OHDSI Newcomer Introductions

OHDSI Community Call
Aug. 12, 2025 • 11 am ET

Upcoming Community Calls

Date	Topic
Aug. 19	Tribute to Jamie Weaver
Aug. 26	Large-Language Model Innovations in OHDSI
Sept. 2	Standardized Vocabulary Summer Refresh Update
Sept. 9	Global Symposium Preview
Sept. 16	TBA
Sept. 23	Recent OHDSI Publications
Sept. 30	OHDSI 2025 Poster Preview Mad Minutes / Symposium Logistics
Oct. 7	No Call – OHDSI Symposium
Oct. 14	Welcome to OHDSI
Oct. 21	Meet the Titans

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **Christine Ramsey-Hardy, Melissa Skanderson, Janet Tate, Amy Justice, Vincent Marconi, Charles Alcorn, Ronald Hauser, Amy Anderson-Mellies, and Kathleen McGinnis** on the publication of **Comparison of 2 electronic health record data extraction methods for laboratory tests used in the Veterans Aging Cohort Study Index** in *JAMIA Open*.

JAMIA Open, 2025, **8(4)**, ooaf074
https://doi.org/10.1093/jamiaopen/ooaf074
Research and Applications

Research and Applications

Comparison of 2 electronic health record data extraction methods for laboratory tests used in the Veterans Aging Cohort Study Index

Christine Ramsey-Hardy, MS, PhD*[,1,2,3], Melissa Skanderson, MSW[2], Janet P. Tate, ScD[2,3], Amy C. Justice, PhD, MD[2,3], Vincent C. Marconi, MD[4,5,6], Charles Alcorn,[7] Ronald G. Hauser, MD[2,8], Amy Anderson-Mellies, MPH[7,8,9], Kathleen A. McGinnis, MS, DrPH*[,2]

[1]Department of Nutritional Sciences, University of Texas at Austin, Austin, TX 78712, United States, [2]VA Connecticut Healthcare System, West Haven, CT 06516, United States, [3]Department of Internal Medicine, Yale School of Medicine, New Haven, CT 06510, United States, [4]Division of Infectious Diseases, Emory University School of Medicine, Atlanta, GA 30322, United States, [5]Department of Global Health, Rollins School of Public Health, Atlanta, GA 30322, United States, [6]Atlanta Veterans Affairs Medical Center, Atlanta, GA 30033, United States, [7]Division of Cardiovascular Medicine, Vanderbilt University Medical Center, Nashville, TN 37232, United States, [8]Vanderbilt Center for Population Science and Randomized Clinical Trials, Vanderbilt University Medical Center, Nashville, TN 37322, United States, [9]VA Tennessee Valley Healthcare System, Nashville, TN 37212, United States

*Corresponding authors: Kathleen A. McGinnis, MS, DrPH, VA Connecticut Healthcare System, 11ACSL-G, 950 Campbell Avenue, West Haven, CT 06516, United States (kathleen.mcginnis3@va.gov) and Christine Ramsey-Hardy, MS, PhD, VA Connecticut Healthcare System, 11ACSL-G, 950 Campbell Avenue, West Haven, CT 06516, United States (Christine.Hardy@va.gov)

Abstract

Objective: To compare Observational Medical Outcomes Partnership (OMOP) Logical Observation Identifiers Names and Codes (LOINC) and Veterans Aging Cohort Study (VACS) methods for extracting laboratory chemistry data from Veterans Health Administration (VA) electronic health records (EHR).

Materials and Methods: Laboratory chemistry test results for 16 laboratory tests commonly assess in Veterans in VACS HIV ($N = 143\,830$) followed in the VA 2015-2019 were extracted from the EHR and compared using 2 different data extraction approaches.

Results: The LOINC approach captured laboratory results from all 130 VA stations for all 16 labs. The VACS approach captured laboratory results from 128 to130 stations. Both approaches yielded laboratory results for a patient on a given date for 97% or more of the observations for 10 of the tests, 94%-97% for 5 of the tests, and 89% for 1 test. The percentage of exact matches on the value of the test result exceeded 99% for 10 of the laboratory tests and 92% for all other laboratory tests.

Discussion: Both approaches resulted in extraction of similar amounts of data in terms of individual patients, VA stations and total observations for all 16 tests. Both approaches yielded high agreement on test results in terms of identical values and correlation of test results for all tests.

Conclusion: The high level of agreement between VACS and LOINC approaches in this study demonstrate the favorable use of the LOINC approach for extracting laboratory results for most tests due to the ease and efficiency of this approach without compromising validity.

Lay Summary

This study compared Observational Medical Outcomes Partnership (OMOP) Logical Observation Identifiers Names and Codes (LOINC) and Veterans Aging Cohort Study (VACS) methods for extracting test results for 16 common laboratory tests (eg, cholesterol, glucose) from Veterans Health Administration (VA) electronic health records (EHR). Both approaches resulted in extraction of similar amounts of data in terms of individual patients, VA stations and total observations for all 16 tests. Both approaches yielded high agreement on test results in terms of identical values and correlation of test results for all tests. The high level of agreement between VACS and LOINC approaches in this study demonstrates the favorable use of the LOINC approach for extracting laboratory results for most tests due to the ease and efficiency of this approach without compromising validity.

Key words: study; Veterans Aging Cohort Study; Veterans Aging Cohort OMOP; laboratory test results.

OHDSI Shoutouts!

Congratulations to the team of **Yilu Fang, Gongbo Zhang, Fangyi Chen, George Hripcsak, Yifan Peng, Patrick Ryan, and Chunhua Weng** on the publication of **CLEAR: A vision to support clinical evidence lifecycle with continuous learning** in the *Journal of Biomedical Informatics.*

ELSEVIER

Journal of Biomedical Informatics

Available online 29 July 2025, 104884

In Press, Journal Pre-proof What's this?

Special Communication

CLEAR: A vision to support clinical evidence lifecycle with continuous learning

Yilu Fang [a], Gongbo Zhang [a], Fangyi Chen [a], George Hripcsak [a 1], Yifan Peng [b 1], Patrick Ryan [a c 1], Chunhua Weng [a]

