Workgroup OKRs + Phenotype Phebruary Update #2

OHDSI Community Call
Feb. 13, 2024 • 11 am ET
## Upcoming Community Calls

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb. 13</td>
<td>Workgroup OKRs / Phenotype Phebruary Update 2</td>
</tr>
<tr>
<td>Feb. 20</td>
<td>Workgroup OKRs / Phenotype Phebruary Update 3</td>
</tr>
<tr>
<td>Feb. 27</td>
<td>Workgroup OKRs / Phenotype Phebruary Update 4</td>
</tr>
<tr>
<td>Mar. 5</td>
<td>New Vocabulary Release Update</td>
</tr>
<tr>
<td>Mar. 12</td>
<td>TBA</td>
</tr>
<tr>
<td>Mar. 19</td>
<td>NO MEETING</td>
</tr>
<tr>
<td>Mar. 26</td>
<td>Recent OHDSI Publications</td>
</tr>
</tbody>
</table>
Three Stages of The Journey

Where Have We Been?
Where Are We Now?
Where Are We Going?
OHDSI Shoutouts!

Congratulations to the team of Xinyuan Zhang, Yixue Feng, Fang Li, Jin Ding, Danyal Tahseen, Ezekiel Hinojosa, Yong Chen, and Cui Tao on the publication of Evaluating MedDRA-to-ICD terminology mappings in *BMC Medical Informatics and Decision Making*.

**Abstract**

**Background** In this era of big data, data harmonization is an important step to ensure reproducible, scalable, and collaborative research. Thus, terminology mapping is a necessary step to harmonize heterogeneous data. Take the Medical Dictionary for Regulatory Activities (MedDRA) and International Classification of Diseases (ICD) for example, the mapping between them is essential for drug safety and pharmacovigilance research. Our main objective is to provide a quantitative and qualitative analysis of the mapping status between MedDRA and ICD.

We focus on evaluating the current mapping status between MedDRA and ICD through the Unified Medical Language System (UMLS) and Observational Medical Outcomes Partnership Common Data Model (OMOP CDM).

We summarized the current mapping statistics and evaluated the quality of the current MedDRA-ICD mapping; for unmapped terms, we used our self-developed algorithm to rank the best possible mapping candidates for additional mapping coverage.

**Results** The identified MedDRA-ICD mapped pairs cover 27.23% of the overall MedDRA preferred terms (PT). The systematic quality analysis demonstrated that, among the mapped pairs provided by UMLS, only 51.44% are considered an exact match. For the 2400 sampled unmapped terms, 56 of the 2400 MedDRA Preferred Terms (PT) could have exact match terms from ICD.
Congratulations, Dr. Chungsoo Kim!

I have graduated with a PhD from the Department of Biomedical Informatics at Ajou University. I am deeply grateful to everyone who collaborated closely and worked intensely with me (including my current/former lab friends and #OHDSI Folks). I couldn't have done it without your help and I will never forget my time in this lab.

I am very much looking forward to my new career, which is a #Postdoctoral Associate at Yale University/Yale New Haven Hospital Center for Outcomes Research and Evaluation (CORE) at Yale University School of Medicine.

Happy Lunar New Year and wish me luck!!
Three Stages of The Journey

Where Have We Been?
Where Are We Now?
Where Are We Going?
# Upcoming Workgroup Calls

