierarchy branches from those with other ingredients

Data from 5 national drug extensions: UK, US, Australia, Netherlands, New Zealand

Attributes: ingredient, dose form, drug strength (shared with RxNorm)

Results

79,646	8,348	98%
Clinical Drugs in RxE (+ 35,756 in RxNorm)	Clinical Drugs in SNOMED CT from 3,914 active ingredients	SNOMED monodrug forms present in RxNorm or RxE

Feature	RxNorm Extension	SNOMED CT Intl.
World assumption	Open-world	Closed-world
Clinical drugs	79,646 (RxE) + 35,756 (RxNorm)	8,348
Ingredients	6,098 (1,871 non-US)	3,914 active
Market coverage	10 countries, 12 vocabularies	5 national drug extensions
Marketed product	Yes (supplier, box size)	In local extensions

Conclusion
Both RxE and SNOMED CT offer comprehensive, internationally-oriented drug vocabularies — but differ in structural philosophy and scale. With 98% overlap at the clinical drug level, the models show strong complementarity. RxE provides broader coverage for global observational research, while SNOMED CT's closed-world framework offers distinct classificational precision. Future work will explore incorporating SNOMED drugs into RxE and comparing SNOMED classes to ATC for therapeutic grouping.

Maryia Khitrun[1] • Anna Ostropolets[2,3]
[1]Odysseus, an EPAM Company, [2]J&J Observational Health Data Analytics, [3]Columbia University Medical Center

#OHDSISocialShowcase This Week

Wednesday

Explainable, Governance-Ready Clinical NLP Pipeline for OMOP-CDM Using Locally Deployed LLMs

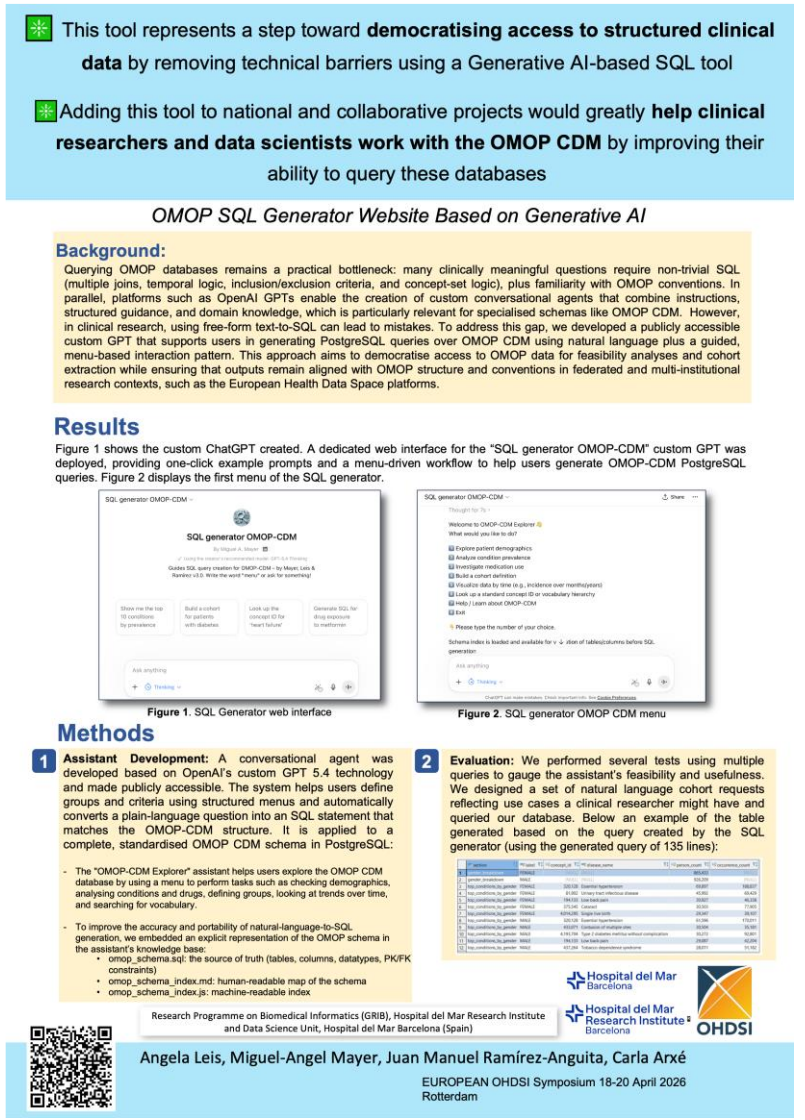
(Alisson Licona, Muhammad Waseem, Alex Loleit)