Show more

Add to Mendeley Share Cite

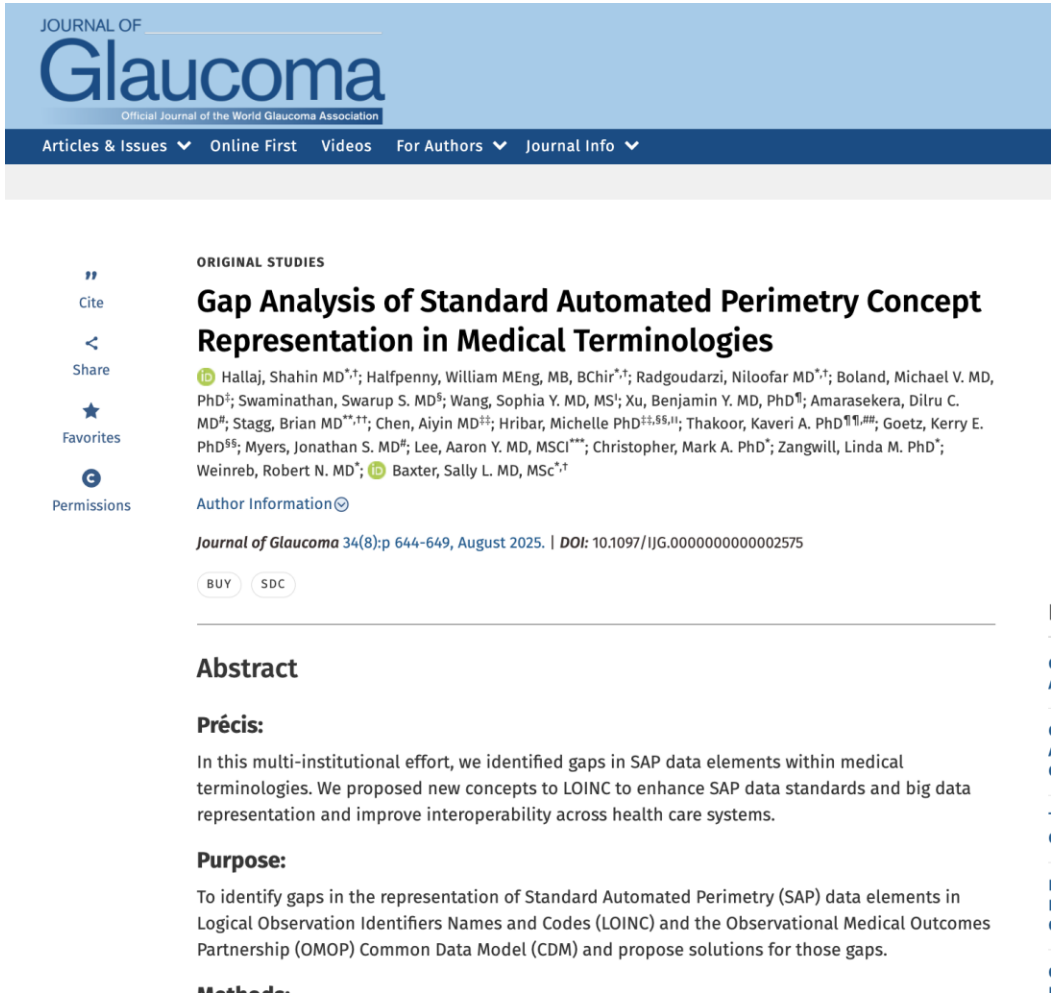
https://doi.org/10.1016/j.jbi.2025.104884 Get rights and content

Abstract

Human knowledge of diseases, treatments, and prevention techniques is constantly evolving. The generation of clinical evidence using randomized controlled trials on human subjects occurs notably slowly and inefficiently. The Learning Health System (LHS) has been proposed to facilitate the continuous improvement of individual and population health through a cycle of knowledge, practice, and data. However, the gap between the demand for high-quality evidence to support clinical decisions and the available evidence continues to enlarge. While the current LHS vision articulates the integration of Real-World Data (RWD), the rapid generation of RWD often outpaces the rate of effective evidence synthesis and implementation. Considering this, we propose a new framework that more effectively leverages RWD to support the entire

OHDSI Shoutouts!

Congratulations to the team of **Shahin Hallaj, William Halfpenny, Niloofar Radgoudarzi, Michael Boland, Swarup Swaminathan, Sophia Wang, Benjamin Xu, Dilru Amarasekera, Brian Stagg, Aiyin Chen, Michelle Hribar, Kaveri Thakoor, Kerry Goetz, Jonathan Myers, Aaron Lee, Mark Christopher, Linda Zangwill, Robert Weinreb, and Sally Baxter** on the publication of **Gap Analysis of Standard Automated Perimetry Concept Representation in Medical Terminologies** in the *Journal of Glaucoma*.



JOURNAL OF Glaucoma
Official Journal of the World Glaucoma Association

Articles & Issues | Online First | Videos | For Authors | Journal Info

Cite
Share
Favorites
Permissions

ORIGINAL STUDIES

Gap Analysis of Standard Automated Perimetry Concept Representation in Medical Terminologies

Hallaj, Shahin MD*,†; Halfpenny, William MEng, MB, BChir*,†; Radgoudarzi, Niloofar MD*,†; Boland, Michael V. MD, PhD‡; Swaminathan, Swarup S. MD§; Wang, Sophia Y. MD, MS‖; Xu, Benjamin Y. MD, PhD¶; Amarasekera, Dilru C. MD#; Stagg, Brian MD**,††; Chen, Aiyin MD‡‡; Hribar, Michelle PhD‡‡,§§,‖‖; Thakoor, Kaveri A. PhD¶¶,##; Goetz, Kerry E. PhD§§; Myers, Jonathan S. MD#; Lee, Aaron Y. MD, MSCI***; Christopher, Mark A. PhD*; Zangwill, Linda M. PhD*; Weinreb, Robert N. MD*; Baxter, Sally L. MD, MSc*,†

Author Information

Journal of Glaucoma 34(8):p 644-649, August 2025. | **DOI:** 10.1097/IJG.0000000000002575

BUY | SDC

Abstract

Précis:

In this multi-institutional effort, we identified gaps in SAP data elements within medical terminologies. We proposed new concepts to LOINC to enhance SAP data standards and big data representation and improve interoperability across health care systems.

Purpose:

To identify gaps in the representation of Standard Automated Perimetry (SAP) data elements in Logical Observation Identifiers Names and Codes (LOINC) and the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) and propose solutions for those gaps.

OHDSI Shoutouts!

Congratulations to the team of **Yoshihiro Aoyagi, Masahiro Baba, Suzue Terao, Yuya Ikeda, Keiichi Nomura, and Akihiro Sato** on the publication of **Feasibility of Converting EMR Data to OMOP CDM and Utilizing OHDSI Analysis Tools in Japan** in *Studies in Health Technology and Information Volume 329: MEDINFO 2025 — Healthcare Smart × Medicine Deep.*

MEDINFO 2025 — Healthcare Smart × Medicine Deep
M.S. Househ et al. (Eds.)

doi:10.3233/SHTI251292

Feasibility of Converting EMR Data to OMOP CDM and Utilizing OHDSI Analysis Tools in Japan