<table>
<thead>
<tr>
<th>Date</th>
<th>Time (ET)</th>
<th>Meeting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuesday</td>
<td>3 pm</td>
<td>OMOP CDM Oncology Outreach/Research Subgroup</td>
</tr>
<tr>
<td>Wednesday</td>
<td>9 am</td>
<td>Patient-Level Prediction</td>
</tr>
<tr>
<td>Wednesday</td>
<td>12 pm</td>
<td>Health Equity</td>
</tr>
<tr>
<td>Wednesday</td>
<td>2 pm</td>
<td>Natural Language Processing</td>
</tr>
<tr>
<td>Wednesday</td>
<td>3 pm</td>
<td>Vulcan/OHDSI Meeting (ZOOM)</td>
</tr>
<tr>
<td>Thursday</td>
<td>9 am</td>
<td>OMOP CDM Oncology Vocabulary/Development Subgroup</td>
</tr>
<tr>
<td>Thursday</td>
<td>9:30 am</td>
<td>Themis</td>
</tr>
<tr>
<td>Thursday</td>
<td>12 pm</td>
<td>HADES</td>
</tr>
<tr>
<td>Thursday</td>
<td>7 pm</td>
<td>Dentistry</td>
</tr>
<tr>
<td>Friday</td>
<td>10 am</td>
<td>GIS – Geographic Information System</td>
</tr>
<tr>
<td>Friday</td>
<td>10:30 am</td>
<td>Open-Source Community</td>
</tr>
<tr>
<td>Friday</td>
<td>11:30 am</td>
<td>Clinical Trials</td>
</tr>
<tr>
<td>Monday</td>
<td>9 am</td>
<td>Vaccine Vocabulary</td>
</tr>
<tr>
<td>Monday</td>
<td>10 am</td>
<td>Africa Chapter</td>
</tr>
<tr>
<td>Monday</td>
<td>11 am</td>
<td>Data Bricks User Group</td>
</tr>
<tr>
<td>Monday</td>
<td>2 pm</td>
<td>Electronic Animal Health Records</td>
</tr>
</tbody>
</table>
Registration is now OPEN for the 2024 OHDSI Europe Symposium, which will be held June 1-3 in Rotterdam, Netherlands.

- June 1 – tutorial/workshop
- June 2 – tutorial/workshop
- June 3 – main conference

@OHDSI www.ohdsi.org #JoinTheJourney
Scientific Review Committee

If you are interested in joining the Scientific Review Committee for the 2024 Global Symposium, you can sign up now.

The deadline to sign up is Feb. 16, and the first meeting will be held March 7.
Phenotype Phebruary 2024

"Phenotype Phebruary" is a community-wide initiative to advance the field of phenotyping in observational studies. The OHDSI community has engaged in Phenotype Phebruary in both 2022 and 2023, and this year the community set a goal to understand what is the current practices in the field and how much researchers introduce variability in the process of phenotype development and evaluation.

Under the leadership of Azza Shoabi, Anna Astropoleos, Gowtham Rac and James Weaver, Phenotype Phebruary 2024 focuses on assessing consistency in phenotype definition components, phenotype representation structure, and phenotype validation methods. The month-long activity empowers OHDSI collaborators to engage with each other while advancing the science of phenotyping and gaining education and training around phenotype development and evaluation.

Throughout the month, collaborators will engage in a month-long study focused on assessing consistency in phenotype definitions and methods. The goal for this is to evaluate reporting patterns and consistency among reported phenotype algorithms for the same clinical phenotype across observational studies.

During the Phenotype Phebruary introductory call, community members voted to focus efforts on four specific phenotypes: Alzheimer’s Disease, pulmonary hypertension, major depression disorder and prostate cancer. Each week, there will be systematic literature search and synthesis, replication using ATLAS and other OHDSI tools, and summarize variations in population characteristics like incidence rates.

There will be consistent updates on the forum post linked below, and weekly updates during February community calls. The working folder is accessible for anybody who wants to read about our community efforts. If you are interested in joining, please consider joining the Phenotype Development & Evaluation workgroup so you have edit access to the working folder. Please join our meetings and identify an area/task you would be interested in helping complete.

ohdsi.org/phenotype-phebruary-2024
New Study: Deep Learning Comparison

Network Study: Deep Learning Comparison

Researchers: patientprediction, networkstudy

😊 It’s been a while since we’ve seen lhjohn — their last post was 2 years ago.

lhjohn Henrik John

We are pleased to announce our network study **Deep Learning Comparison**.

**Study leads:** Henrik John (@lhjohn), Chungsoo Kim (@Chungsoo_Kim), Jenna Reps (@jennareps), and Egill Fridgeirsson (@egillax)

**GitHub:** Deep Learning Comparison - GitHub Repository

**Protocol:** Deep Learning Comparison - Protocol

**Infrastructure:** To execute the analysis an Nvidia GPU with CUDA support is required. We recommend a minimum of 12 GB video memory; more is preferred to speed up analysis.

**Participant deadline:** Please let us know before 1 March, if you are interested in joining the study.

**Aim:** Assess the value of deep learning methods over conventional methods for the development of clinical prediction models. The specific diseases under consideration are dementia in individuals over 55, lung cancer in those over 45, and bipolar disorder in patients misdiagnosed with major depressive disorder.

