LLM Research in the OHDSI Community (Session 2 of 3)

OHDSI Community Call
June 9, 2026 • 11 am ET

OHDSI
Observational Health Data Sciences and Informatics

Upcoming Community Calls

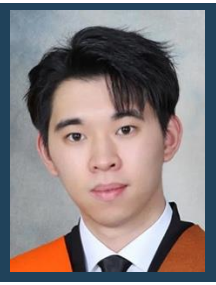
Date	Topic
June 9	LLM Research Around The World, Session 2
June 16	LLM Research Around The World, Session 3
June 23	CANCELLED: OHDSI Summer School at Columbia University
June 30	OMOP & OHDSI Research Spotlight

June 16: LLM Research in the OHDSI Community

Richard Boyce

University of Pittsburgh

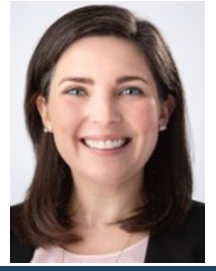
Study Agent support for users of Hades



Shihao Shenzhang

King’s College, London

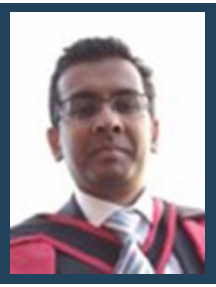
FastSSV, a semantic static validator for LLM Generated queries



Brooke Wolford

Norwegian University of Science and Technology / University of Oslo

Implementation of foundation models in the Trøndelag Health Study



Vishnu V Chandrabalan

Lancashire Teaching Hospitals NHS Foundation Trust

FastOMOP: Agent Harness and Ecosystem for RWE

Three Stages of The Journey

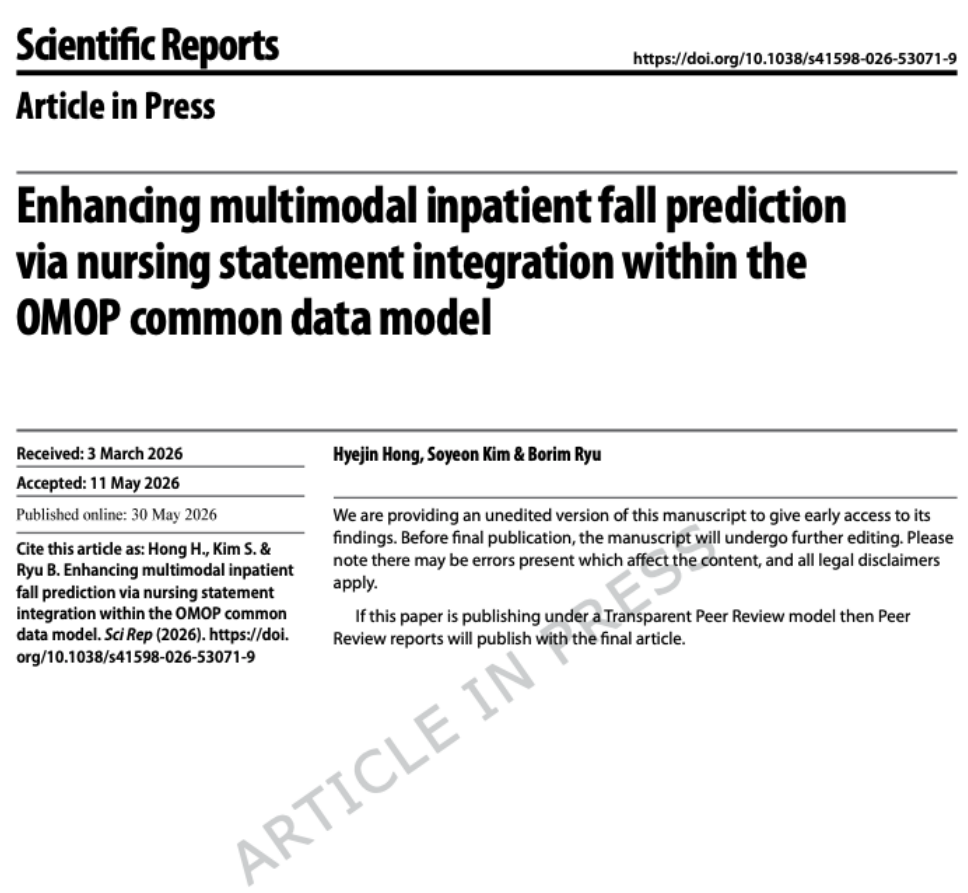
Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **Hyejin Hong, Soyeon Kim, and Borim Ryu** on the recent publication of **Enhancing multimodal inpatient fall prediction via nursing statement integration within the OMOP common data model** in *JAMIA Open*.