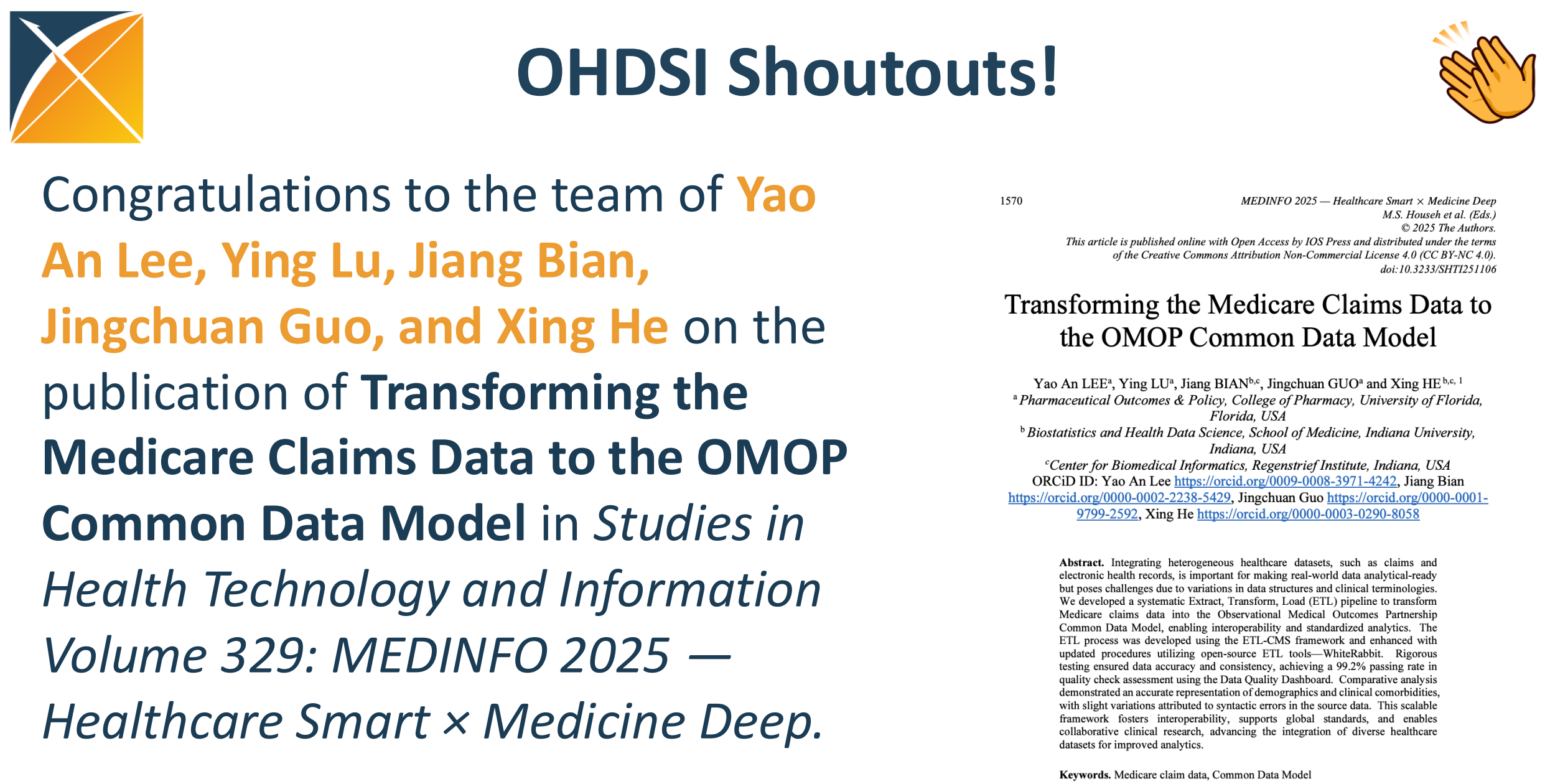
Yoshihiro AOYAGI[a,b], Masahiro BABA[b], Suzue TERAO[b], Yuya IKEDA[b], Keiichi NOMURA[a], Akihiro SATO[b]
a *Department of Medical Information, National Cancer Center Hospital East*
b *Clinical Research Management Division, Clinical Research Support Office, National Cancer Center Hospital East*
ORCiD ID: Yoshihiro Aoyagi https://orcid.org/0000-0003-1496-1112

Abstract. The study investigates the feasibility of converting electronic medical record (EMR) data to the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM) and utilizing Observational Health Data Sciences and Informatics (OHDSI) analysis tools in Japan. Data from 8,387 breast cancer patients at the National Cancer Centre Hospital East were extracted and transformed into OMOP CDM. OHDSI software and standard vocabulary were installed, and the transformed data were displayed on ATLAS for analysis. The quality of the converted datasets was assessed using the Data Quality Dashboard developed by the OHDSI community. Mapping to OHDSI standard vocabulary was performed for 7,259 terms, with SNOMED-CT, RxNorm, and LOINC frequently used. The conversion program was applied to all cases, but there were challenges, such as race data and anticancer drug specification selection. ATLAS displayed reports on the observation period and drug use. The study demonstrated the feasibility of converting EMR data to OMOP CDM and utilizing ATLAS in Japan. However, there were challenges in software installation, vocabulary mapping, and data quality evaluation.

Keywords. EMR, OMOP CDM, OHDSI, SNOMED-CT, RxNorm, LOINC

OHDSI Shoutouts!

Congratulations to the team of **Yao An Lee, Ying Lu, Jiang Bian, Jingchuan Guo, and Xing He** on the publication of **Transforming the Medicare Claims Data to the OMOP Common Data Model** in *Studies in Health Technology and Information Volume 329: MEDINFO 2025 — Healthcare Smart × Medicine Deep.*



1570

MEDINFO 2025 — Healthcare Smart × Medicine Deep
M.S. Househ et al. (Eds.)

doi:10.3233/SHTI251106

Transforming the Medicare Claims Data to the OMOP Common Data Model

Yao An LEE[a], Ying LU[a], Jiang BIAN[b,c], Jingchuan GUO[a] and Xing HE[b,c, 1]
[a] *Pharmaceutical Outcomes & Policy, College of Pharmacy, University of Florida, Florida, USA*
[b] *Biostatistics and Health Data Science, School of Medicine, Indiana University, Indiana, USA*
[c] *Center for Biomedical Informatics, Regenstrief Institute, Indiana, USA*
ORCiD ID: Yao An Lee https://orcid.org/0009-0008-3971-4242, Jiang Bian https://orcid.org/0000-0002-2238-5429, Jingchuan Guo https://orcid.org/0000-0001-9799-2592, Xing He https://orcid.org/0000-0003-0290-8058

Abstract. Integrating heterogeneous healthcare datasets, such as claims and electronic health records, is important for making real-world data analytical-ready but poses challenges due to variations in data structures and clinical terminologies. We developed a systematic Extract, Transform, Load (ETL) pipeline to transform Medicare claims data into the Observational Medical Outcomes Partnership Common Data Model, enabling interoperability and standardized analytics. The ETL process was developed using the ETL-CMS framework and enhanced with updated procedures utilizing open-source ETL tools—WhiteRabbit. Rigorous testing ensured data accuracy and consistency, achieving a 99.2% passing rate in quality check assessment using the Data Quality Dashboard. Comparative analysis demonstrated an accurate representation of demographics and clinical comorbidities, with slight variations attributed to syntactic errors in the source data. This scalable framework fosters interoperability, supports global standards, and enables collaborative clinical research, advancing the integration of diverse healthcare datasets for improved analytics.

Keywords. Medicare claim data, Common Data Model

OHDSI Shoutouts!