**Rationale:** Deep learning techniques have proven to be highly effective for prediction on unstructured data, such as image and text. However, when applied to structured, sparse, and high-dimensional healthcare data deep learning often yields results comparable to those of simpler, conventional prediction methods. In this study we develop and validate clinical prediction models using deep learning and conventional approaches to compare their discriminatory power and calibration on OMOP CDM data.
March 14: Current Approaches for Distributed Analysis

Federated Analysis
State of the Science Collective Learning Series

Panel Discussion:
Current Approaches for Distributed Analysis

Thursday, March 14
10:00 a.m. PT | 1:00 p.m. ET

Dr. Judith Maro
James Weaver
Michael Paterson
Toward a General-Purpose Geography-Focused OHDSI Infrastructure

(Kyle Zollo-Venecek, Robert Miller, William G. Adams, Jay Greenfield, Timothy B Norris, Polina Talapova, Maksym Trofymenko, Andrew Williams)

**Gaia**

Gaia introduces a geography-focused infrastructure to the OMOP CDM including a universal representation for geospatial data, software tools for data ingestion, an OMOP-aligned GIS vocabulary package, and a new event table for capturing person-level exposures.

**Methods**

The OHDSI Workgroup is building support that complements prior OHDSI GIS efforts by enabling analysis of region attributes like poverty in conjunction with OMOP clinical data. Development in the workgroup can be thought of as two segments: the foundational work for a geography-focused infrastructure and OHDSI integration. Foundational work involved a data catalog for managing compatible external sources, a harmonized dataset model for staging geospatial data, and software tools for data ingestion and transformation. OHDSI integration has so far focused on the development of a just-in-time event table exposure occurrence, which relates regional information to social determinants of health (SDOH) and environmental pollutants to patients and their annotations over time.

**Results**

Gaia introduces a geography-focused infrastructure to the OMOP CDM including a universal representation for geospatial data, software tools for data ingestion, an OMOP-aligned GIS vocabulary package, and a new event table for capturing person-level exposures.

**Keywords**

Gaia, geography, OMOP, GIS, exposure, social determinants of health, environmental pollutants.
The Development and Validation of an Individual-Level Socioeconomic Deprivation Index (ISDI) with OMOP in the NIH's All of Us Data Network

(Nripendra Acharya, Karthik Natarajan)
Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(Xiaoyu Wang, Jenna Reps, Anthony Sena, James Gilbert, Marc Suchard)

Real World Evidence Using Measurement Values for Patient-Level Prediction Models: A Feasibility Study

Real World Evidence Using Measurement Values for Patient-Level Prediction Models: A Feasibility Study

Xiaoyu Wang1, Jenna Reps2, Anthony Sena2, James Gilbert2, Marc A Suchard2

1Division of Biostatistics, University of California, Berkeley, CA
2Department of Statistics, University of California, Berkeley, CA

This Week

#OHDSISocialShowcase

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(WEDNESDAY)

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(Xiaoyu Wang, Jenna Reps, Anthony Sena, James Gilbert, Marc Suchard)

#OHDSI

www.ohdsi.org

#JoinTheJourney

Real World Evidence Using Measurement Values for Patient-Level Prediction Models: A Feasibility Study

Xiaoyu Wang1, Jenna Reps2, Anthony Sena2, James Gilbert2, Marc A Suchard2

1Division of Biostatistics, University of California, Berkeley, CA
2Department of Statistics, University of California, Berkeley, CA

This Week

#OHDSISocialShowcase

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(WEDNESDAY)

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(Xiaoyu Wang, Jenna Reps, Anthony Sena, James Gilbert, Marc Suchard)

#OHDSI

www.ohdsi.org

#JoinTheJourney

Real World Evidence Using Measurement Values for Patient-Level Prediction Models: A Feasibility Study

Xiaoyu Wang1, Jenna Reps2, Anthony Sena2, James Gilbert2, Marc A Suchard2

1Division of Biostatistics, University of California, Berkeley, CA
2Department of Statistics, University of California, Berkeley, CA

This Week

#OHDSISocialShowcase

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(WEDNESDAY)

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(Xiaoyu Wang, Jenna Reps, Anthony Sena, James Gilbert, Marc Suchard)

#OHDSI

www.ohdsi.org

#JoinTheJourney

Real World Evidence Using Measurement Values for Patient-Level Prediction Models: A Feasibility Study

