Standardized Vocabularies 2026 Winter Refresh

OHDSI Community Call
March 3, 2026 • 11 am ET

OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

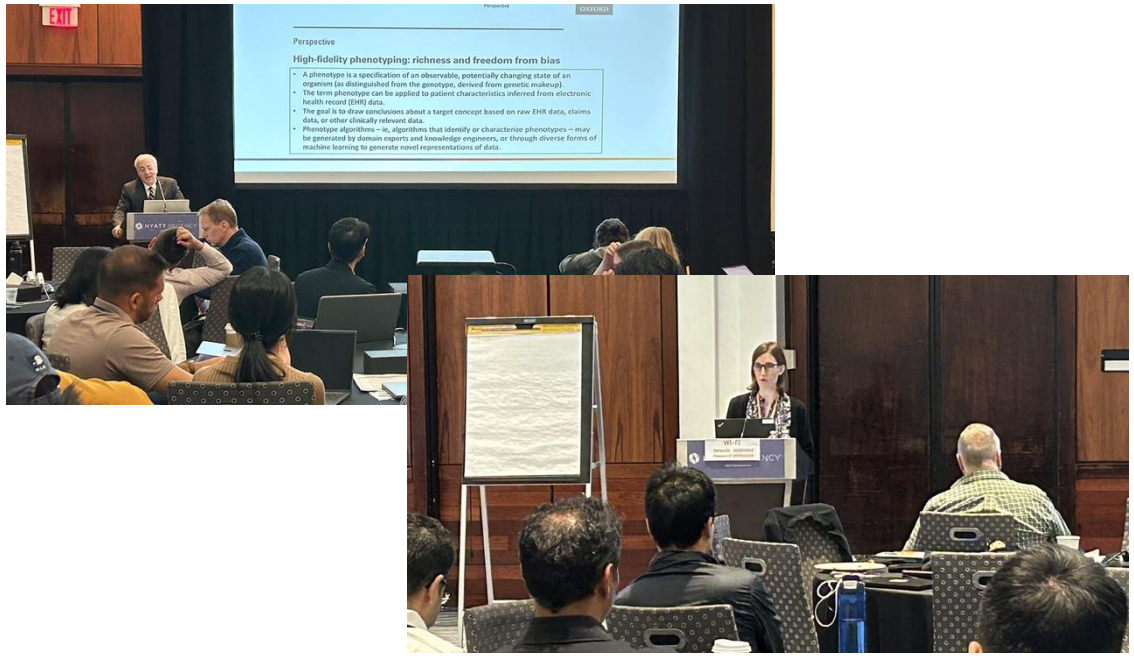
Upcoming Community Calls

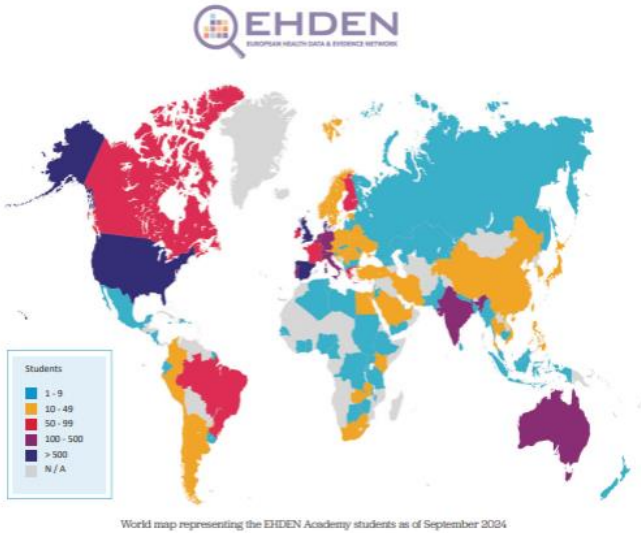
Date	Topic
Mar. 10	The Future of Education in OHDSI
Mar. 17	A Closer Look: Common Data Model, HADES
Mar. 24	OHDSI/OMOP Research Spotlight
Mar. 31	Kickoff to Phenotype April
Apr. 7	Phenotype Phebruary, Week 1
Apr. 14	Phenotype Phebruary, Week 2
Apr. 21	NO MEETING / EUROPE SYMPOSIUM
Apr. 28	Pheotype Phebruary, Week 4
May 5	Europe Symposium Review/Phenotype Phebruary Finale

March 10: The Future of Education in OHDSI

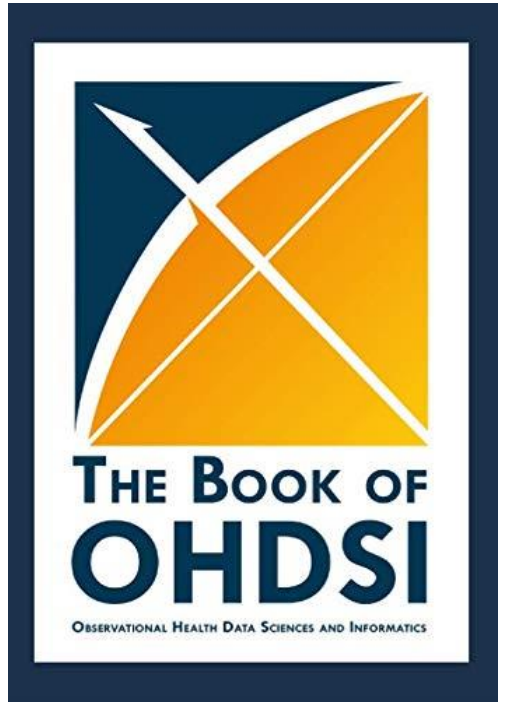
Paul Nagy

Deputy Director, Johns Hopkins Medicine Technology Innovation Center
Director of Education, Biomedical Informatics and Data Science Graduate Training Programs





World map representing the EHDEN Academy students as of September 2024



Three Stages of The Journey

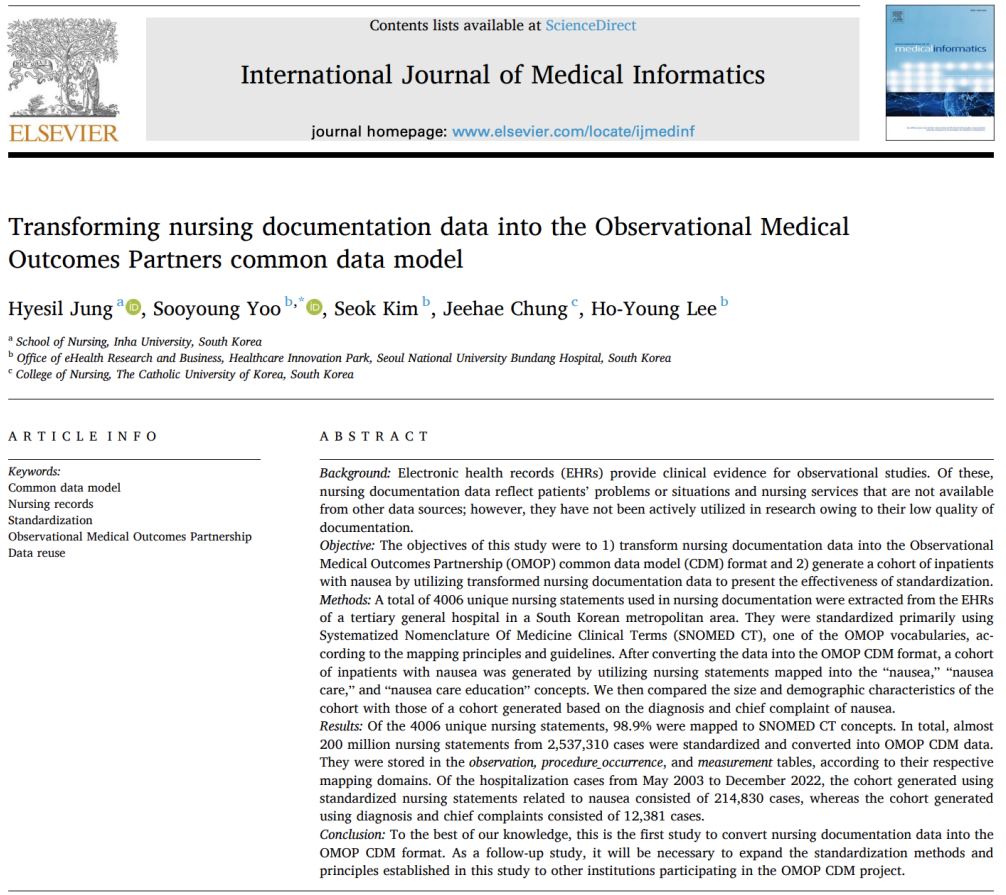
Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **Hyesil Jung, Sooyoung Yoo, Seok Kim, Jeehae Chung, and Ho-Young Lee** on the recent publication of **Transforming nursing documentation data into the Observational Medical Outcomes Partners common data model** in the *International Journal of Medical Informatics*.