Scientific Reports

https://doi.org/10.1038/s41598-026-53071-9

Article in Press

Enhancing multimodal inpatient fall prediction via nursing statement integration within the OMOP common data model

Received: 3 March 2026

Accepted: 11 May 2026

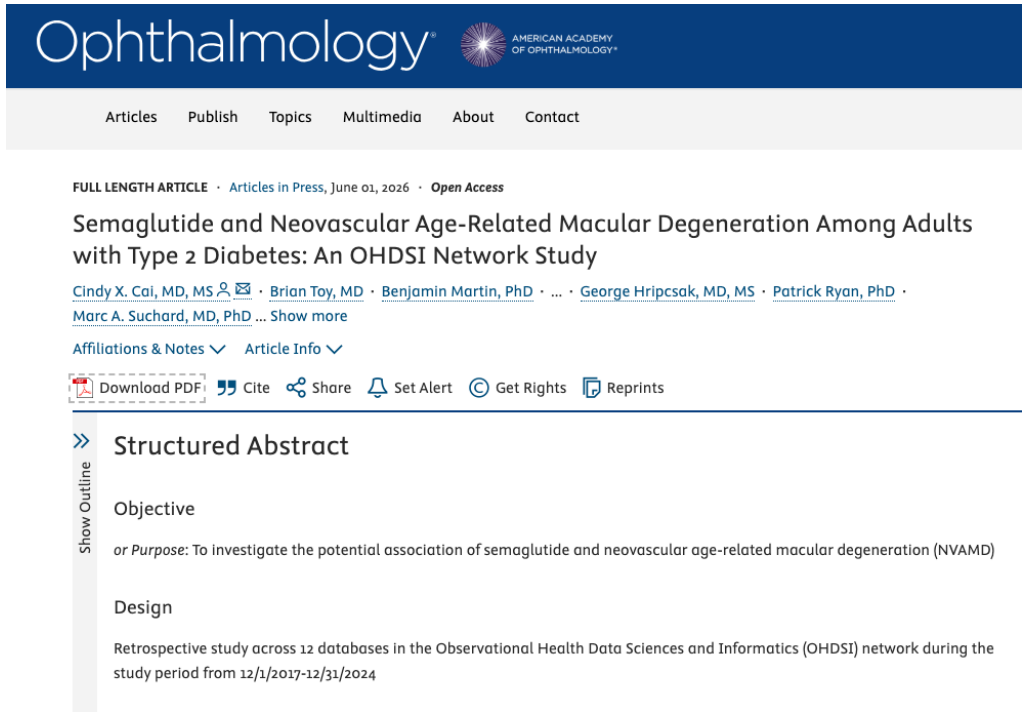
Published online: 30 May 2026

Cite this article as: Hong H., Kim S. & Ryu B. Enhancing multimodal inpatient fall prediction via nursing statement integration within the OMOP common data model. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-53071-9

Hyejin Hong, Soyeon Kim & Borim Ryu

OHDSI Shoutouts!

Congratulations to the team of **Cindy X Cai, Brian Toy, Benjamin Martin, Ruochong Fan, Erik Westlund, Diep Tran, Akihiko Nishimura, Haeun Lee, Theodore Leng, Paul Nagy, Nestoras Mathioudakis, Linying Zhang, Michelle Hribar, Aiyin Chen, Karen Armbrust, Kerry Goetz, Sally Baxter, Michael V Boland, Eric N Brown, Edmund Tsui, Andrew J Barkmeier, Sophia Wang, Nitish Mehta, Jacqueline C Stocking, Ghazala O'Keefe, Cecilia S Lee, Philip R O Payne, William J O'Brien, Scott DuVall, Thamir Alshammari, Thomas Falconer, David A Dorr, Izabelle Humes, David McCoy, Mohammed Adibuzzaman, Rumel Mahmood, Hannah Morgan-Cooper, Priya Desai, Shikha Yashwant Kothari, Anthony Sena, Clair Blacketer, Anna Ostropolets, Azza Shoaibi, Gowtham Rao, George Hripcsak, Patrick Ryan, and Marc A Suchard** on the recent publication of **Semaglutide and Neovascular Age-Related Macular Degeneration Among Adults with Type 2 Diabetes: An OHDSI Network Study** in *Ophtalmology*.



Ophthalmology

AMERICAN ACADEMY OF OPHTHALMOLOGY

Articles Publish Topics Multimedia About Contact

FULL LENGTH ARTICLE · Articles in Press, June 01, 2026 · *Open Access*

Semaglutide and Neovascular Age-Related Macular Degeneration Among Adults with Type 2 Diabetes: An OHDSI Network Study

Cindy X. Cai, MD, MS · Brian Toy, MD · Benjamin Martin, PhD · ... · George Hripcsak, MD, MS · Patrick Ryan, PhD · Marc A. Suchard, MD, PhD ... Show more

Affiliations & Notes Article Info

Download PDF Cite Share Set Alert Get Rights Reprints

Show Outline

Structured Abstract

Objective

or Purpose: To investigate the potential association of semaglutide and neovascular age-related macular degeneration (NVAMD)

Design

Retrospective study across 12 databases in the Observational Health Data Sciences and Informatics (OHDSI) network during the study period from 12/1/2017-12/31/2024