Congratulations to the team of **Yan Hu, Na Hong, Yiming Li, Xueqing Peng, Yong Chen, and Hua Xu** on the publication of **PheCatcher: Leveraging LLM-Generated Synthetic Data for Automated Phenotype Definition Extraction from Biomedical Literature** in *Studies in Health Technology and Information Volume 329: MEDINFO 2025 — Healthcare Smart × Medicine Deep.*

MEDINFO 2025 — Healthcare Smart × Medicine Deep
M.S. Househ et al. (Eds.)

doi:10.3233/SHTI250934

PheCatcher: Leveraging LLM-Generated Synthetic Data for Automated Phenotype Definition Extraction from Biomedical Literature

Yan HU[a], Na HONG[b], Yiming LI[a], Xueqing PENG[b], Yong CHEN[c], Hua XU[b,1]
[a] *McWilliams School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, USA*
[b] *Department of Biomedical Informatics and Data Science, Yale School of Medicine, Yale University, New Haven, CT, USA*
[c] *Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA*
ORCiD ID: Yan Hu https://orcid.org/0009-0008-2413-5918, Na Hong https://orcid.org/0000-0001-6798-1761, Yiming Li https://orcid.org/0009-0009-8784-1745, Xueqing Peng https://orcid.org/0009-0004-1484-0622, Yong Chen https://orcid.org/0000-0003-0835-0788, Hua Xu https://orcid.org/0000-0002-5274-4672

Abstract. Phenotype definitions are crucial for the progression of precision and personalized medicine. Although phenotype knowledge bases such as PheKB and the OHDSI library are available, they rely heavily on manual input. This study introduces PheCatcher, an automated pipeline that integrates BiomedBERT-based Named Entity Recognition (NER) and Relation Extraction (RE) to extract phenotypes and standardized codes from biomedical literature. To complement human annotation, GPT-4 was utilized to generate synthetic data, which improved model performance. The NER model's F1 score for "phenotype" entities increased from 0.616 to 0.800, and the RE model achieved an F1 score of 0.901. The application of the pipeline to the PubMed Central (PMC) articles resulted in the extraction of 173,283 phenotype definitions, which are now publicly accessible. Our study underscores the potential of synthetic data for information extraction (IE) and offers the first evidence of the feasibility of leveraging synthetic data to build a complete IE system.

Keywords. Phenotype Definition, Information Extraction, Synthetic Data

OHDSI Shoutouts!

Congratulations to the team of **Clair Blacketer, Martijn Schuemie, Maxim Moinat, Erica Voss, Montse Camprubi, Peter Rijnbeek, and Patrick Ryan** on the publication of **Advancing Real-World Evidence Through a Federated Health Data Network (EHDEN): Descriptive Study** in the *Journal of Medical Internet Research*.

JOURNAL OF MEDICAL INTERNET RESEARCH Blacketer et al

Original Paper

Advancing Real-World Evidence Through a Federated Health Data Network (EHDEN): Descriptive Study

Clair Blacketer[1,2,3], MPH; Martijn J Schuemie[1,3,4], PhD, MS; Maxim Moinat[1,2], MS; Erica A Voss[1,3], PhD, MPH; Montse Camprubi[1,5], MS; Peter R Rijnbeek[1,2*], PhD, MS; Patrick B Ryan[1,3,6*], PhD, MS

[1]OHDSI Collaborators, New York, NY, United States
[2]Department of Medical Informatics, Erasmus MC, Rotterdam, The Netherlands
[3]Johnson & Johnson (United States), Raritan, NJ, United States
[4]Department of Biostatistics, University of California, Los Angeles, Los Angeles, CA, United States
[5]Synapse (Spain), Madrid, Spain
[6]Department of Biomedical Informatics, Columbia University, New York, NY, United States
*these authors contributed equally

Corresponding Author:

Clair Blacketer, MPH
Johnson & Johnson (United States)
920 US Route 202
Raritan, NJ 08869
United States
Phone: 1 7573345788
Email: mblacke@its.jnj.com

Abstract

Background: Real-world data (RWD) are increasingly used in health research and regulatory decision-making to assess the effectiveness, safety, and value of interventions in routine care. However, the heterogeneity of European health care systems, data capture methods, coding standards, and governance structures poses challenges for generating robust and reproducible real-world evidence. The European Health Data & Evidence Network (EHDEN) was established to address these challenges by building a large-scale federated data infrastructure that harmonizes RWD across Europe.

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
Tuesday	12 pm	Generative AI and Analytics
Tuesday	3 pm	Oncology Outreach/Research Subgroup
Wednesday	9 am	Prediction
Wednesday	7 pm	Eyecare and Vision Research
Thursday	9:30 am	Network Data Quality
Thursday	10 am	Rare Diseases
Thursday	10:30 am	Evidence Network
Friday	10 am	GIS-Geographic Information System
Friday	10:30 am	Open-Source Community
Friday	11:30 am	Steering
Monday	9 am	Vaccine Vocabulary
Monday	10 am	Africa Chapter
Monday	10 am	Getting Started (Health Systems Subgroup)
Monday	11 am	Data Bricks User Group
Monday	2 pm	Electronic Animal Health Records
Tuesday	9:30 am	CDM Survey Subgroup

Titan Award Nominations Are Open

The Titan Awards have been handed out annually since 2018 to recognize OHDSI collaborators (or collaborating institutions) for their contributions towards OHDSI's mission.

Nominations for the 2025 Titan Awards are now open. **Please complete your nominations by our Sept. 9 (8 pm ET) deadline!**

ohdsi.org/titan-awards

August Newsletter is Available

OHDSI

The Journey Newsletter (August 2025)

The second half of 2025 will feature OHDSI symposia around the world, and we got off to a fantastic start with the Europe Symposium, held in Hasselt, Belgium, last month. Videos, posters and more are now available from this event. We also look ahead to events in the USA, Uganda and China. There is also an opportunity to be part of the FHIR to OMOP conversation, which you can learn more about in this newsletter. #JoinTheJourney

Podcast: Europe Symposium Review, Revelations from OHDSI Summer School

In the August 2025 On The Journey podcast, Patrick Ryan and Craig Sachson reflect on the collaborative nature of the recent Europe Symposium, as well as some of the scientific output from the event, including a panel around the European Health Data Space. They also discuss the first Summer School in Observational Health Data Science & Informatics, AI, and Real World Evidence, hosted by OHDSI faculty at Columbia University. (if video does not appear, click *View this email in your browser*)

Community Updates

Where Have We Been?

- Thank you to everybody in the community who submitted brief reports for the 2025 Global Symposium. We had nearly 150 submissions for our OHDSI2025 Collaborator Showcase, which will be held Oct. 8 as part of the Global Symposium (more information below). All collaborators will receive notifications about their submissions in mid-August.
- The 6th European OHDSI Symposium took place July 5-7 at the historic Old Prison building of Hasselt University in Belgium, and all resources can now be found on the event homepage. The program included two full days of tutorials and workshops, followed by a main symposium featuring presentations, a panel discussion, poster and lightning talks, and opportunities to network with current or future collaborators. Over 300 researchers, data scientists, healthcare professionals, and policymakers attended to explore developments in OMOP-CDM, collaborative health data projects, and regional efforts aligned with the emerging European Health Data Space.
- The first Summer School in Observational Health Data Science & Informatics, AI, and Real World Evidence was held in July at the Department of Biomedical Informatics at Columbia University. OHDSI faculty (Patrick Ryan, George Hripcsak, Anna Ostropolets and Karthik Natarajan) led this week-long event which provided health professionals, researchers, and industry practitioners with an immersive, hands-on training to working with real-world health data and generating real-world evidence.

Where Are We Now?

- **On August 12, we will host our annual "Meet 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. For scheduling purposes, please note your interest in taking part in this brief survey.
- During the July 29 community call, Davera Gabriel and Jean Duteau provided an update on the FHIR-to-OMOP Implementation Guide. They discussed what is in the guide and how you can provide feedback by joining the formal review before the ballot pool closes on August 7. Learn more about how to provide feedback later in this newsletter.
- Curious about the latest developments and innovations around large-language models in OHDSI? The August 26 community call will feature some of the leaders in our community to highlight some of the recent work. The link to join upcoming community calls, as well as videos and slides from past presentations, can be found here.