International Journal of Medical Informatics 213 (2026) 106358

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.elsevier.com/locate/ijmedinf

Transforming nursing documentation data into the Observational Medical Outcomes Partners common data model

Hyesil Jung [a], Sooyoung Yoo [b,*], Seok Kim [b], Jeehae Chung [c], Ho-Young Lee [b]

[a] School of Nursing, Inha University, South Korea
[b] Office of eHealth Research and Business, Healthcare Innovation Park, Seoul National University Bundang Hospital, South Korea
[c] College of Nursing, The Catholic University of Korea, South Korea

ARTICLE INFO

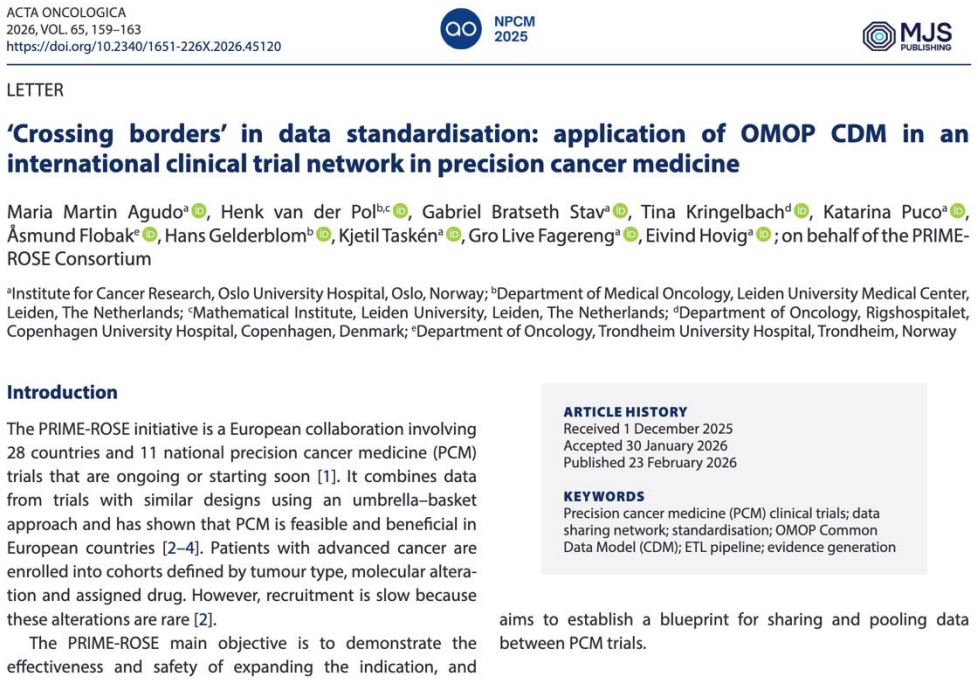
Keywords:
Common data model
Nursing records
Standardization
Observational Medical Outcomes Partnership
Data reuse

ABSTRACT

Background: Electronic health records (EHRs) provide clinical evidence for observational studies. Of these, nursing documentation data reflect patients' problems or situations and nursing services that are not available from other data sources; however, they have not been actively utilized in research owing to their low quality of documentation.
Objective: The objectives of this study were to 1) transform nursing documentation data into the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) format and 2) generate a cohort of inpatients with nausea by utilizing transformed nursing documentation data to present the effectiveness of standardization.
Methods: A total of 4006 unique nursing statements used in nursing documentation were extracted from the EHRs of a tertiary general hospital in a South Korean metropolitan area. They were standardized primarily using Systematized Nomenclature Of Medicine Clinical Terms (SNOMED CT), one of the OMOP vocabularies, according to the mapping principles and guidelines. After converting the data into the OMOP CDM format, a cohort of inpatients with nausea was generated by utilizing nursing statements mapped into the "nausea," "nausea care," and "nausea care education" concepts. We then compared the size and demographic characteristics of the cohort with those of a cohort generated based on the diagnosis and chief complaint of nausea.
Results: Of the 4006 unique nursing statements, 98.9% were mapped to SNOMED CT concepts. In total, almost 200 million nursing statements from 2,537,310 cases were standardized and converted into OMOP CDM data. They were stored in the *observation*, *procedure_occurrence*, and *measurement* tables, according to their respective mapping domains. Of the hospitalization cases from May 2003 to December 2022, the cohort generated using standardized nursing statements related to nausea consisted of 214,830 cases, whereas the cohort generated using diagnosis and chief complaints consisted of 12,381 cases.
Conclusion: To the best of our knowledge, this is the first study to convert nursing documentation data into the OMOP CDM format. As a follow-up study, it will be necessary to expand the standardization methods and principles established in this study to other institutions participating in the OMOP CDM project.

OHDSI Shoutouts!

Congratulations to the team of **Maria Martin Agudo, Henk Van der Pol, Gabriel Bratseth Stav, Tina Kringelbach, Katarina Puco, Åsmund Flobak, Hans Gelderblom, Kjetil Taskén, Gro Live Fagereng, Eivind Hovig, and the PRIME-ROSE Consortium** on the recent publication of **'Crossing borders' in data standardisation: application of OMOP CDM in an international clinical trial network in precision cancer medicine** in *Acta Oncologica*.



ACTA ONCOLOGICA
2026, VOL. 65, 159–163
https://doi.org/10.2340/1651-226X.2026.45120

NPCM 2025

MJS PUBLISHING

LETTER

'Crossing borders' in data standardisation: application of OMOP CDM in an international clinical trial network in precision cancer medicine

Maria Martin Agudo[a], Henk van der Pol[b,c], Gabriel Bratseth Stav[a], Tina Kringelbach[d], Katarina Puco[a], Åsmund Flobak[e], Hans Gelderblom[b], Kjetil Taskén[a], Gro Live Fagereng[a], Eivind Hovig[a]; on behalf of the PRIME-ROSE Consortium

[a]Institute for Cancer Research, Oslo University Hospital, Oslo, Norway; [b]Department of Medical Oncology, Leiden University Medical Center, Leiden, The Netherlands; [c]Mathematical Institute, Leiden University, Leiden, The Netherlands; [d]Department of Oncology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; [e]Department of Oncology, Trondheim University Hospital, Trondheim, Norway

Introduction

The PRIME-ROSE initiative is a European collaboration involving 28 countries and 11 national precision cancer medicine (PCM) trials that are ongoing or starting soon [1]. It combines data from trials with similar designs using an umbrella–basket approach and has shown that PCM is feasible and beneficial in European countries [2–4]. Patients with advanced cancer are enrolled into cohorts defined by tumour type, molecular alteration and assigned drug. However, recruitment is slow because these alterations are rare [2].