#OHDSISocialShowcase This Week

Thursday

OMOP SQL Generator Website Based on Generative AI

(Angela Leis, Miguel-Angel Mayer, Juan Manuel Ramírez-Anguita, Carla Arxé)

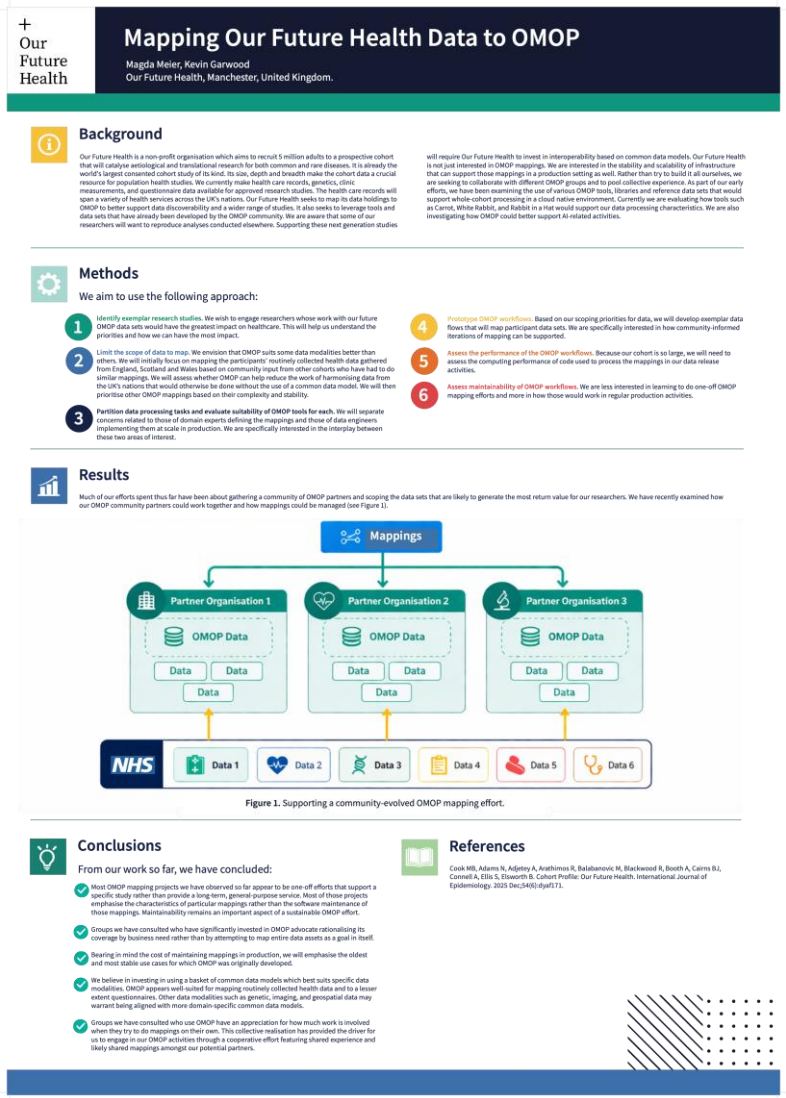


#OHDSISocialShowcase This Week

Friday

Mapping Our Future Health data to OMOP

(Magda Meier, Kevin Garwood)



Our Future Health

Mapping Our Future Health Data to OMOP

Magda Meier, Kevin Garwood
Our Future Health, Manchester, United Kingdom.

Background

Our Future Health is a non-profit organisation which aims to recruit 5 million adults to a prospective cohort that will catalyse aetiological and translational research for both common and rare diseases. It is already the world's largest consented cohort study of its kind. Its size, depth and breadth make the cohort data a crucial resource for population health studies. We currently make health care records, genetics, clinic measurements, and questionnaire data available for approved research studies. The health care records will span a variety of health services across the UK's nations. Our Future Health seeks to map its data holdings to OMOP to better support data discoverability and a wider range of studies. It also seeks to leverage tools and data sets that have already been developed by the OMOP community. We are aware that some of our researchers will want to reproduce analyses conducted elsewhere. Supporting these next generation studies will require Our Future Health to invest in interoperability based on common data models. Our Future Health is not just interested in OMOP mappings. We are interested in the stability and scalability of infrastructure that can support those mappings in a production setting as well. Rather than try to build it all ourselves, we are seeking to collaborate with different OMOP groups and to pool collective experience. As part of our early efforts, we have been examining the use of various OMOP tools, libraries and reference data sets that would support whole-cohort processing in a cloud native environment. Currently we are evaluating how tools such as Carrot, White Rabbit, and Rabbit in a Hat would support our data processing characteristics. We are also investigating how OMOP could better support AI-related activities.