Hasselt Is Home For 2025 Europe Symposium, Which Highlights Advancements, Collaboration

The 2025 OHDSI Europe Symposium took place July 5-7 at Hasselt University in Belgium, marking the first time the event was held outside of Rotterdam. Hosted in the historic Old Prison building, the symposium attracted more than 350 participants from across Europe and beyond, including researchers, clinicians, policymakers, and data scientists. Thank you to Liesbet Peeters, Ilse Vermeulen and the entire Hasselt team for your amazing leadership.

July Publications

Haber R, Webster-Clark M, Pratt N, Barclay N, Li X, Maro JC, Platt RW, Prieto-Alhambra D, Filion KB. Core Concepts in Pharmacoepidemiology: Multi-Database Distributed Data Networks. Pharmacoepidemiol Drug Saf. 2025 Jul;34(7):e70177. doi: 10.1002/pds.70177. PMID: 40619597; PMCID: PMC12230205.

Spotnitz M, Giannini J, Clark E, Ostchega Y, Litwin TR, Goff SL, Berman L. Assessing the Data Quality Dimensions of Surgical Oncology Cohorts in the *All of Us* Research Program. JCO Clin Cancer Inform. 2025 Jul;9:e2500078. doi: 10.1200/CCI-25-00078. Epub 2025 Jul 8. PMID: 40627823; PMCID: PMC12240465.

Trofymenko M, Korchmar E, Kaduk D, Vikhrak M, Khilchevskyi B, Nesmiian T, Talapova P, Ved M, Ageeva I. Jackalope Plus tool for post-coordination, ontology development, and precise mapping in observational health studies. Sci Rep. 2025 Jul 2;15(1):23674. doi: 10.1038/s41598-025-04046-9. PMID: 40603892; PMCID: PMC12223214.

Mendez N, Alonso E, Espejo MD, López M, Alonso-Arce M, Rojo E, Garin-Muga A. Harmonizing Electronic Health Records to the OMOP Common Data Model: A Case Study on Surgical Complications. Stud Health Technol Inform. 2025 Jun 26;328:261-265. doi: 10.3233/SHTI250715. PMID: 40588922.

Park WY, Sippel Schmidt T, Salvador G, O'Donnell K, Genereaux B, Jeon K, You SC, Dewey BE, Nagy P; Alzheimer's Disease Neuroimaging Initiative. Breaking data silos: incorporating the DICOM imaging standard into the OMOP CDM to enable multimodal research. J Am Med Inform Assoc. 2025 Jul 18:ocaf091. doi: 10.1093/jamia/ocaf091. Epub ahead of print. PMID: 40680297.

Tong J, Reps JM, Luo C, Lu Y, Li L, Ramirez-Anguita JM, Brand MT, DuVall SL, Falconer T, Fuentes AM, He X, Matheny ME, Mayer MA, Patel BK, Simon KR, Suchard MA, Tang G, Viernes B, Williams RD, van Zandt M, Wang F, Bian J, Zhou J, Asch DA, Chen Y. Unlocking efficiency in real-world collaborative studies: a multi-site international study with one-shot lossless GLMM algorithm. NPJ Digit Med. 2025 Jul 19;8(1):457. doi: 10.1038/s41746-025-01846-1. PMID: 40681761; PMCID: PMC12274576.

Trezena S, Martelli DRB, Bonan PRF, Graner E, Sobrinho LMF, Alawi F, Coletta RD, Martelli-Júnior H. Knowledge and attitudes about rare genetic diseases among practitioners of oral medicine/pathology in Brazil: a cross-sectional study. Front Oral Health. 2025 Jul 7;6:1573355. doi: 10.3389/froh.2025.1573355. PMID: 40692699; PMCID: PMC12277291.

August Newsletter is Available

Spotlight: Priya Desai

Priya Desai is a physicist by training and a data scientist at heart. As the Biomedical Informatics R&D lead of the Research IT team at Stanford University, she is passionate about the mission of building the next generation of products to leverage healthcare research and analytics.

In the latest edition of the Collaborator Spotlight, Desai discusses her career journey, the STARR-OMOP initiative, why she believes medicine mirrors astrophysics, watching her daughter enter the health data science research field, and more.

ohdsi.org/spotlight-priya-desai

SURVEY DATA AND THE OMOP CDM: LANDSCAPE ASSESSMENT

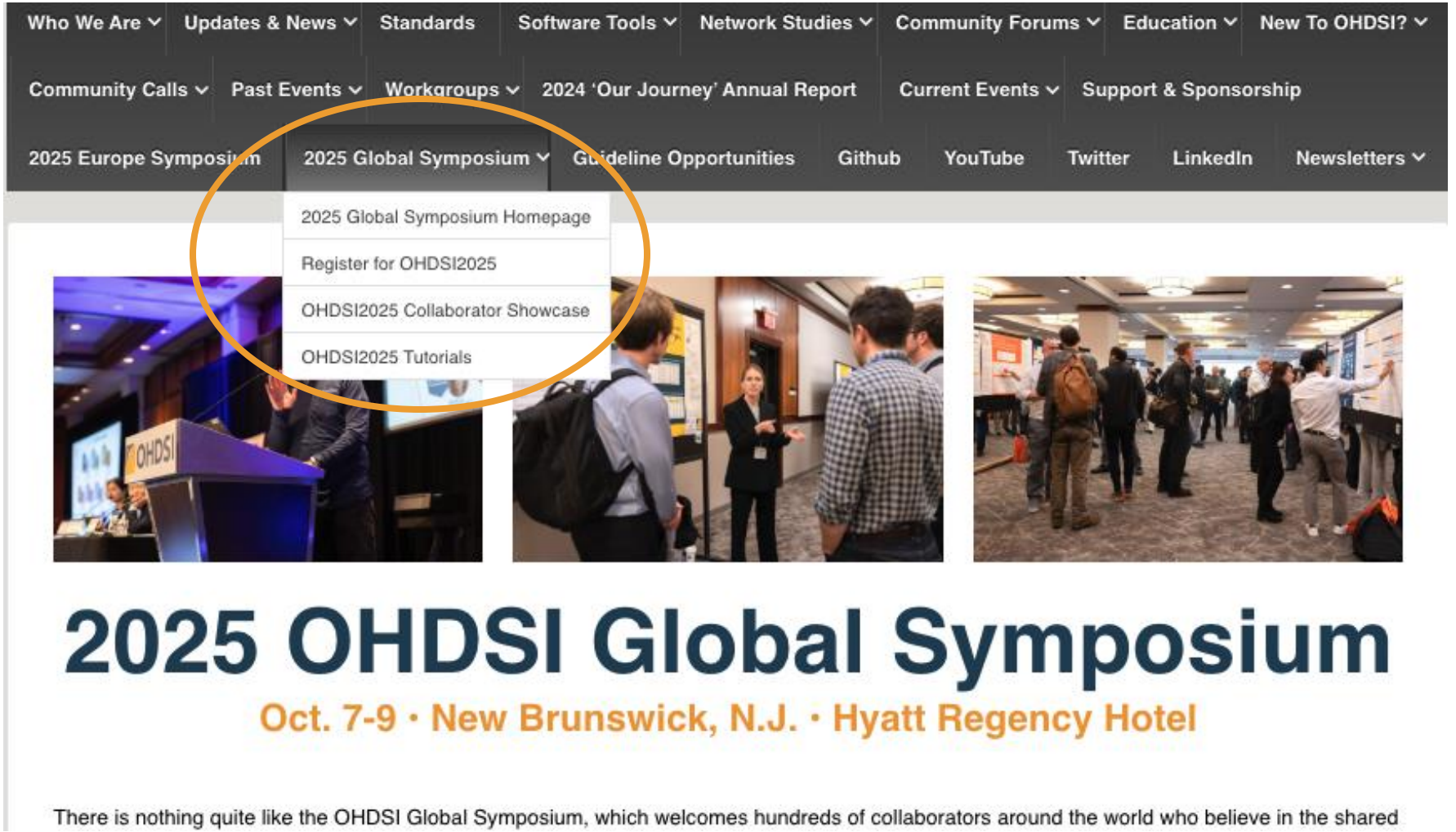
- https://forms.gle/f18ufspAFT3jSYrk6
- **Open through August 31, 2025**