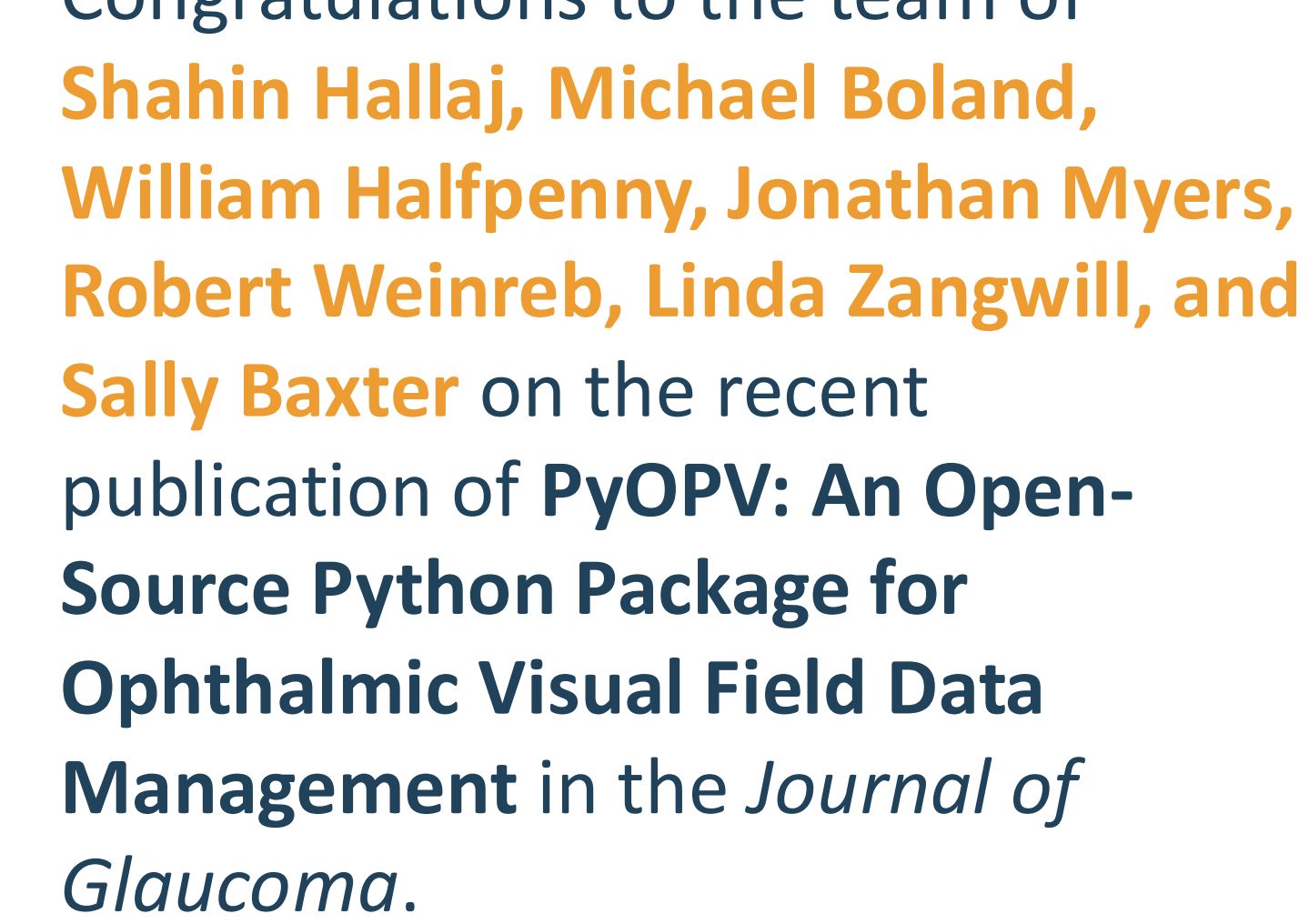
The PRIME-ROSE main objective is to demonstrate the effectiveness and safety of expanding the indication, and aims to establish a blueprint for sharing and pooling data between PCM trials.

ARTICLE HISTORY
Received 1 December 2025
Accepted 30 January 2026
Published 23 February 2026

KEYWORDS
Precision cancer medicine (PCM) clinical trials; data sharing network; standardisation; OMOP Common Data Model (CDM); ETL pipeline; evidence generation

OHDSI Shoutouts!

Congratulations to the team of **Shahin Hallaj, Michael Boland, William Halfpenny, Jonathan Myers, Robert Weinreb, Linda Zangwill, and Sally Baxter** on the recent publication of **PyOPV: An Open-Source Python Package for Ophthalmic Visual Field Data Management** in the *Journal of Glaucoma*.



ORIGINAL STUDY

PyOPV: An Open-Source Python Package for Ophthalmic Visual Field Data Management

Shahin Hallaj, MD,† Michael V. Boland, MD, PhD,‡ William Halfpenny, MD, MS,*† Jonathan S. Myers, MD,§ Robert N. Weinreb, MD,* Linda M. Zangwill, PhD,* and Sally L. Baxter, MD, MSc*†*

Précis: PyOPV is a software designed and validated for handling standard visual field DICOM files, enabling multiple functionalities for glaucoma researchers.

Purpose: To introduce PyOPV, a novel vendor-agnostic Python-based software package we designed for the management and analysis of OPhthalmic Visual field (OPV) DICOM data. PyOPV addresses limitations in interoperability and data accessibility encountered by vision researchers by providing tools that check DICOM compliance, parse, and convert OPV DICOM files into formats easily usable for research and integration with research data systems (eg, Pandas Dataframes, JSON).

Methods: PyOPV was developed using Python 3.8.2. It uses Supplement 146 of the DICOM standard to check compliance, which defines the "ophthalmic-visual-field-static-perimetry-measurements" Composite Information Object Definition. Sample OPV DICOM files from 3 vendors that provide perimetry devices were used to design the package and analyzed for DICOM. The functionalities were then validated at 2 different institutions.

Results: PyOPV successfully extracted and converted OPV DICOM data into Pandas DataFrames and JSON formats, facilitating data access, analysis, and visualization. The validation on longitudinal files from different protocols demonstrated excellent agreement between PyOPV outputs and ground truth data extracted using in-place workflows of each institution. Further, it highlighted significant interoperability challenges by demonstrating missing attributes across vendors, with a considerable proportion (range: 17%–51%) of the required tags missing from the files.

Conclusions: PyOPV provides an efficient solution for handling ophthalmic visual field data, bridging a critical gap in data interoperability and research scalability. It can incorporate OPV files from different vendors and distinct protocols in bulk, thereby enhancing the ability to analyze and integrate visual field data into large-scale health data warehouses, supporting ophthalmic informatics and advancing clinical research. PyOPV is limited by the vendors' failure to provide all data elements.

Key Words: standard automated perimetry, OPV DICOM, PyOPV, visual field data, interoperability, OMOP CDM, ophthalmic informatics

(*J Glaucoma* 2026;35:150–156)

OHDSI Shoutouts!

Congratulations to the team of **Borham Kim, Wongeun Song, Eunsil Yoon, Seok Kim, Ho-Young Lee, Jee Hyun Kim, Koung Jin Suh, Kwang-Il Kim, So Yeon Park, Eun-Kyu Kim, Se Hyun Kim, Seonghae Yoon and Sooyoung Yoo** on the recent publication of **Transforming unstructured breast cancer pathology reports into the Observational Medical Outcomes Partnership Common Data Model** in *BMC Medical Informatics and Decision-Making*.

BMC Med Inform Decis Mak https://doi.org/10.1186/s12911-026-03375-7

Article in Press

Transforming unstructured breast cancer pathology reports into the Observational Medical Outcomes Partnership Common Data Model

Received: 2 October 2025

Accepted: 4 February 2026

Published online: 28 February 2026

Cite this article as: Kim B., Song W., Yoon E. *et al.* Transforming unstructured breast cancer pathology reports into the Observational Medical Outcomes Partnership Common Data Model. *BMC Med Inform Decis Mak* (2026). https://doi.org/10.1186/s12911-026-03375-7

Borham Kim, Wongeun Song, Eunsil Yoon, Seok Kim, Ho-Young Lee, Jee Hyun Kim, Koung Jin Suh, Kwang-Il Kim, So Yeon Park, Eun-Kyu Kim, Se Hyun Kim, Seonghae Yoon & Sooyoung Yoo

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

Upcoming Workgroup Calls

Date	Time (ET)	Meeting
Wednesday	8 am	Psychiatry
Wednesday	7 pm	Medical Imaging
Thursday	10 am	GIS – Geographic Information System
Thursday	11 am	Industry
Thursday	12 pm	Methods Research
Thursday	1 pm	Oncology Vocabulary/Development Subgroup
Thursday	2 pm	Early-Stage Researchers
Thursday	7 pm	Dentistry
Friday	9 am	Waveform
Friday	10 am	Transplant
Friday	11:30 am	Steering
Tuesday	9 am	Oncology Genomic Subgroup
Tuesday	9 am	Data2Evidence

2026 Global Symposium
Scientific Review Committee

The deadline to sign up for the Scientific Review Committee is March 3. The first meeting will take place on Thursday, March 5th at 11 am ET.