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

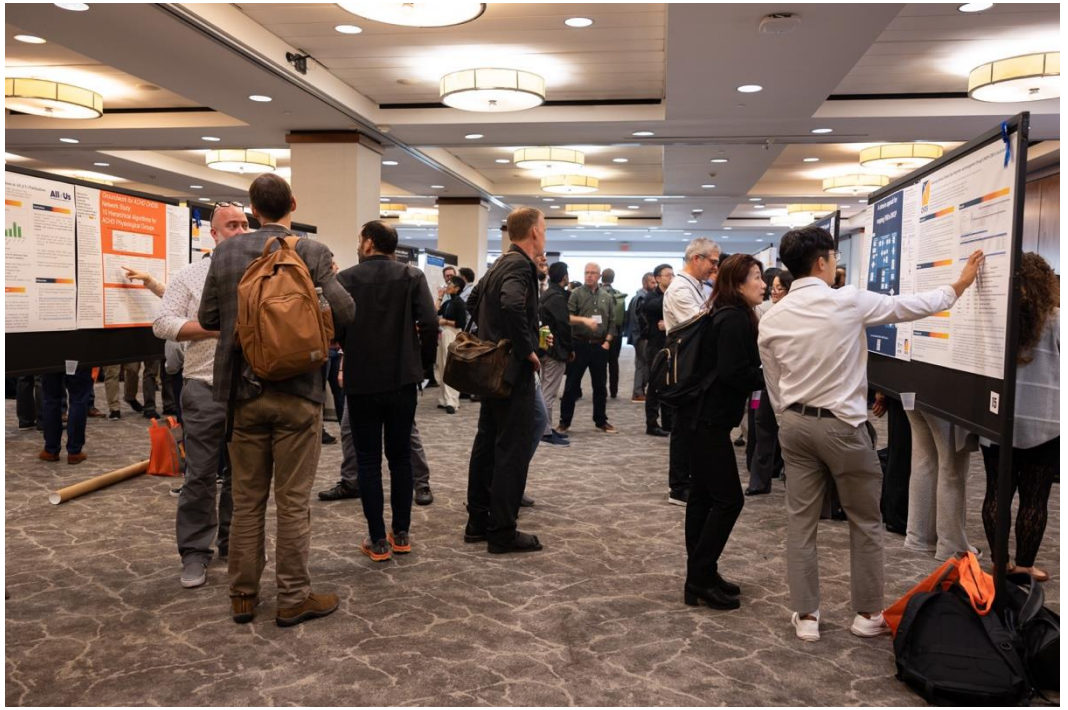
Where Are We Going?

Upcoming Workgroup Calls

Date	Time (ET)	Meeting
Tuesday	12 pm	Generative AI and Analytics
Wednesday	7 am	Medical Imaging
Wednesdy	8 am	Psychiatry
Wednesday	9 am	Patient-Level Prediction
Wednesday	10 am	Common Data Model
Wednesday	2 pm	Natural Language Processing
Wednesday	7 pm	Eyecare and Vision Research
Thursday	7 am	Europe Community Call
Thursday	10 am	GIS-Geographic Information System
Thursday	10 am	Rare Diseases
Thursday	7 pm	Dentistry
Friday	9 am	Waveform
Friday	10 am	Transplant
Friday	11 am	Clinical Trials
Friday	11:30 am	Steering Group
Friday	11 pm	China Chapter (ZOOM)
Monday	2 pm	Electronic Animal Health Records
Tuesday	9 am	Data2Evidence

THANK YOU!

The **call for participation** for the 2026 Global Symposium has ended.



We received another record-setting amount of submissions!

Thank you to the Scientific Review Committee for leading the effort to build our Collaborator Showcase over the next month!