Nicole Gerlanc, PhD

Data Analyst Lead, Connect Study
Trans-Divisional Research Program
Division of Cancer Epidemiology and Genetics
National Cancer Institute

Email: nicole.gerlanc@nih.gov
Wiki: https://github.com/OHDSI/CdmSurveySubWg/wiki

Global Symposium: Oct. 7-9



ohdsi.org/ohdsi2025

Africa Symposium: Nov. 10-12

Who We Are | Updates & News | Standards | Software Tools | Network Studies | Community Forums | Education | New To OHDSI?
Community Calls | Past Events | Workgroups | 2024 'Our Journey' Annual Report | Current Events | Support & Sponsorship
2025 Europe Symposium | 2025 Global Symposium | 2025 Africa Symposium | Github | YouTube | Twitter | LinkedIn | Newsletters

Join Us At The Inaugural OHDSI Africa Symposium

Nov. 10-12, 2025 · Joint Clinical Research Centre (JCRC) & Mestil Hotel Kampala




The inaugural OHDSI Africa Symposium will be held in Kampala at the Joint Clinical Research Centre (JCRC) and Mestil Hotel. Our community is delighted to introduce a new face-to-face opportunity in Africa, where OHDSI is growing at an exciting pace. We hope you will join us for this historical moment.

The first OHDSI Africa symposium will be hosted by JCRC and will begin with a dedicated one-day training course at JCRC, followed by a two-day main conference at Mestil hotel. Below are some important dates for you to save to your calendar:

Collaborator Showcase
- Submissions deadline: September 10
- Submissions review: September 11-30
- Notification of acceptance: October 5

Symposium
- Tutorial: November 10 at JCRC
- Main conference: November 11-12 at Mestil Hotel

ohdsi.org/africa2025

Africa Symposium: Nov. 10-12

The 2025 OHDSI APAC Symposium will be held in Shanghai, China at the Shanghai Jiao Tong University featuring a 1-day tutorial and a 1-day main conference. Here are some important dates for you to save to your calendar:

Collaborator Showcase

- Submissions deadline: September 7
- Submissions review: September 8 – October 9
- Notification of acceptance: October 17

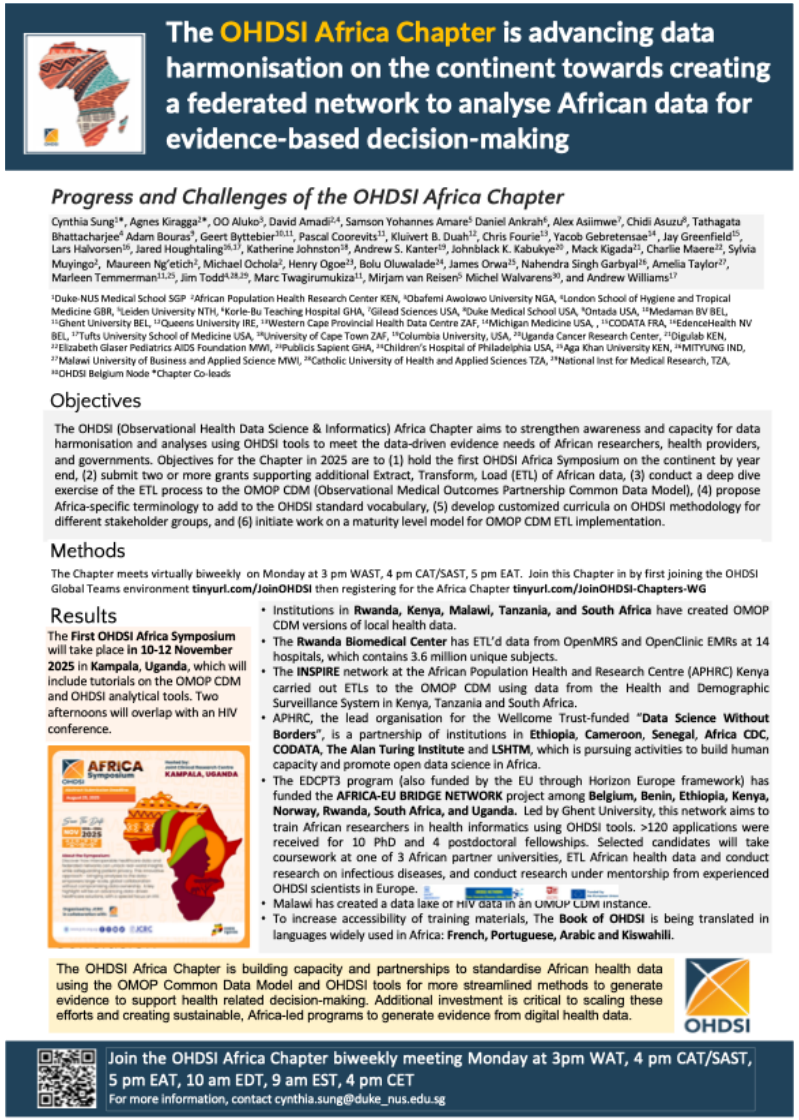
ohdsi.org/africa2025

#OHDSISocialShowcase This Week

Monday

Progress and Challenges of the OHDSI Africa Chapter

(Cynthia Sung, Agnes Kiragga, OO Aluko, David Amadi, Samson Yohannes Amare, Daniel Ankrah, Alex Asiimwe, Chidi Asuzu, Tathagata Bhattacharjee, Adam Bouras, Geert Byttebie, Pascal Coorevits, Kluivert B. Duah, Chris Fourie, Yacob Gebretensae, Jay Greenfield, Lars Halvorsen, Jared Houghtaling, Katherine Johnston, Andrew S. Kanter, Johnblack Kabukye, Mack Kigada, Charlie Maere, Sylvia Muyingo, Maureen Ng'etich, Michael Ochola, Henry Ogoe, Bolu Oluwalade, James Orwa, Nahendra Singh Garbyal, Amelia Taylor, Marleen Temmerman, Jim Todd, Marc Twagirumukiza, Mirjam van Reisen, Michel Walravens, Andrew Williams)