Save The Date: UK Symposium

The 2026 UK Symposium will be held Sept. 18 at the University of Nottingham. There will be an an OMOP Training Day on Sept. 17.

2026 Europe Symposium

The 2026 OHDSI Europe Symposium returns to Rotterdam next year and will be held **April 18-20.**

Registration is open on the **OHDSI & OHDSI Europe** web sites.

Time	Symposium Agenda - Monday April 20, 2026	Location
8:00	***Registration and Coffee***	Queen's Lounge
9:00	**Welcome to OHDSI Europe** Dr. Renske Los, Department of Medical Informatics, *Erasmus MC* Dr. Aniek Markus, Department of Medical Informatics, *Erasmus MC*	Theatre
9:05	**Journey of OHDSI** Prof. Peter Rijnbeek, Chair Department of Medical Informatics, *Erasmus MC*	Theatre
9:30	**Collaborator Showcase - part 1** Moderated by Dr. Egill Fridgeirsson, Department of Medical Informatics, *Erasmus MC*	Theatre
10:00	**Speed networking**	Theatre
10:15	***Coffee Break & posters National Nodes***	Queen's Lounge
11:15	**Collaborator Showcase - part 2** Moderated by Dr. Egill Fridgeirsson, Department of Medical Informatics, *Erasmus MC*	Theatre
11:45	**Dreaming about the OHDSI journey ahead** Dr. Patrick Ryan, Vice President, Observational Health Data Analytics, *Johnson & Johnson* Dr. Renske Los, Department of Medical Informatics, *Erasmus MC*	Theatre
12:15	***Lunch break & networking & posters/demo's*** *(Early investigator meeting - 13:00-13:45 Queen's Lounge)*	La Fontaine & Odyssee Room
13:45	**From dreams to reality** OHDSI Titan Award winners	Theatre
14:30	**Propositions for collaboration from the National Nodes** National Node leads	Theatre
14:45	***Coffee break & posters/demo's***	La Fontaine & Odyssee Room
16:15	**The OH Factor** To be announced	Theatre
17:00	**Closing**	Theatre
17:15	***Networking reception***	Queen's Lounge

March Newsletter Is Available

On The Journey: March 2026

The momentum of 2026 is building across our community, and this month we celebrate the collaborative engine that makes our mission possible: our workgroups. Many of these teams have now finalized their Objectives and Key Results (OKRs) for the year, providing a clear roadmap for the impactful research and development ahead. As we look toward applying these goals in person, we are happy to announce that the agenda for the 2026 OHDSI Europe Symposium in Rotterdam is officially posted and registration is now open. Whether you are aligning with a workgroup's 2026 vision or planning your trip to the SS Rotterdam this April, there are more ways than ever to find your home in the OHDSI journey. #JoinTheJourney

Podcast: 2026 Workgroups, Europe Symposium

In the March 2026 On The Journey podcast, Patrick Ryan and Craig Sachson discuss the impact of workgroups within the OHDSI community, and some of the exciting research goals for the coming year. They also look at the upcoming Europe Symposium (April 18-20, Rotterdam, Netherlands), and highlight some of the exciting research submitted for the collaborator showcase. *(If video does not appear, please click 'view this email in your browser.)*

Community Updates

Where Have We Been?

- **OHDSI Europe Registration:** Registration is officially open for the OHDSI Europe Symposium, held April 18–20 in Rotterdam. We were thrilled to receive over 140 collaborator showcase submissions. You can find more details on the event further down in this newsletter.
- **Workgroup Momentum:** With nearly 40 active workgroups, there are endless opportunities for veterans and newcomers alike to contribute their talents. Several workgroup leads recently shared their 2026 research goals—read more about their visions below.
- **System Restored:** We sincerely apologize for the technical disruptions experienced throughout February. All MS Teams accounts have been fully restored with enhanced security layers now in place, and we hope to avoid future challenges. We appreciate your patience and continued support during this process.

Where Are We Now?

- **Vocabulary Refresh:** The team has released the Winter 2026 Vocabulary Refresh. Join us for the March 3 community call, where workgroup leaders will highlight critical domain changes and newly added concepts.
- **Educational Pathways:** Training the next generation of RWE researchers is vital to our mission. Our March 10 community call will focus on diverse pathways for knowledge sharing and sustainability within the community.
- **Scientific Review Committee:** We are seeking volunteers for the Scientific Review Committee to help evaluate submissions for the OHDSI Europe Collaborator Showcase. If you are interested, please fill out this form by March 3. Our first meeting is Thursday, March 5, at 11 am ET.

Where Are We Going?

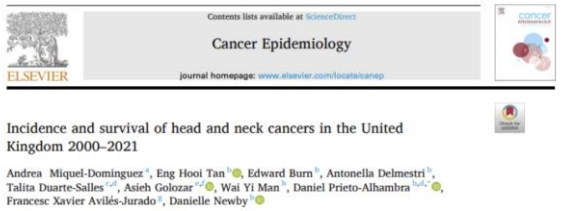
- **OHDSI UK 2026:** Mark your calendars for September 18 at the University of Nottingham, preceded by an OMOP Training Day on September 17. Please register to stay updated on the agenda, training sessions, registration and the call for abstracts.
- **Columbia University Summer School:** Registration is open for the 2026 Summer School in Observational Health Data Science & Informatics, AI, and RWE. This immersive, hands-on program runs June 22-26 at Columbia University's DBMI. Now in its second year, it is a premier opportunity for health professionals and researchers to master real-world data generation.
- **2026 Global Symposium:** Save the date! The 2026 Global Symposium returns to the Hyatt Regency in New Brunswick, NJ, from October 20-22. We will share agenda and showcase details as they become available.

My Journey: Sarah Seager

In the latest installment of our "My Journey" series, Sarah Seager (Senior Director of Business Development at EPAM) shares how she found a global "family" within the OHDSI community. From the impressive scale of multi-country data systems to the creative "artistic flavor" she brings to the science, Sarah discusses why OHDSI is a place where everyone—from newcomers to "OGs"—can contribute to a shared mission. *(If video does not appear, please click 'view this email in your browser.)*

Find Your Home in OHDSI: Workgroups Highlight Exciting Research Initiatives for 2026

February Publications



Razzaghi H, Dickinson K, Wieand K, Boss S, Weidlich H, Huang Y, Morse K, Mutyala SK, Nandagopal JPA, Viswanathan K, Forrest CB, Bailey LC. A multifaceted approach to advancing data quality and fitness standards in multi-institutional networks. J Am Med Inform Assoc. 2026 Feb 1;33(2):371-382. doi: 10.1093/jamia/ocaf181. PMID: 41128352; PMCID: PMC12844579.

Adams MCB, Hurley RW, Bartels K, Perkins ML, Hudson C, Topaloglu U, Cobb JP, Reuter-Rice K, Stocking JC, Khanna AK. Extending the Observational Medical Outcomes Partnership (OMOP) Common Data Model for Critical Care Medicine: A Framework for Standardizing Complex ICU Data Using the Society of Critical Care Medicine's Critical Care Data Dictionary (C2D2). Crit Care Med. 2026 Feb 1;54(2):270-279. doi: 10.1097/CCM.0000000000006969. Epub 2025 Nov 21. PMID: 41269063.