Xiaoyu Wang1, Jenna Reps2, Anthony Sena2, James Gilbert2, Marc A Suchard2

1Division of Biostatistics, University of California, Berkeley, CA
2Department of Statistics, University of California, Berkeley, CA

This Week

#OHDSISocialShowcase

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(WEDNESDAY)

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(Xiaoyu Wang, Jenna Reps, Anthony Sena, James Gilbert, Marc Suchard)

#OHDSI

www.ohdsi.org

#JoinTheJourney

Real World Evidence Using Measurement Values for Patient-Level Prediction Models: A Feasibility Study

Xiaoyu Wang1, Jenna Reps2, Anthony Sena2, James Gilbert2, Marc A Suchard2

1Division of Biostatistics, University of California, Berkeley, CA
2Department of Statistics, University of California, Berkeley, CA

This Week

#OHDSISocialShowcase

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(WEDNESDAY)

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(Xiaoyu Wang, Jenna Reps, Anthony Sena, James Gilbert, Marc Suchard)

#OHDSI

www.ohdsi.org

#JoinTheJourney

Real World Evidence Using Measurement Values for Patient-Level Prediction Models: A Feasibility Study

Xiaoyu Wang1, Jenna Reps2, Anthony Sena2, James Gilbert2, Marc A Suchard2

1Division of Biostatistics, University of California, Berkeley, CA
2Department of Statistics, University of California, Berkeley, CA

This Week

#OHDSISocialShowcase

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(WEDNESDAY)

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(Xiaoyu Wang, Jenna Reps, Anthony Sena, James Gilbert, Marc Suchard)

#OHDSI

www.ohdsi.org

#JoinTheJourney

Real World Evidence Using Measurement Values for Patient-Level Prediction Models: A Feasibility Study

Xiaoyu Wang1, Jenna Reps2, Anthony Sena2, James Gilbert2, Marc A Suchard2

1Division of Biostatistics, University of California, Berkeley, CA
2Department of Statistics, University of California, Berkeley, CA

This Week

#OHDSISocialShowcase

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(WEDNESDAY)

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(Xiaoyu Wang, Jenna Reps, Anthony Sena, James Gilbert, Marc Suchard)

#OHDSI

www.ohdsi.org

#JoinTheJourney

Real World Evidence Using Measurement Values for Patient-Level Prediction Models: A Feasibility Study

Xiaoyu Wang1, Jenna Reps2, Anthony Sena2, James Gilbert2, Marc A Suchard2

1Division of Biostatistics, University of California, Berkeley, CA
2Department of Statistics, University of California, Berkeley, CA

This Week

#OHDSISocialShowcase

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(WEDNESDAY)

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(Xiaoyu Wang, Jenna Reps, Anthony Sena, James Gilbert, Marc Suchard)

#OHDSI

www.ohdsi.org

#JoinTheJourney

Real World Evidence Using Measurement Values for Patient-Level Prediction Models: A Feasibility Study

Xiaoyu Wang1, Jenna Reps2, Anthony Sena2, James Gilbert2, Marc A Suchard2

1Division of Biostatistics, University of California, Berkeley, CA
2Department of Statistics, University of California, Berkeley, CA

This Week

#OHDSISocialShowcase

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(WEDNESDAY)

Incorporating measurement values into patient-level prediction with missing entries: a feasibility study

(Xiaoyu Wang, Jenna Reps, Anthony Sena, James Gilbert, Marc Suchard)
The Use of the Julia Programming Language for Global Health Informatics and Observational Health Research

**Presented by:** Jacob S. Zelko

**Introduction:**
Performing real-world data-based studies in the healthcare domain has become increasingly common. The Julia programming language is one of the most promising languages for data science and scientific computing. It is designed to be easy to learn and use, while still providing high performance and scalability.

**Methods:**
This paper describes a project to create an open-source Julia package for data analysis in the field of global health informatics and observational health research.