Methods

We aim to use the following approach:

1. **Identify exemplar research studies.** We wish to engage researchers whose work with our future OMOP data sets would have the greatest impact on healthcare. This will help us understand the priorities and how we can have the most impact.
2. **Limit the scope of data to map.** We envision that OMOP suits some data modalities better than others. We will initially focus on mapping the participants' routinely collected health data gathered from England, Scotland and Wales based on community input from other cohorts who have had to do similar mappings. We will assess whether OMOP can help reduce the work of harmonising data from the UK's nations that would otherwise be done without the use of a common data model. We will then prioritise other OMOP mappings based on their complexity and stability.
3. **Partition data processing tasks and evaluate suitability of OMOP tools for each.** We will separate concerns related to those of domain experts defining the mappings and those of data engineers implementing them at scale in production. We are specifically interested in the interplay between these two areas of interest.
4. **Prototype OMOP workflows.** Based on our scoping priorities for data, we will develop exemplar data flows that will map participant data sets. We are specifically interested in how community-informed iterations of mapping can be supported.
5. **Assess the performance of the OMOP workflows.** Because our cohort is so large, we will need to assess the computing performance of code used to process the mappings in our data release activities.
6. **Assess maintainability of OMOP workflows.** We are less interested in learning to do one-off OMOP mapping efforts and more in how those would work in regular production activities.

Results

Much of our efforts spent thus far have been about gathering a community of OMOP partners and scoping the data sets that are likely to generate the most return value for our researchers. We have recently examined how our OMOP community partners could work together and how mappings could be managed (see Figure 1).

Figure 1. Supporting a community-evolved OMOP mapping effort.

Conclusions

From our work so far, we have concluded:

- Most OMOP mapping projects we have observed so far appear to be one-off efforts that support a specific study rather than provide a long-term, general-purpose service. Most of those projects emphasise the characteristics of particular mappings rather than the software maintenance of those mappings. Maintainability remains an important aspect of a sustainable OMOP effort.
- Groups we have consulted who have significantly invested in OMOP advocate rationalising its coverage by business need rather than by attempting to map entire data assets as a goal in itself.
- Bearing in mind the cost of maintaining mappings in production, we will emphasise the oldest and most stable use cases for which OMOP was originally developed.
- We believe in investing in using a basket of common data models which best suits specific data modalities. OMOP appears well-suited for mapping routinely collected health data and to a lesser extent questionnaires. Other data modalities such as genetic, imaging, and geospatial data may warrant being aligned with more domain-specific common data models.
- Groups we have consulted who use OMOP have an appreciation for how much work is involved when they try to do mappings on their own. This collective realisation has provided the driver for us to engage in our OMOP activities through a cooperative effort featuring shared experience and likely shared mappings amongst our potential partners.

References

Cook MB, Adams N, Adjetey A, Arathimos R, Balabanovic M, Blackwood R, Booth A, Cairns BJ, Connell A, Ellis S, Elsworth B. Cohort Profile: Our Future Health. International Journal of Epidemiology. 2025 Dec;54(6):dyaf171.

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?

July 28: Newcomer Introductions

Ryan Chou
Louise Sun
András Novoszáth
Zongzhi Zach Liu
Marianna Macera
Pulkit Gairola
Asha Mahesh
Asif Syed
Michael Gaile
Anne Marie Chan Cheung Shing
Carla Laladula
Jonathan Lowe
Sandy Estremera-Zink
Beatrice Okrah

Welcome to the annual **Welcome OHDSI Newcomers** community call! If you are relatively new to the community or hope to have a bigger role, this is a great opportunity to introduce yourself.

Please tell us who you are and where you work, what are your research interests, how you hope to help the community, and how OHDSI can help your own research journey.

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-2026

Find your workgroup.

Fuel our mission.

ohdsi.org/workgroups