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 Global Symposium

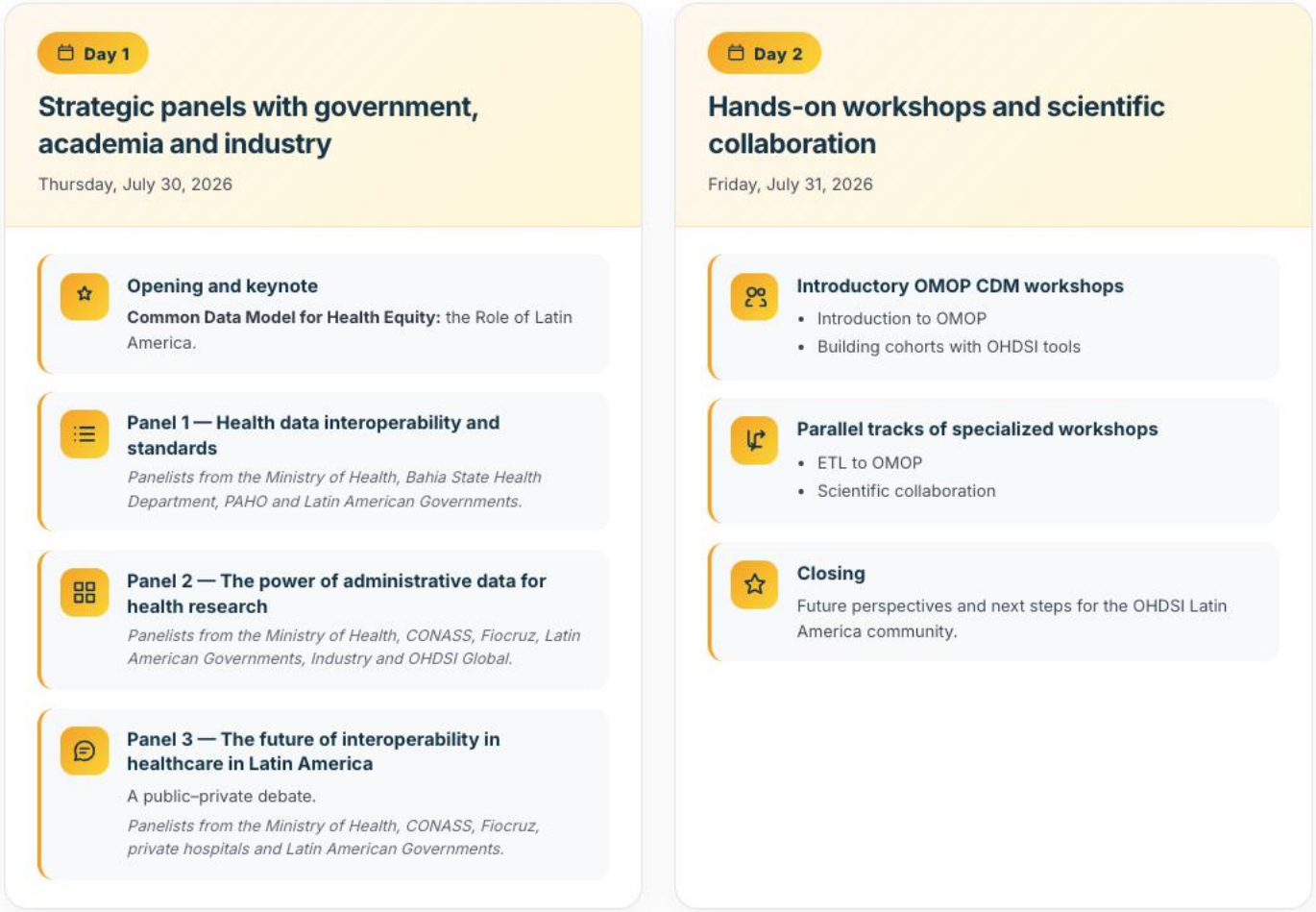
There are opportunities to be both a **sponsor** and an **exhibitor** at the Global Symposium.

Please reach out to **symposium@OHDSI.org** for more information.

ohdsi.org/OHDSI2026

First Latin America Symposium – July 30-31

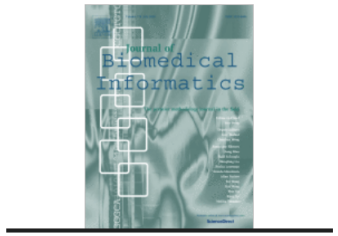
Registration is open for the first OHDSI Latin America Symposium, taking place July 30-31 in Salvador, Brazil.



ohdsilatam.org

Call for Papers: Agentic AI for RWE Life Cycle Support

The Journal of Biomedical Informatics has opened a call for papers focused on Agentic AI For RWE Life-Cycle Support — scaling trustworthy evaluation for treatment safety and effectiveness.



Journal of Biomedical Informatics

Date: Available online 8 June 2026

Article: 105062

In Press, Journal Pre-proof

Published by

ELSEVIER

Editorial • Full text access Get rights and content ↗

Editorial

Call for papers: agentic AI for real-world evidence life cycle support – scaling trustworthy evaluation of treatment safety and effectiveness

Show more

View PDF Cite Add to Mendeley Share 10.1016/j.jbi.2026.105062 ↗

Article

Guest editors:

Richard Boyce (Contact) rdb20@pitt.edu; Department of Biomedical Informatics, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA.

Scott Alexander Malec, smalec@salud.unm.edu; Translational Informatics Division, School of Medicine, University of New Mexico, Albuquerque, NM, USA.

Chunhua Weng. cw2384@cumc.columbia.edu; Biomedical Informatics, Data Science Institute, Columbia University, New York, NY, USA.