**Julia Ecosystem Tools:**
- **OMOP:** A comprehensive toolkit for working with data in the OMOP CDM
- **Data Cleaning and Transformation:** Tools for cleaning and transforming data
- **Plotting and Visualization:** Tools for creating visualizations and plots
- **Simulation:** Tools for simulating data

**General Julia Ecosystem Tools:**
- **OMOP:** A comprehensive toolkit for working with data in the OMOP CDM
- **Data Cleaning and Transformation:** Tools for cleaning and transforming data
- **Plotting and Visualization:** Tools for creating visualizations and plots
- **Simulation:** Tools for simulating data

**Sample PhenoDefinition:**
- **Definition:** A definition of a phenotype based on a combination of data from different sources
- **Implementation:** Implementation of the definition in the OMOP CDM
- **Validation:** Validation of the definition in the OMOP CDM

**Calculating Crude Prevalence Rates:**
- **Definition:** The proportion of individuals in a population who have a specific condition
- **Implementation:** Implementation of the definition in the OMOP CDM
- **Validation:** Validation of the definition in the OMOP CDM

**Final Results:**
- **Presentation:** A summary of the project's goals and outcomes
- **Discussion:** A discussion of the project's impact and implications
- **Acknowledgments:** Acknowledgment of contributions from team members and partners

**References:**
- Leemor Jenkins, author2, author3, author4, author5, author6, author7, author8, author9, author10, author11, author12, author13, author14, author15, author16, author17, author18, author19, author20, author21, author22, author23, author24, author25, author26, author27, author28, author29, author30, author31, author32, author33, author34, author35, author36, author37, author38, author39, author40, author41, author42

---

**Preparation Synthetic OMOP CDM Data for Analysis:**

**Finding Patients Who Have Had Strep Throat:**

**Characterizing Patients That Have Had Strep Throat:**

**Discussion:**
- FunSQL
- Age groupings
- [b, b + 4] for in 0:5:119

**Next Steps:**
- Key Takeaways, Summarized:
- **Acknowledgments:**
- **References:**
FRIDAY

Analyzing a Tabloid Headline with Real-World Data: A Summer Intern’s Investigation

(Delia Harms, Kristin Kostka)

With **misinformation** on the rise, OHDSI tools help us **interrogate headlines** and understand **real-world prevalence rates**.
Opening: Three Positions at Gilead

Sr. Director, Head of Data Office

Job Description:
As a Senior Director in our Data Office, you will play a pivotal role in shaping and executing our data strategy. In this leadership position, you will oversee and drive activities related to data sharing, governance, and access across the organization. Working closely with cross-functional teams, you will define and implement data acquisition policies and practices, ensuring the efficient and effective use of data to support our scientific and business objectives.

Director, Data Acquisition - Clinical Data Science

This role reports to the Head of Gilead data office, RWE Generation, Clinical Data Science and is based at different Gilead sites. This individual has responsibility for acquiring all data across clinical, development, medical affairs function and Gilead affiliates. This individual will work in close collaboration with the Development organization, Commercial, Procurement, Medical Affairs, IT, and other functions at Gilead in implementing data acquisition processes and is expected to operate with a “one Gilead” mindset & play a key role in the global Gilead Data Office set up.

Director, RWE - Data Science - OHDSI

Responsibilities:
Collaborate with researchers and data scientists to understand project requirements and translate them into OHDSI-compatible solutions. Work with databases, ensuring data integrity and optimization for OHDSI-related queries and analyses. Perform data analyses in OHDSI-related tools like ATLAS. Customize and extend OHDSI tools and applications to meet specific project needs. Collaborate with cross-functional teams to troubleshoot and resolve technical issues related to OHDSI implementations. Stay informed about OHDSI community updates, best practices, and emerging trends in observational health data research. Contribute to the development and documentation of data standards and conventions within the OHDSI community.
The Zhang Lab at Washington University School of Medicine in St. Louis has one postdoc/senior data analyst position to work on causal machine learning and responsible AI for reliable real-world evidence generation.

- More details at https://linyingzhang.com
  - Postdoc: https://linyingzhang.com/files/Postdoc.pdf
  - Data analyst: https://linyingzhang.com/files/Analyst.pdf
- If interested, please send CV and cover letter to linyingz@wustl.edu
Opening: Epidemiology UX/Web Design Intern at J&J