Hamamoto R, Koyama T, Takahashi S, Yasuda T, Kobayashi K, Akagi Y, Kouno N, Sudo K, Hirata M, Sunami K, Kubo T, Katayama H, Takashima A, Taniguchi T, Matsumoto H, Shibaki R, Asada K, Komatsu M, Kaneko S, Yamada M, Horinouchi H, Tanaka K, Goto Y, Kato K, Saito Y, Nakamura K, Yamamoto N. Implementing generative artificial intelligence in precision oncology: safety, governance, and significance. J Hematol Oncol. 2026 Feb 9;19(1):14. doi: 10.1186/s13045-026-01781-y. PMID: 41664164; PMCID: PMC12896320.

Miquel-Dominguez A, Tan EH, Burn E, Delmestri A, Duarte-Salles T, Golozar A, Man WY, Prieto-Alhambra D, Avilés-Jurado FX, Newby D. Incidence and survival of head and neck cancers in the United Kingdom 2000-2021. Cancer Epidemiol. 2026 Feb 11;101:103018. doi: 10.1016/j.canep.2026.103018. Epub ahead of print. PMID: 41678907.

March Newsletter Is Available

#OHDSISocialShowcase This Week

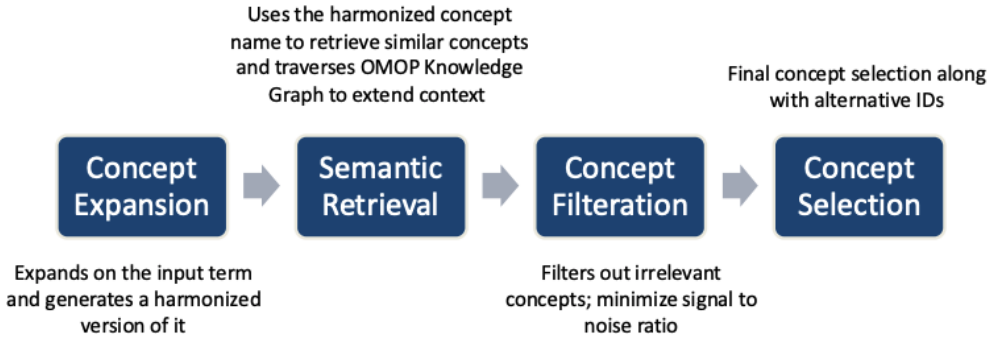
Monday

Automated OMOP Concept Mapping Using Multi-Agent Large Language Models and Graph-Enhanced Semantic Retrieval

(Adil Ahmed, Selvin Soby, Boudewijn Aasman, Parsa Mirhaji)

Summary

A LLM-workflow that maps clinical terminologies to standard OMOP concepts. The pipeline consists of 4 stages:

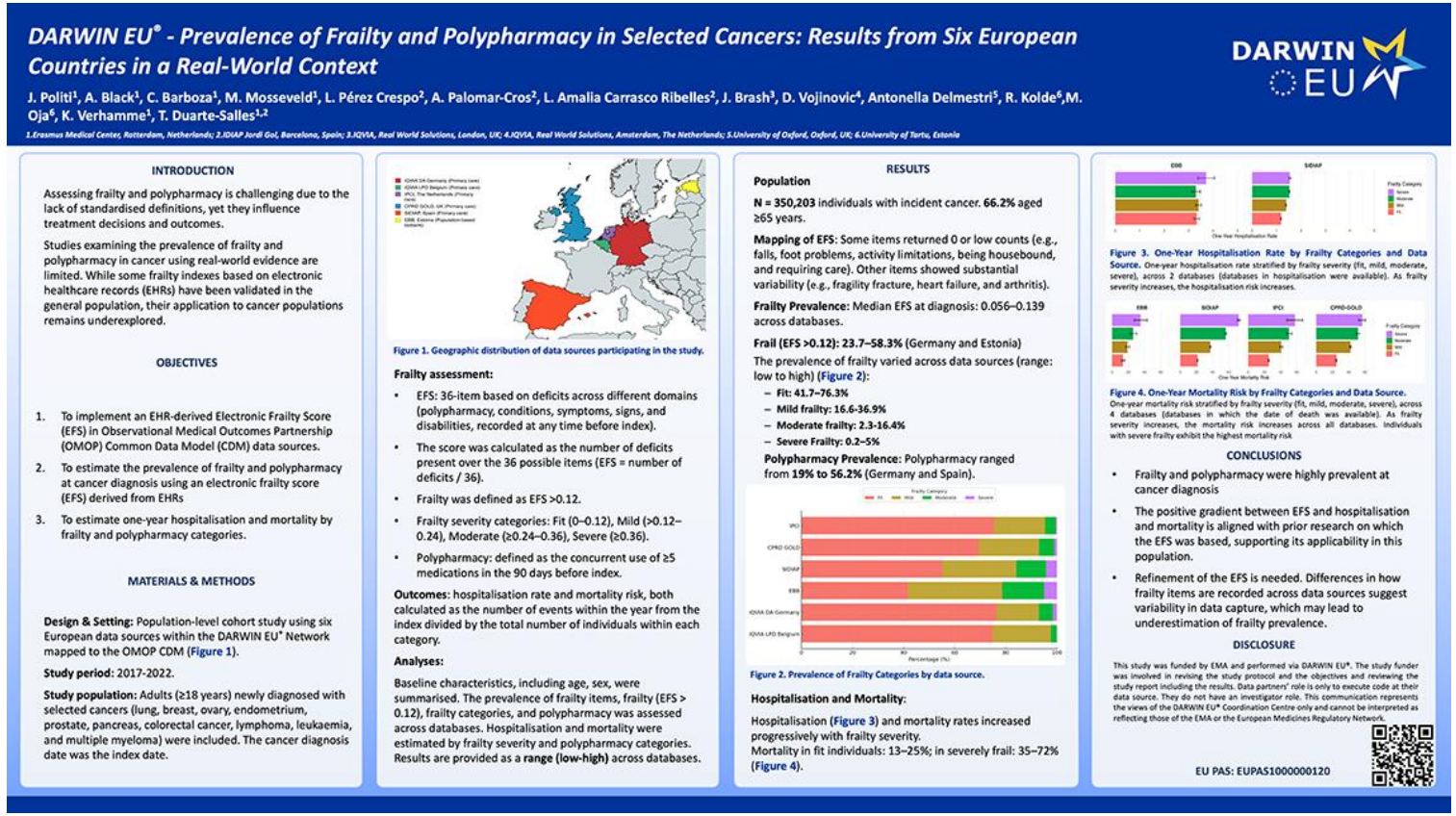


#OHDSISocialShowcase This Week

Tuesday

DARWIN EU® – Assessing Frailty and Polypharmacy in Oncology within OMOP Common Data Model: A Real-World Data Approach

(Julieta Politi, Adam Black, Cesar Barboza, Mees Mosseveld, Laura Pérez Crespo, Ana Palomar-Cros, Lucia Carrasco Ribelles, James Brash, Dragana Vojinovic, Antonella Delmestri, Raivo Kolde, Marek Oja, Katia Verhamme, Talita Duarte-Salles)



DARWIN EU® - Prevalence of Frailty and Polypharmacy in Selected Cancers: Results from Six European Countries in a Real-World Context

J. Politi[1], A. Black[1], C. Barboza[1], M. Mosseveld[1], L. Pérez Crespo[2], A. Palomar-Cros[2], L. Amalia Carrasco Ribelles[2], J. Brash[3], D. Vojinovic[4], Antonella Delmestri[5], R. Kolde[6], M. Oja[6], K. Verhamme[1], T. Duarte-Salles[1,2]

1.Erasmus Medical Center, Rotterdam, Netherlands; 2.IDIAP Jordi Gol, Barcelona, Spain; 3.IQVIA, Real World Solutions, London, UK; 4.IQVIA, Real World Solutions, Amsterdam, The Netherlands; 5.University of Oxford, Oxford, UK; 6.University of Tartu, Estonia

DARWIN EU

INTRODUCTION

Assessing frailty and polypharmacy is challenging due to the lack of standardised definitions, yet they influence treatment decisions and outcomes.