Outline

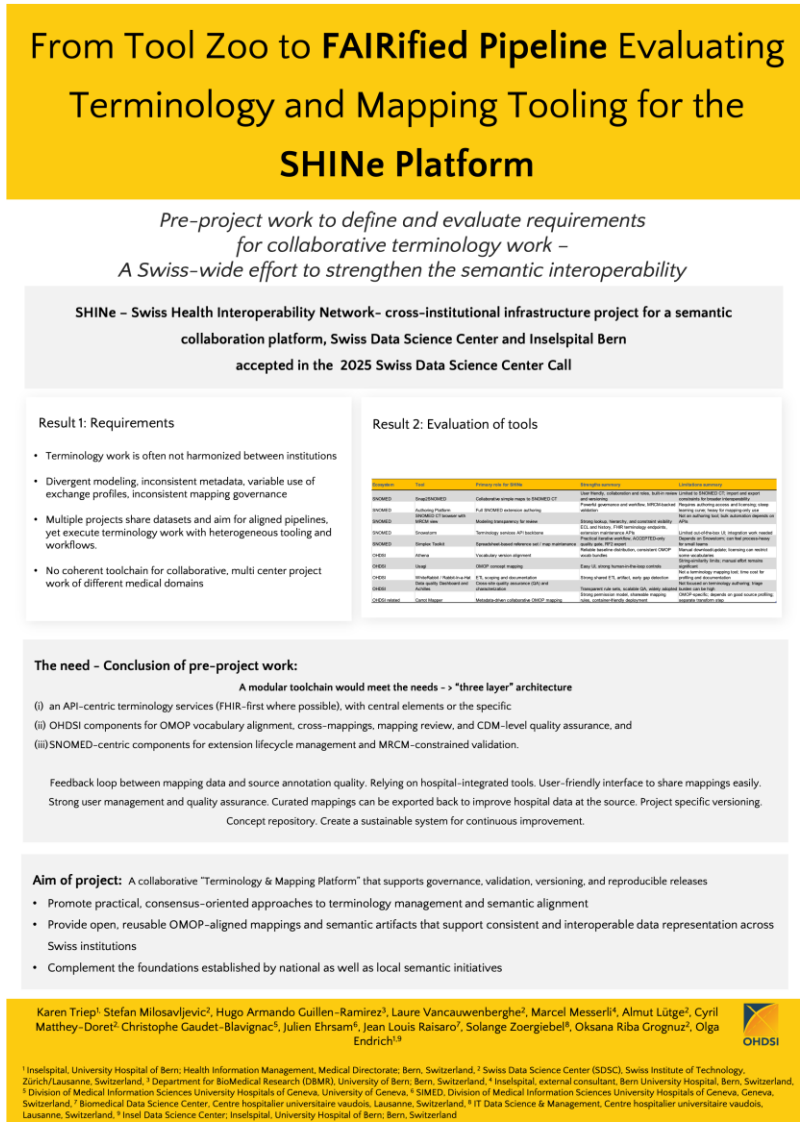
Submissions will be accepted from June 15 through November 15, 2026, with expected publication date of September 15, 2027.

#OHDSISocialShowcase This Week

Monday

From Tool Zoo to FAIRified Pipeline: Evaluating Terminology and Mapping Tooling for the SHINe Platform

(Karen Triep, Stefan Milosavljevic, Hugo Armando Guillen-Ramirez, Laure Vancauwenberghe, Marcel Messerli, Almut Lütge, Cyril Matthey-Doret, Christophe Gaudet-Blavignac, Julien Ehrsam, Jean Louis Raisaro, Solange Zoergiebel, Oksana Riba Grognuz, Olga Endrich)



From Tool Zoo to **FAIRified Pipeline** Evaluating Terminology and Mapping Tooling for the **SHINe Platform**

Pre-project work to define and evaluate requirements for collaborative terminology work – A Swiss-wide effort to strengthen the semantic interoperability

SHINe – Swiss Health Interoperability Network- cross-institutional infrastructure project for a semantic collaboration platform, Swiss Data Science Center and Inselspital Bern accepted in the 2025 Swiss Data Science Center Call

Result 1: Requirements

- Terminology work is often not harmonized between institutions
- Divergent modeling, inconsistent metadata, variable use of exchange profiles, inconsistent mapping governance
- Multiple projects share datasets and aim for aligned pipelines, yet execute terminology work with heterogeneous tooling and workflows.
- No coherent toolchain for collaborative, multi center project work of different medical domains

Result 2: Evaluation of tools

The need - Conclusion of pre-project work:

A modular toolchain would meet the needs - > "three layer" architecture

(i) an API-centric terminology services (FHIR-first where possible), with central elements or the specific
(ii) OHDSI components for OMOP vocabulary alignment, cross-mappings, mapping review, and CDM-level quality assurance, and
(iii) SNOMED-centric components for extension lifecycle management and MRCM-constrained validation.

Feedback loop between mapping data and source annotation quality. Relying on hospital-integrated tools. User-friendly interface to share mappings easily. Strong user management and quality assurance. Curated mappings can be exported back to improve hospital data at the source. Project specific versioning. Concept repository. Create a sustainable system for continuous improvement.

Aim of project: A collaborative "Terminology & Mapping Platform" that supports governance, validation, versioning, and reproducible releases

- Promote practical, consensus-oriented approaches to terminology management and semantic alignment
- Provide open, reusable OMOP-aligned mappings and semantic artifacts that support consistent and interoperable data representation across Swiss institutions
- Complement the foundations established by national as well as local semantic initiatives

Karen Triep[1] Stefan Milosavljevic[2], Hugo Armando Guillen-Ramirez[3], Laure Vancauwenberghe[2], Marcel Messerli[4], Almut Lütge[3], Cyril Matthey-Doret[2] Christophe Gaudet-Blavignac[5], Julien Ehrsam[6], Jean Louis Raisaro[7], Solange Zoergiebel[8], Oksana Riba Grognuz[2], Olga Endrich[1,9]

[1] Inselspital, University Hospital of Bern; Health Information Management, Medical Directorate; Bern, Switzerland, [2] Swiss Data Science Center (SDSC), Swiss Institute of Technology, Zürich/Lausanne, Switzerland, [3] Department for BioMedical Research (DBMR), University of Bern; Bern, Switzerland, [4] Inselspital, external consultant, Bern University Hospital, Bern, Switzerland, [5] Division of Medical Information Sciences University Hospitals of Geneva, University of Geneva, [6] SIMED, Division of Medical Information Sciences University Hospitals of Geneva, Geneva, Switzerland, [7] Biomedical Data Science Center, Centre hospitalier universitaire vaudois, Lausanne, Switzerland, [8] IT Data Science & Management, Centre hospitalier universitaire vaudois, Lausanne, Switzerland, [9] Insel Data Science Center; Inselspital, University Hospital of Bern; Bern, Switzerland