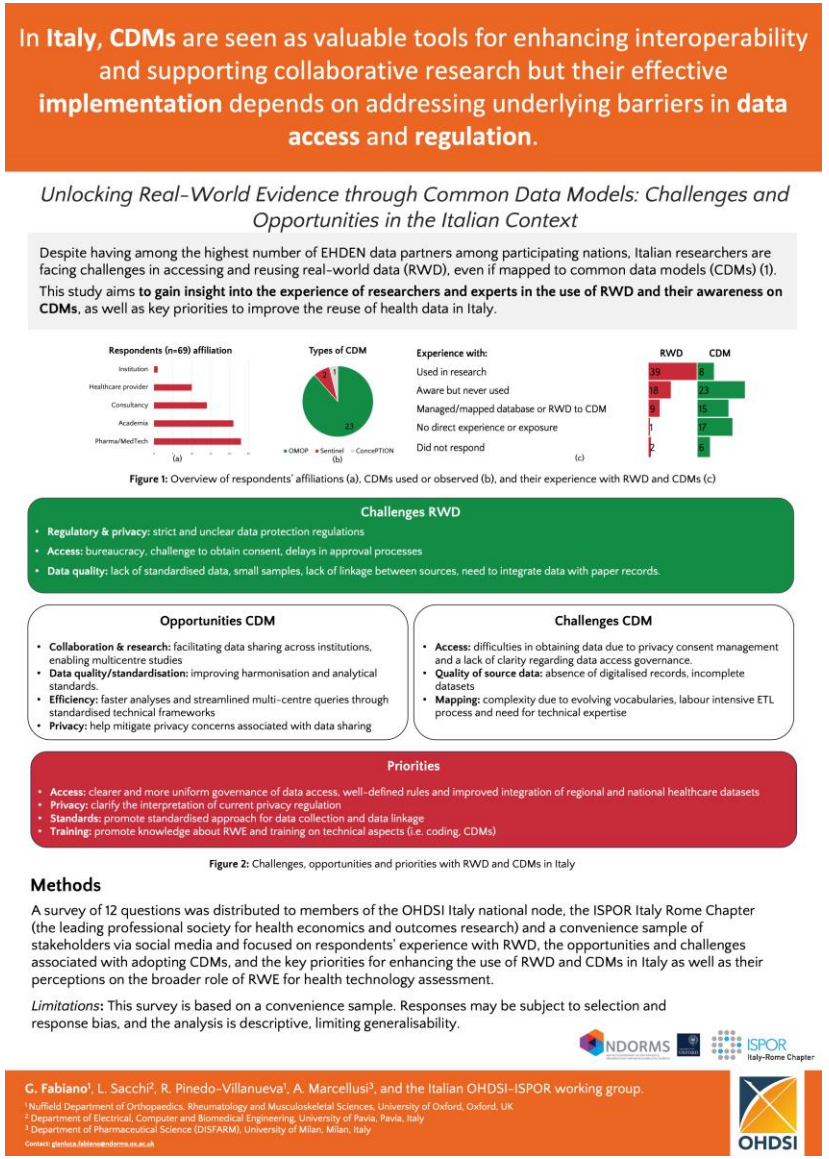


#OHDSISocialShowcase This Week

Tuesday

Unlocking Real-World Evidence through Common Data Models: Challenges and Opportunities in the Italian Context

(Gianluca Fabiano, Lucia Sacchi, Rafael Pinedo-Villanueva, Andrea Marcellusi)



In **Italy**, **CDMs** are seen as valuable tools for enhancing interoperability and supporting collaborative research but their effective **implementation** depends on addressing underlying barriers in **data access** and **regulation**.

Unlocking Real-World Evidence through Common Data Models: Challenges and Opportunities in the Italian Context

Despite having among the highest number of EHDEN data partners among participating nations, Italian researchers are facing challenges in accessing and reusing real-world data (RWD), even if mapped to common data models (CDMs) (1).

This study aims **to gain insight into the experience of researchers and experts in the use of RWD and their awareness on CDMs**, as well as key priorities to improve the reuse of health data in Italy.

Figure 1: Overview of respondents' affiliations (a), CDMs used or observed (b), and their experience with RWD and CDMs (c)

Challenges RWD

- **Regulatory & privacy:** strict and unclear data protection regulations
- **Access:** bureaucracy, challenge to obtain consent, delays in approval processes
- **Data quality:** lack of standardised data, small samples, lack of linkage between sources, need to integrate data with paper records.

Opportunities CDM

- **Collaboration & research:** facilitating data sharing across institutions, enabling multicentre studies
- **Data quality/standardisation:** improving harmonisation and analytical standards.
- **Efficiency:** faster analyses and streamlined multi-centre queries through standardised technical frameworks
- **Privacy:** help mitigate privacy concerns associated with data sharing

Challenges CDM

- **Access:** difficulties in obtaining data due to privacy consent management and a lack of clarity regarding data access governance.
- **Quality of source data:** absence of digitalised records, incomplete datasets
- **Mapping:** complexity due to evolving vocabularies, labour intensive ETL process and need for technical expertise

Priorities

- **Access:** clearer and more uniform governance of data access, well-defined rules and improved integration of regional and national healthcare datasets
- **Privacy:** clarify the interpretation of current privacy regulation
- **Standards:** promote standardised approach for data collection and data linkage
- **Training:** promote knowledge about RWE and training on technical aspects (i.e. coding, CDMs)

Figure 2: Challenges, opportunities and priorities with RWD and CDMs in Italy

Methods

A survey of 12 questions was distributed to members of the OHDSI Italy national node, the ISPOR Italy Rome Chapter (the leading professional society for health economics and outcomes research) and a convenience sample of stakeholders via social media and focused on respondents' experience with RWD, the opportunities and challenges associated with adopting CDMs, and the key priorities for enhancing the use of RWD and CDMs in Italy as well as their perceptions on the broader role of RWE for health technology assessment.

Limitations: This survey is based on a convenience sample. Responses may be subject to selection and response bias, and the analysis is descriptive, limiting generalisability.

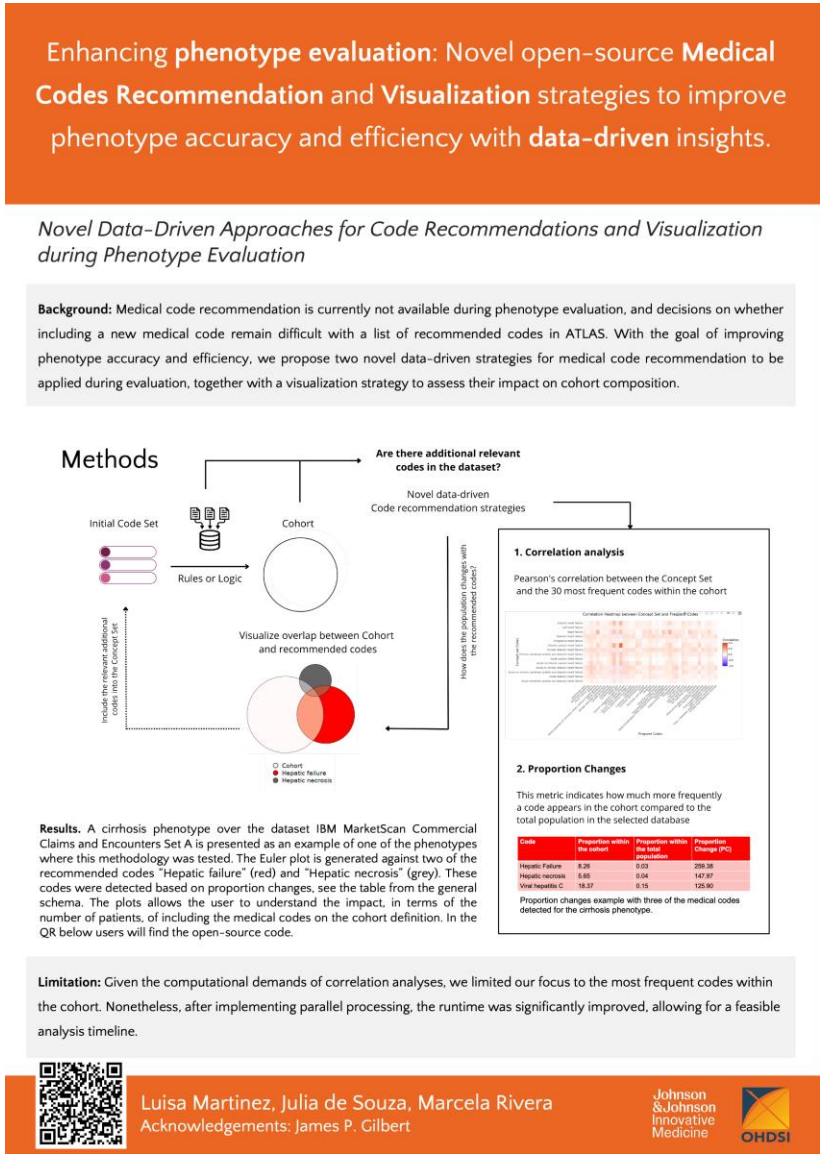
G. Fabiano[1], L. Sacchi[2], R. Pinedo-Villanueva[1], A. Marcellusi[3], and the Italian OHDSI-ISPOR working group.
[1] Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
[2] Department of Electrical, Computer and Biomedical Engineering, University of Pavia, Pavia, Italy
[3] Department of Pharmaceutical Science (DISFARM), University of Milan, Milan, Italy
Contact: gianluca.fabiano@ndorms.ox.ac.uk