Studies examining the prevalence of frailty and polypharmacy in cancer using real-world evidence are limited. While some frailty indexes based on electronic healthcare records (EHRs) have been validated in the general population, their application to cancer populations remains underexplored.

OBJECTIVES

1. To implement an EHR-derived Electronic Frailty Score (EFS) in Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) data sources.
2. To estimate the prevalence of frailty and polypharmacy at cancer diagnosis using an electronic frailty score (EFS) derived from EHRs
3. To estimate one-year hospitalisation and mortality by frailty and polypharmacy categories.

MATERIALS & METHODS

Design & Setting: Population-level cohort study using six European data sources within the DARWIN EU® Network mapped to the OMOP CDM (Figure 1).

Study period: 2017-2022.

Study population: Adults (≥18 years) newly diagnosed with selected cancers (lung, breast, ovary, endometrium, prostate, pancreas, colorectal cancer, lymphoma, leukaemia, and multiple myeloma) were included. The cancer diagnosis date was the index date.

Figure 1. Geographic distribution of data sources participating in the study.

Frailty assessment:

- EFS: 36-item based on deficits across different domains (polypharmacy, conditions, symptoms, signs, and disabilities, recorded at any time before index).
- The score was calculated as the number of deficits present over the 36 possible items (EFS = number of deficits / 36).
- Frailty was defined as EFS >0.12.
- Frailty severity categories: Fit (0–0.12), Mild (>0.12–0.24), Moderate (≥0.24–0.36), Severe (≥0.36).
- Polypharmacy: defined as the concurrent use of ≥5 medications in the 90 days before index.

Outcomes: hospitalisation rate and mortality risk, both calculated as the number of events within the year from the index divided by the total number of individuals within each category.

Analyses:

Baseline characteristics, including age, sex, were summarised. The prevalence of frailty items, frailty (EFS > 0.12), frailty categories, and polypharmacy was assessed across databases. Hospitalisation and mortality were estimated by frailty severity and polypharmacy categories. Results are provided as a **range (low-high)** across databases.

RESULTS

Population

N = 350,203 individuals with incident cancer. **66.2%** aged ≥65 years.

Mapping of EFS: Some items returned 0 or low counts (e.g., falls, foot problems, activity limitations, being housebound, and requiring care). Other items showed substantial variability (e.g., fragility fracture, heart failure, and arthritis).

Frailty Prevalence: Median EFS at diagnosis: 0.056–0.139 across databases.

Frail (EFS >0.12): 23.7–58.3% (Germany and Estonia)

The prevalence of frailty varied across data sources (range: low to high) (Figure 2):

- **Fit: 41.7–76.3%**
- **Mild frailty: 16.6-36.9%**
- **Moderate frailty: 2.3-16.4%**
- **Severe Frailty: 0.2–5%**

Polypharmacy Prevalence: Polypharmacy ranged from **19% to 56.2%** (Germany and Spain).

Figure 2. Prevalence of Frailty Categories by data source.

Hospitalisation and Mortality:

Hospitalisation (Figure 3) and mortality rates increased progressively with frailty severity.
Mortality in fit individuals: 13–25%; in severely frail: 35–72% (Figure 4).

Figure 3. One-Year Hospitalisation Rate by Frailty Categories and Data Source. One-year hospitalisation rate stratified by frailty severity (fit, mild, moderate, severe), across 2 databases (databases in hospitalisation were available). As frailty severity increases, the hospitalisation risk increases.

Figure 4. One-Year Mortality Risk by Frailty Categories and Data Source. One-year mortality risk stratified by frailty severity (fit, mild, moderate, severe), across 4 databases (databases in which the date of death was available). As frailty severity increases, the mortality risk increases across all databases. Individuals with severe frailty exhibit the highest mortality risk

CONCLUSIONS

- Frailty and polypharmacy were highly prevalent at cancer diagnosis
- The positive gradient between EFS and hospitalisation and mortality is aligned with prior research on which the EFS was based, supporting its applicability in this population.
- Refinement of the EFS is needed. Differences in how frailty items are recorded across data sources suggest variability in data capture, which may lead to underestimation of frailty prevalence.

DISCLOSURE

This study was funded by EMA and performed via DARWIN EU®. The study funder was involved in revising the study protocol and the objectives and reviewing the study report including the results. Data partners' role is only to execute code at their data source. They do not have an investigator role. This communication represents the views of the DARWIN EU® Coordination Centre only and cannot be interpreted as reflecting those of the EMA or the European Medicines Regulatory Network.

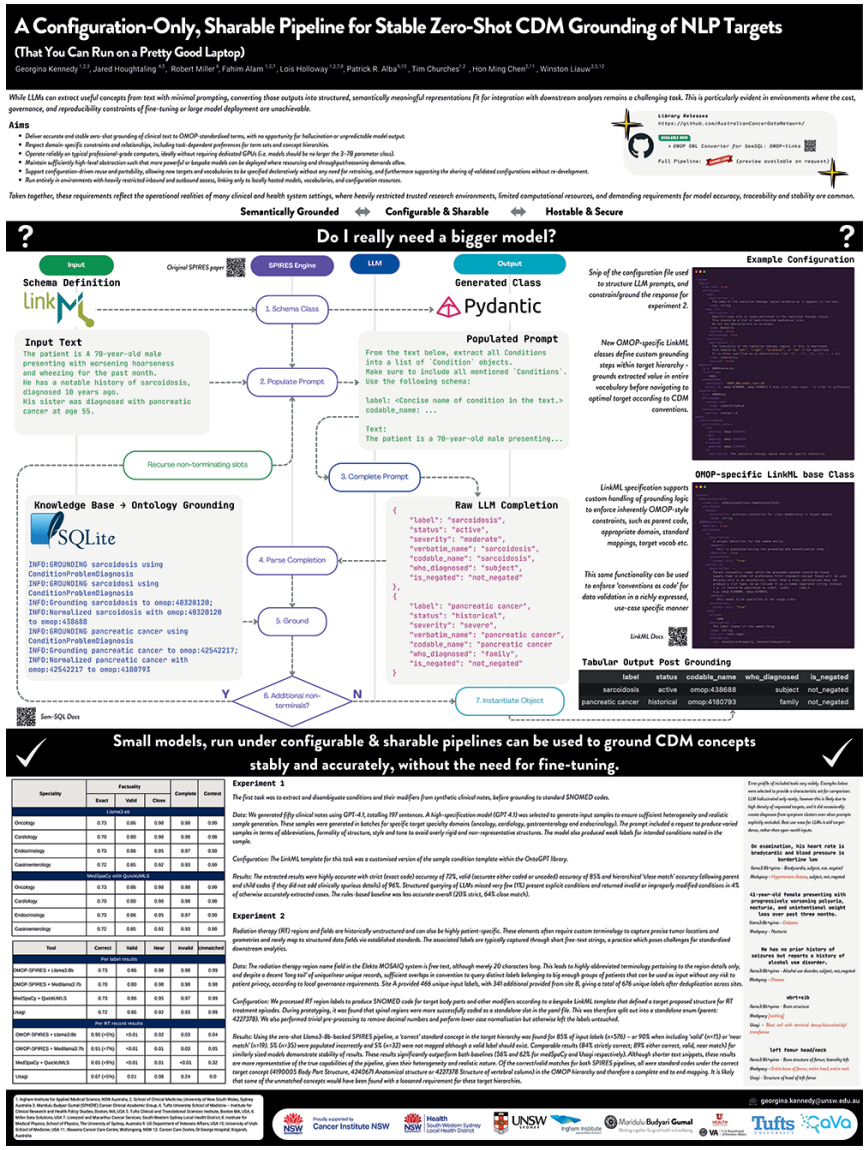
EU PAS: EUPAS1000000120

#OHDSISocialShowcase This Week

Wednesday

A Configuration-Only, Sharable Pipeline for Stable Zero-Shot CDM Grounding of NLP Targets (That You Can Run on a Pretty Good Laptop)