#OHDSISocialShowcase This Week

Tuesday

Mitos: A Python implementation of OHDSI Circe cohort expressions with schema fidelity and cross-engine validation

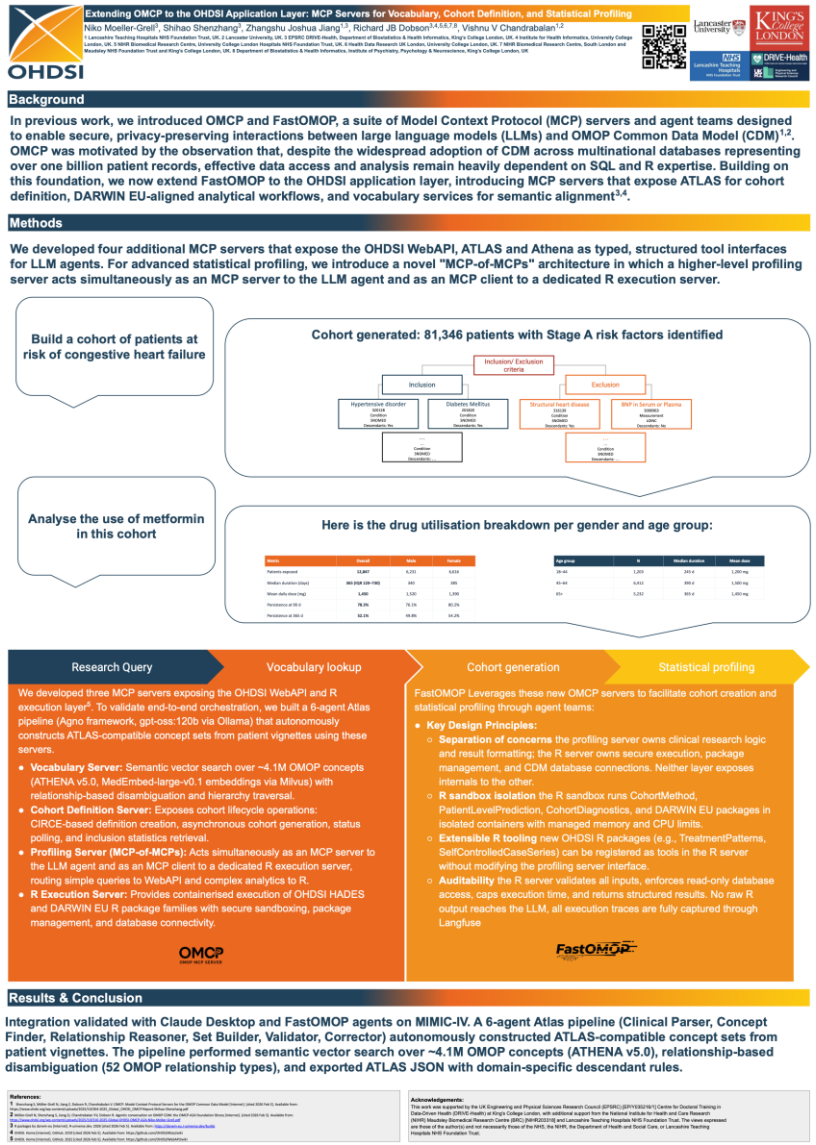
(Egill Fridgeirsson, Peter Rijnbeek)

#OHDSISocialShowcase This Week

Wednesday

Extending OMCP to the OHDSI Application Layer: MCP Servers for Vocabulary, Cohort Definition, and Statistical Profiling

(Niko Möller-Grell, Shihao Shenzhang, Zhangshu Joshua Jiang, Richard Dobson, Vishnu V Chandrabalan)



Extending OMCP to the OHDSI Application Layer: MCP Servers for Vocabulary, Cohort Definition, and Statistical Profiling

Niko Moeller-Grell[1], Shihao Shenzhang[1], Zhangshu Joshua Jiang[1,3], Richard JB Dobson[2,4,5,6,7,8], Vishnu V Chandrabalan[1,2]

OHDSI

Background

In previous work, we introduced OMCP and FastOMOP, a suite of Model Context Protocol (MCP) servers and agent teams designed to enable secure, privacy-preserving interactions between large language models (LLMs) and OMOP Common Data Model (CDM)[1,2]. OMCP was motivated by the observation that, despite the widespread adoption of CDM across multinational databases representing over one billion patient records, effective data access and analysis remain heavily dependent on SQL and R expertise. Building on this foundation, we now extend FastOMOP to the OHDSI application layer, introducing MCP servers that expose ATLAS for cohort definition, DARWIN EU-aligned analytical workflows, and vocabulary services for semantic alignment[3,4].

Methods

We developed four additional MCP servers that expose the OHDSI WebAPI, ATLAS and Athena as typed, structured tool interfaces for LLM agents. For advanced statistical profiling, we introduce a novel "MCP-of-MCPs" architecture in which a higher-level profiling server acts simultaneously as an MCP server to the LLM agent and as an MCP client to a dedicated R execution server.