#OHDSISocialShowcase This Week

Wednesday

Novel Data-Driven Approaches for Code Recommendations and Visualization during Phenotype Evaluation

(Luisa Martinez, Julia Souza, Marcela Rivera)



#OHDSISocialShowcase This Week

Thursday

“Studyathons” can catalyse rapid progress in OMOP-enabled oncology research: insights from DigiONE I3

(Xosé Fernández, Jie Yeap, James Brash, Lauren Revie, James Anderson)

#OHDSISocialShowcase This Week

Friday

The REACT-TIA Study-a-thon: Conducting a rapid multi-database cohort study in a pharmaceutical company - Lessons learned

(Elif Inan-Eroglu, Daloha Rodriguez-Molina, Khaled Abdelgawwad, Kim López-Güell, Carina Dinkel-Keuthage, Fernando Dursi, Tomasz Dyszynski, Melanie Tuchardt, Juho Immonen, Gilbert Ko, Samu Kurki, Amy Mulick, Montse Soriano Gabarro, Jason Xeni, Pareen Vora, Gianmario Candore, Katrin Manlik, David Vizcaya)

The studyathon at **Bayer** showed that a **structured, collaborative** approach effectively accelerates research execution while maintaining **scientific rigor**, proving its feasibility and value in a **pharmaceutical setting**

The REACT-TIA Studyathon: Conducting a rapid multi-database cohort study in a pharmaceutical company - Lessons learned

Figure 1: Bayer studyathon logo

Background: A studyathon is an innovative research method where experts collaborate intensively over a short period to generate reliable evidence on specific medical topics. In July 2024, Bayer conducted its first studyathon to assess Recurrent Events following a Transient Ischemic Attack (REACT-TIA), testing the feasibility of this format within a pharmaceutical setting and sharing insights to enhance evidence generation for future studies using harmonized data and analytical tools (Figure 1).

Methods: The studyathon was planned with an 8-week preparation phase, 1-week execution, and 3-week post-event phase (Figure 2), using Real-World Data sources in OMOP CDM format and tools provided by OHDSI and DARWIN EU.

- **Preparatory Phase:** A core team of seven defined the research question and coordinated internally to align with study objectives and four sub-teams were formed (Figure 3).
- **Execution Phase:** Conducted in a hybrid setting, participants committed 2-8 hours/day. Sub-teams continued coordinating activities, tracked via a dashboard. Daily lessons were captured for future improvements, culminating in "Big Medical Round Table" sessions to interpret the results and discuss findings.
- **Post-Event Phase:** Post-event responsibilities were assigned to maintain high engagement.

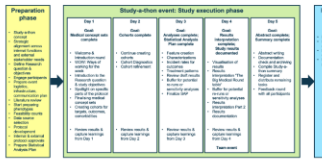


Figure 2: Agenda for the studyathon

Figure 3: Organizational setup of the studyathon team

Results: The event successfully concluded with a conference abstract presented at the **International Stroke Conference** in February 2025. The studyathon yielded key insights for the REACT-TIA study and future events.

The preparatory phase was vital, with the division into four sub-teams facilitating coordination of activities. The execution phase was intense, culminating in a draft abstract created collaboratively by 18 participants, including epidemiologists, researchers, data scientists, clinicians, and pharmacovigilance experts (Figure 4).

Focusing on a single primary objective while minimizing secondary and exploratory objectives.
Allowing for an extended timeline to address technical issues, suggesting splitting study execution activities over 2-3 weeks.
The four sub-teams helped distribute workload, establish clear deliverables, and enhance engagement; appointing a coordinator for logistics and scientific inquiries is essential.
A hybrid setup engaged more participants but challenged consistent communication.
Diverse expertise fostered valuable discussions and allowed for task sharing; organizers should ensure the right skillsets are present.
Daily lessons-learned sessions encouraged reflection and allowed to immediate implementation of suggestions.
Assigning follow-up responsibilities before the event's conclusion ensured continuity.

Figure 4: Lessons learned from the studyathon

Conclusion: Organizing the studyathon in 3 phases ensured structure, speed, and focus. This approach accelerated study execution while maintaining **scientific rigor**. Implementing studyathons in a pharmaceutical company proved **feasible and valuable** for enhancing collaboration. Future work includes creating a "**studyathon in a box**" toolkit with organizational setup, agenda samples, a deliverables dashboard, and checklists to help teams efficiently implement future studyathons.

OHDSI

Elif Inan Eroglu[1], Daloha Rodriguez-Molina[1], Khaled Abdelgawwad[1], Kim López-Güell[1], Carina Dinkel-Keuthage[1], Fernando Dursi[2], Tomasz Dyszynski[1], Melanie Tuchardt[1], Juho Immonen[3], Gilbert Ko[4], Samu Kurki[3], Amy Mulick[1], Montse Soriano Gabarro[1], Jason Xeni[5], Pareen Vora[1], Gianmario Candore[1], Katrin Manlik[1], David Vizcaya[2]

elif.inaneroglu@bayer.com

BAYER

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?

Meet The OHDSI Newcomers

Olga Endrich
Sam Fisch
Gaurav Dravida
Pranav Adiga
Mostafa Samy
Ugochukwu Onyeonoro
Natalie Provenzale
Ousmane Diop
Nhung
Somayeh Abedian
Ashley Marte
Lucia
Jiwoo Seo
Jill Vestal
Ashley Marte
Ozioma Stellamaris Uwaezu
Daniel Nsanzabandi
Vinay Pai
Kayoko Shioda
Christian Högberg
Farnoosh
Kyungseon Yohan Choi
Adeel Arif
Delaney Ding
Francis Kanyike
Charlotte Heaven
Craig Rothenberg
Saba Sohail

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