(Georgina Kennedy, Jared Houghtaling, Robert Miller, Fahim Alam, Lois Holloway, Tim Churches, Winston Liauw)

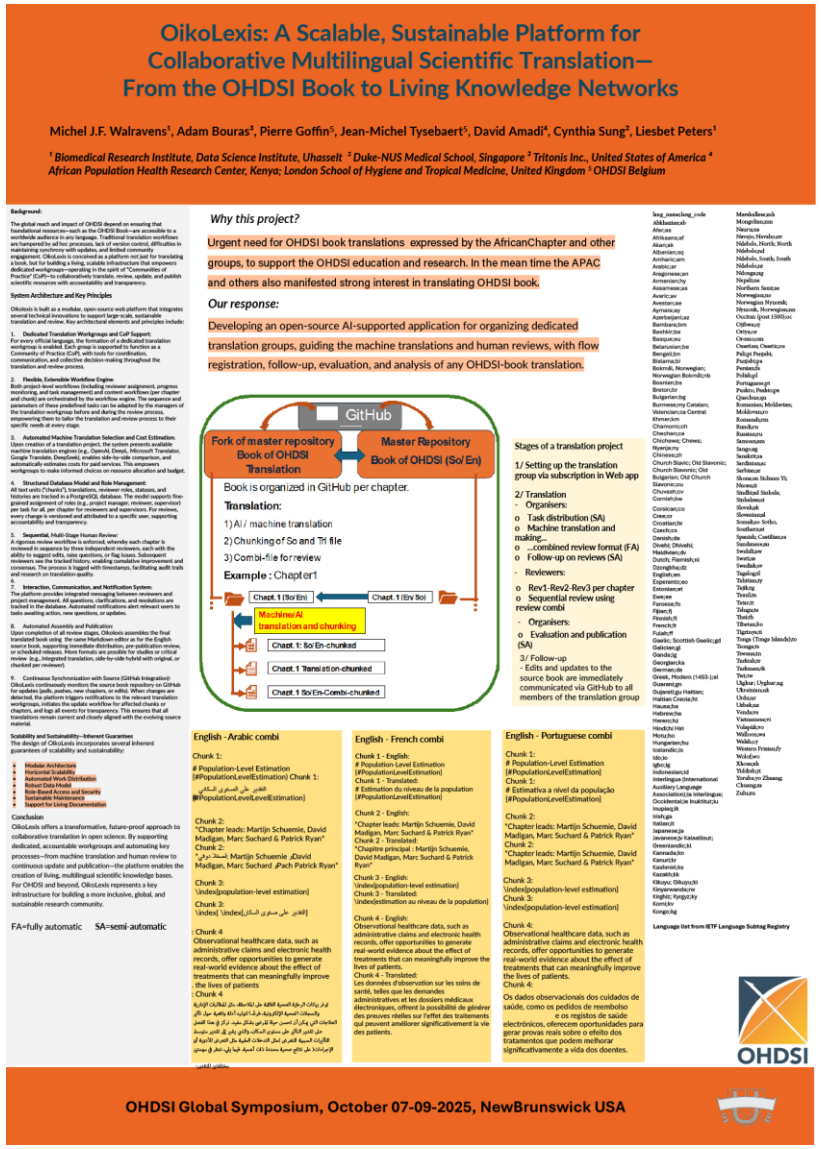


#OHDSISocialShowcase This Week

Thursday

Metalexis, A Scalable, Sustainable Platform for Collaborative Multilingual Scientific Translation of the OHDSI Book; potential evolution towards AI-enabled Living Knowledge Networks

(Michel J.F. Walravens, Adam Bouras, Pierre Goffin, Jean-Michel Tysebaert, David Amadi, Cynthia Sung, Liesbet Peeters)

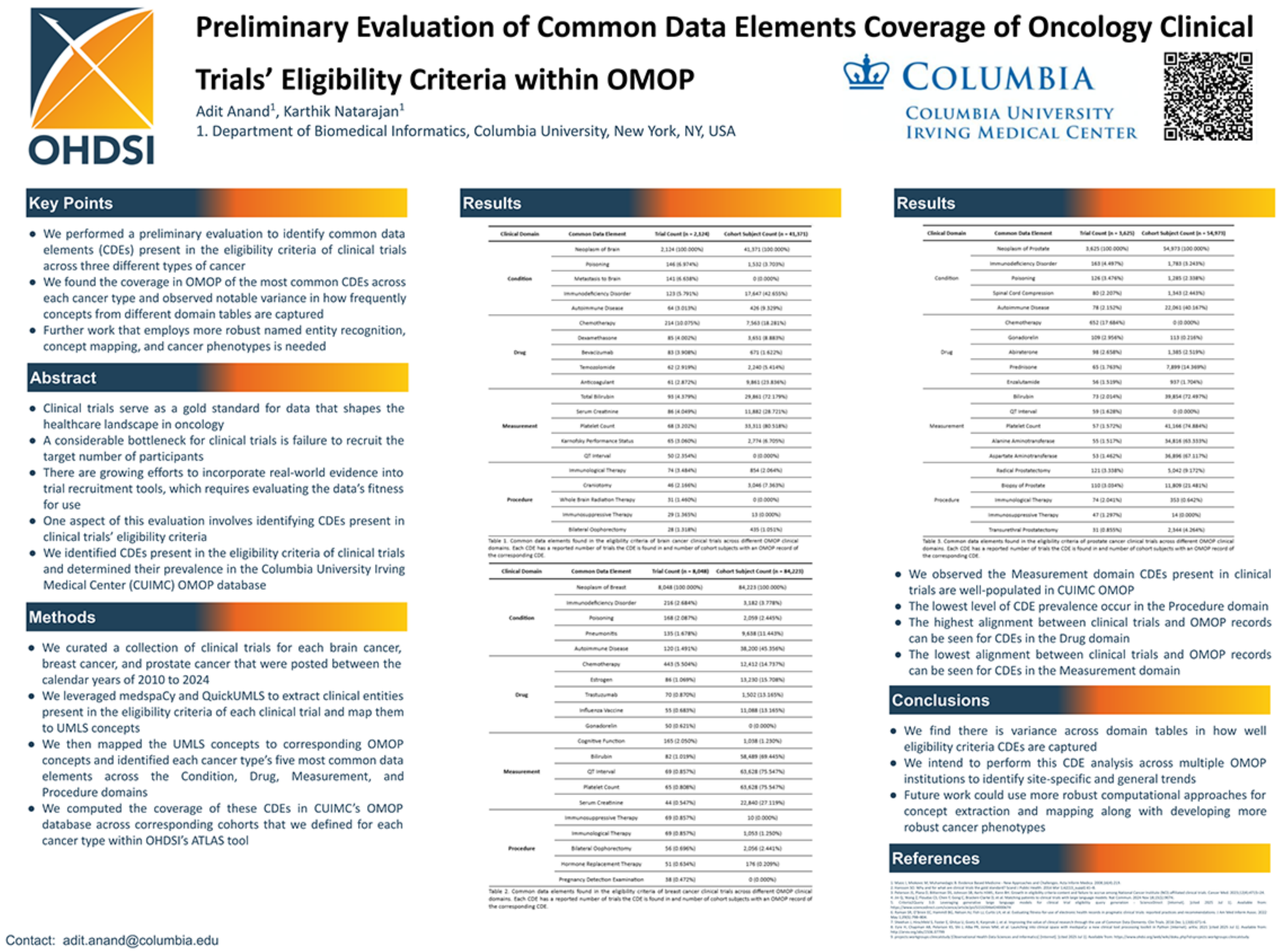


#OHDSISocialShowcase This Week

Friday

Preliminary Evaluation of Common Data Elements Coverage of Oncology Clinical Trials' Eligibility Criteria within OMOP

(Adit Anand, Karthik Natarajan)



Preliminary Evaluation of Common Data Elements Coverage of Oncology Clinical Trials' Eligibility Criteria within OMOP