Build a cohort of patients at risk of congestive heart failure

Cohort generated: 81,346 patients with Stage A risk factors identified

Analyse the use of metformin in this cohort

Here is the drug utilisation breakdown per gender and age group:

Research Query	Vocabulary lookup	Cohort generation	Statistical profiling

We developed three MCP servers exposing the OHDSI WebAPI and R execution layer[4]. To validate end-to-end orchestration, we built a 6-agent Atlas pipeline (Agno framework, gpt-oss:120b via Ollama) that autonomously constructs ATLAS-compatible concept sets from patient vignettes using these servers.

- **Vocabulary Server:** Semantic vector search over ~4.1M OMOP concepts (ATHENA v5.0, MedEmbed-large-v0.1 embeddings via Milvus) with relationship-based disambiguation and hierarchy traversal.
- **Cohort Definition Server:** Exposes cohort lifecycle operations: CIRCE-based definition creation, asynchronous cohort generation, status polling, and inclusion statistics retrieval.
- **Profiling Server (MCP-of-MCPs):** Acts simultaneously as an MCP server to the LLM agent and as an MCP client to a dedicated R execution server, routing simple queries to WebAPI and complex analytics to R.
- **R Execution Server:** Provides containerised execution of OHDSI HADES and DARWIN EU R package families with secure sandboxing, package management, and database connectivity.

OMCP

FastOMOP Leverages these new OMCP servers to facilitate cohort creation and statistical profiling through agent teams:

- **Key Design Principles:**
  - **Separation of concerns** the profiling server owns clinical research logic and result formatting; the R server owns secure execution, package management, and CDM database connections. Neither layer exposes internals to the other.
  - **R sandbox isolation** the R sandbox runs CohortMethod, PatientLevelPrediction, CohortDiagnostics, and DARWIN EU packages in isolated containers with managed memory and CPU limits.
  - **Extensible R tooling** new OHDSI R packages (e.g., TreatmentPatterns, SelfControlledCaseSeries) can be registered as tools in the R server without modifying the profiling server interface.
  - **Auditability** the R server validates all inputs, enforces read-only database access, caps execution time, and returns structured results. No raw R output reaches the LLM, all execution traces are fully captured through Langfuse

FastOMOP

Results & Conclusion

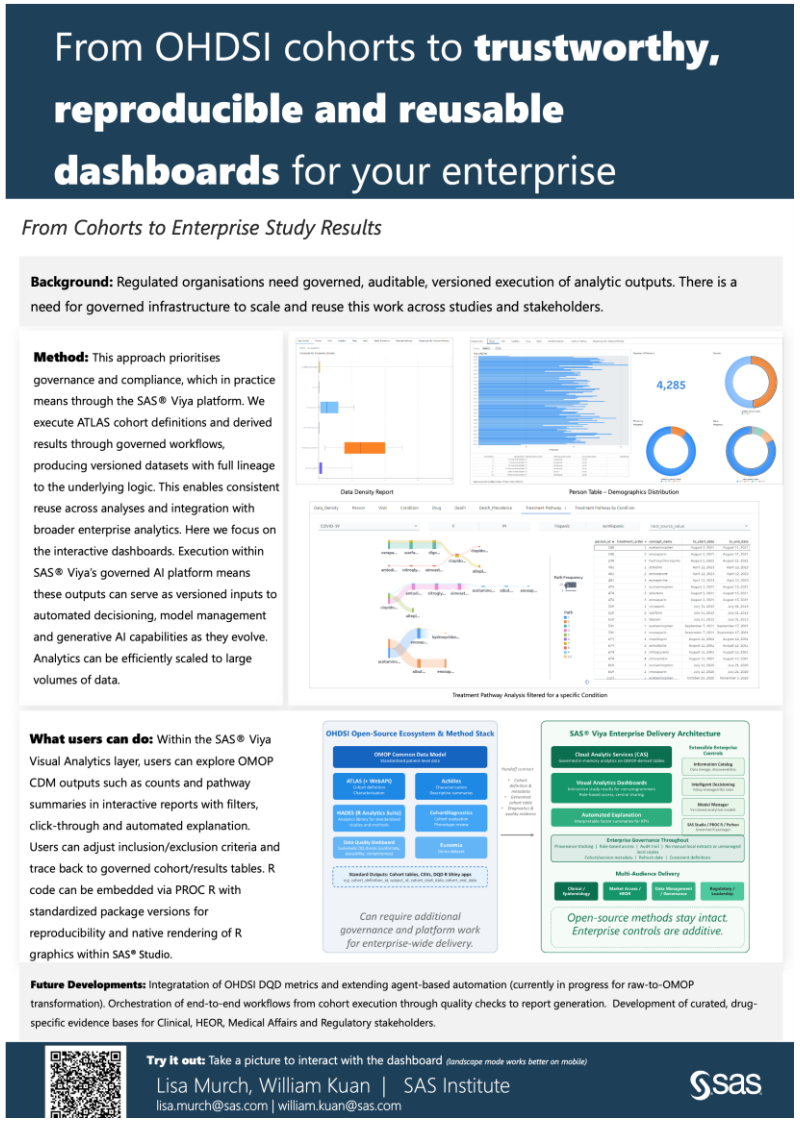
Integration validated with Claude Desktop and FastOMOP agents on MIMIC-IV. A 6-agent Atlas pipeline (Clinical Parser, Concept Finder, Relationship Reasoner, Set Builder, Validator, Corrector) autonomously constructed ATLAS-compatible concept sets from patient vignettes. The pipeline performed semantic vector search over ~4.1M OMOP concepts (ATHENA v5.0), relationship-based disambiguation (52 OMOP relationship types), and exported ATLAS JSON with domain-specific descendant rules.

#OHDSISocialShowcase This Week

Thursday

From OHDSI Cohorts to Trustworthy, Reproducible and Reusable Dashboards for your Enterprise

(Lisa Murch, William Kuan)

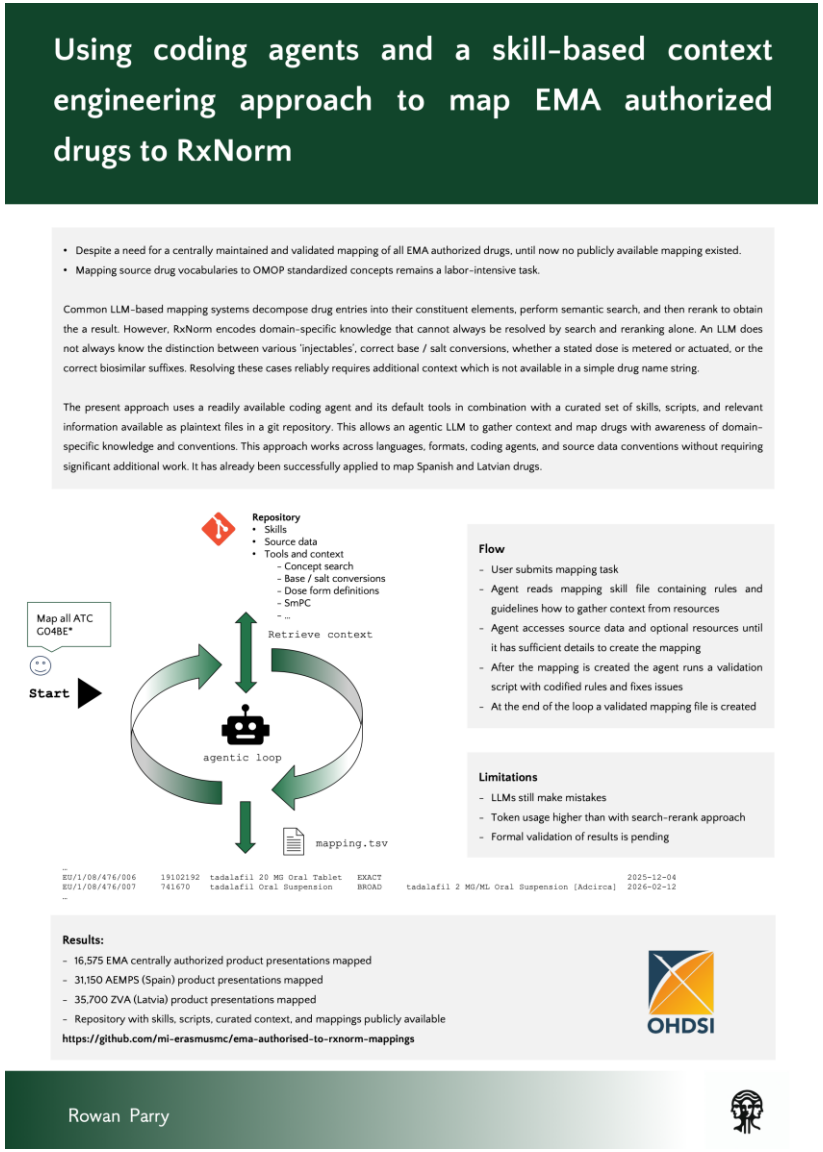


#OHDSISocialShowcase This Week

Friday

Using coding agents and a skill-based context engineering approach to map EMA authorized drugs to RxNorm

(Rowan Parry)



Where Are We Going?

Any other announcements of upcoming work, events, deadlines, etc?

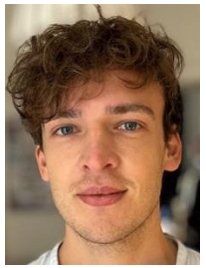
Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

June 9: LLM Research in the OHDSI Community



Niko Moeller-Grell

King's College

FastOMOP - multi agent cohort creation

Joel Swerdel

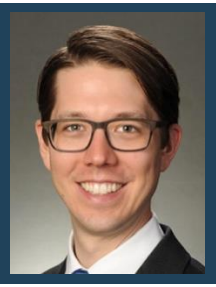
Johnson & Johnson

Phenelope – tool for developing concept sets using LLM

Jared Houghtaling

Johnson & Johnson

LLM-Based Phenotype Refinement via CAPR



Adam Johnson

Duke University

Using Synthetic Data and Claude Code to Develop Transportable Analytic Code

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