Adit Anand[1], Karthik Natarajan[1]
1. Department of Biomedical Informatics, Columbia University, New York, NY, USA

COLUMBIA
COLUMBIA UNIVERSITY IRVING MEDICAL CENTER

Key Points

- We performed a preliminary evaluation to identify common data elements (CDEs) present in the eligibility criteria of clinical trials across three different types of cancer
- We found the coverage in OMOP of the most common CDEs across each cancer type and observed notable variance in how frequently concepts from different domain tables are captured
- Further work that employs more robust named entity recognition, concept mapping, and cancer phenotypes is needed

Abstract

- Clinical trials serve as a gold standard for data that shapes the healthcare landscape in oncology
- A considerable bottleneck for clinical trials is failure to recruit the target number of participants
- There are growing efforts to incorporate real-world evidence into trial recruitment tools, which requires evaluating the data's fitness for use
- One aspect of this evaluation involves identifying CDEs present in clinical trials' eligibility criteria
- We identified CDEs present in the eligibility criteria of clinical trials and determined their prevalence in the Columbia University Irving Medical Center (CUIMC) OMOP database

Methods

- We curated a collection of clinical trials for each brain cancer, breast cancer, and prostate cancer that were posted between the calendar years of 2010 to 2024
- We leveraged medspaCy and QuickUMLS to extract clinical entities present in the eligibility criteria of each clinical trial and map them to UMLS concepts
- We then mapped the UMLS concepts to corresponding OMOP concepts and identified each cancer type's five most common data elements across the Condition, Drug, Measurement, and Procedure domains
- We computed the coverage of these CDEs in CUIMC's OMOP database across corresponding cohorts that we defined for each cancer type within OHDSI's ATLAS tool

Contact: adit.anand@columbia.edu

Results

Clinical Domain	Common Data Element	Trial Count (n = 2,124)	Cohort Subject Count (n = 41,371)
Condition	Neoplasm of Brain	2,124 (100.000%)	41,371 (100.000%)
	Poisoning	146 (6.974%)	1,532 (3.703%)
	Metastasis to Brain	141 (6.638%)	0 (0.000%)
	Immunodeficiency Disorder	123 (5.791%)	17,647 (42.655%)
	Autoimmune Disease	64 (3.013%)	426 (9.329%)
Drug	Chemotherapy	214 (10.075%)	7,563 (18.281%)
	Dexamethasone	85 (4.002%)	3,651 (8.883%)
	Bevacizumab	83 (3.908%)	671 (1.622%)
	Temozolomide	62 (2.919%)	2,240 (5.414%)
	Anticoagulant	61 (2.872%)	9,861 (23.836%)
Measurement	Total Bilirubin	93 (4.379%)	29,861 (72.179%)
	Serum Creatinine	86 (4.049%)	11,882 (28.721%)
	Platelet Count	68 (3.202%)	33,311 (80.518%)
	Karnofsky Performance Status	65 (3.060%)	2,774 (6.705%)
	QT Interval	50 (2.354%)	0 (0.000%)
Procedure	Immunological Therapy	74 (3.484%)	854 (2.064%)
	Craniotomy	46 (2.166%)	3,046 (7.363%)
	Whole Brain Radiation Therapy	31 (1.460%)	0 (0.000%)
	Immunosuppressive Therapy	29 (1.365%)	13 (0.000%)
	Bilateral Oophorectomy	28 (1.318%)	435 (1.051%)

Table 1. Common data elements found in the eligibility criteria of brain cancer clinical trials across different OMOP clinical domains. Each CDE has a reported number of trials the CDE is found in and number of cohort subjects with an OMOP record of the corresponding CDE.

Clinical Domain	Common Data Element	Trial Count (n = 8,048)	Cohort Subject Count (n = 84,223)
Condition	Neoplasm of Breast	8,048 (100.000%)	84,223 (100.000%)
	Immunodeficiency Disorder	216 (2.684%)	3,182 (3.778%)
	Poisoning	168 (2.087%)	2,059 (2.445%)
	Pneumonitis	135 (1.678%)	9,638 (11.443%)
	Autoimmune Disease	120 (1.491%)	38,200 (45.356%)
Drug	Chemotherapy	443 (5.504%)	12,412 (14.737%)
	Estrogen	86 (1.069%)	13,230 (15.708%)
	Trastuzumab	70 (0.870%)	1,502 (13.165%)
	Influenza Vaccine	55 (0.683%)	11,088 (13.165%)
	Gonadorelin	50 (0.621%)	0 (0.000%)
Measurement	Cognitive Function	165 (2.050%)	1,038 (1.230%)
	Bilirubin	82 (1.019%)	58,489 (69.445%)
	QT Interval	69 (0.857%)	63,628 (75.547%)
	Platelet Count	65 (0.808%)	63,628 (75.547%)
	Serum Creatinine	44 (0.547%)	22,840 (27.119%)
Procedure	Immunosuppressive Therapy	69 (0.857%)	10 (0.000%)
	Immunological Therapy	69 (0.857%)	1,053 (1.250%)
	Bilateral Oophorectomy	56 (0.696%)	2,056 (2.441%)
	Hormone Replacement Therapy	51 (0.634%)	176 (0.209%)
	Pregnancy Detection Examination	38 (0.472%)	0 (0.000%)

Table 2. Common data elements found in the eligibility criteria of breast cancer clinical trials across different OMOP clinical domains. Each CDE has a reported number of trials the CDE is found in and number of cohort subjects with an OMOP record of the corresponding CDE.

Results

Clinical Domain	Common Data Element	Trial Count (n = 3,625)	Cohort Subject Count (n = 54,973)
Condition	Neoplasm of Prostate	3,625 (100.000%)	54,973 (100.000%)
	Immunodeficiency Disorder	163 (4.497%)	1,783 (3.243%)
	Poisoning	126 (3.476%)	1,285 (2.338%)
	Spinal Cord Compression	80 (2.207%)	1,343 (2.443%)
	Autoimmune Disease	78 (2.152%)	22,061 (40.167%)
Drug	Chemotherapy	652 (17.684%)	0 (0.000%)
	Gonadorelin	109 (2.956%)	113 (0.216%)
	Abiraterone	98 (2.658%)	1,385 (2.519%)
	Prednisone	65 (1.763%)	7,899 (14.369%)
	Enzalutamide	56 (1.519%)	937 (1.704%)
Measurement	Bilirubin	73 (2.014%)	39,854 (72.497%)
	QT Interval	59 (1.628%)	0 (0.000%)
	Platelet Count	57 (1.572%)	41,166 (74.884%)
	Alanine Aminotransferase	55 (1.517%)	34,816 (63.333%)
	Aspartate Aminotransferase	53 (1.462%)	36,896 (67.117%)
Procedure	Radical Prostatectomy	121 (3.338%)	5,042 (9.172%)
	Biopsy of Prostate	110 (3.034%)	11,809 (21.481%)
	Immunological Therapy	74 (2.041%)	353 (0.642%)
	Immunosuppressive Therapy	47 (1.297%)	14 (0.000%)
	Transurethral Prostatectomy	31 (0.855%)	2,344 (4.264%)

Table 3. Common data elements found in the eligibility criteria of prostate cancer clinical trials across different OMOP clinical domains. Each CDE has a reported number of trials the CDE is found in and number of cohort subjects with an OMOP record of the corresponding CDE.

- We observed the Measurement domain CDEs present in clinical trials are well-populated in CUIMC OMOP
- The lowest level of CDE prevalence occur in the Procedure domain
- The highest alignment between clinical trials and OMOP records can be seen for CDEs in the Drug domain
- The lowest alignment between clinical trials and OMOP records can be seen for CDEs in the Measurement domain

Conclusions

- We find there is variance across domain tables in how well eligibility criteria CDEs are captured
- We intend to perform this CDE analysis across multiple OMOP institutions to identify site-specific and general trends
- Future work could use more robust computational approaches for concept extraction and mapping along with developing more robust cancer phenotypes

References

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?

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