Epidemiology UX/Web Design Intern

**JOB TITLE**  Epidemiology UX/Web Design Intern

**FUNCTION**  Career Programs

**SUB FUNCTION**  Non-LDP Intern/Co-Op

**LOCATION**  Raritan, New Jersey, United States

**DATE POSTED**  Jan 19 2024

**REQUISITION NUMBER**  2406163977W

**DESCRIPTION**

Janssen Research & Development, L.L.C., a division of Johnson & Johnson's Family of Companies is recruiting for Epidemiology UX/Web Design Intern. This position is a member of the Observational Health Data Analytics (OHDA) team. OHDA's mission is to improve the lives of individuals and quality of healthcare by efficiently generating real-world evidence from the world's observational health data, transparently disseminating evidence-based insights to real-world decision-makers, and objectively advancing the science and technology behind reliable evidence.
Opening: Research Information Specialist at UNC

Responsibilities include:

* Perform SQL-based programming against UNC’s clinical data warehouse to identify patient cohorts and develop patient datasets.
* Consult with and collaborate with researchers to ensure programming work aligns with project needs.
* Develop ETL (extract, transform, and load) and data integration processes to support common data models (OMOP, PCORnet) using appropriate technologies (SQL, Python, etc.).
* Carefully following UNC’s regulatory and governance policy to ensure data integrity and security.
* In collaboration with IDSci team, identify potential enhancements in current workflows and data architecture.
* Implement quality assurance strategies, such as data validation and peer code review.
* Write and maintain up-to-date supporting documentation. Ensure code is well-commented and use GitLab/GitHub to manage code changes and track data lineage.
* Provide technical leadership and direction for assigned projects and/or data requests.

Master’s and 1-2 years’ experience; or Bachelors and 2-4 years’ experience; or will accept a combination of related education and experience in substitution.

This position requires two or more years of relevant work experience and:

* Expert-level knowledge of SQL programming, data modeling, and relational database systems such as Oracle, Microsoft SQL Server, MySQL, etc.
* Demonstrable past experience in supporting technical projects in terms of length of time, competencies and cost. Individual will be expected to manage multiple projects at one time while delivering high-quality work on time.
* Excellent written and oral business communication skills. Public speaking at meetings and conferences may be required. The ability to clearly convey technical concepts to non-technical clients is a must.

Minimum Education and Experience Requirements:

- Position Number: 2060092
- Vacancy ID: N90007640
- Full Time/Part Time: Full-Time Permanent
- FTE: 1
Opening: Data Steward at EBMD

Description

Are you looking for a job where you can make a difference and work in a non-profit? Would you like to be a part of an ambitious and international organisation on the cutting edge of science? Then this position might be right up your alley.

The EBMT is a non-profit medical and scientific organisation which hosts a unique patient registry providing a pool of data to perform studies and assess new trends.

OUR MISSION
Save and improve the lives of patients with blood-related disorders.

The Registry
Holding the data of over half a million patients, the EBMT registry is the starting point for all studies carried out through the EBMT working parties. The department focuses on data collection processes, data quality monitoring, and maintenance of the database.

YOUR MISSION
Responsible for collecting, collating, and evaluating issues and problems with data and enforcing data usage policies.

RESPONSIBILITIES AND TASKS

Data Stewardship:
- Design, implementation and testing of new data collection processes including data collection forms (DCFs) development.
- Take care of the mapping of new items from DCFs to the OMOP CDM
- Providing input on data quality reports
- Check and clean data on request and ad hoc.
- Data retrieval including designing data reports and data report running.
- Carry out computerized system validation activities.
- Supporting consolidation/harmonization of data
- Creating standard data definitions, and maintain a consistent use of data assets across the organization
- Documenting data policies and data standards
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?
Learn more about all of the OHDSI workgroups

ohdsi.org/workgroups
Common Data Model Workgroup

2024 OKR Update
Purpose

The CDM workgroup exists to maintain and improve the use of the OMOP Common Data Model to make it the premier observational health data model in the world. We ensure the integrity and usability of the OMOP CDM in relation to other working groups by providing guidance on data standardization best practices.
2024 Objectives and Key Results

Objective 1: Facilitate collaboration and alignment between the CDM and other OHDSI working groups

• Host a hack-a-thon to collaborate with THEMIS, DQD, and Vocabulary WGs, aligning the community on data standards, conventions, and evaluation
  • Clarify data standardization best practices and share with other workgroups
  • Document prior decisions made by the CDM WG
2024 Objectives and Key Results

Objective 2: Make the OMOP CDM the premier observation health data model by reducing technical debt and improving documentation

- Get the CDM package onto CRAN
- Clean up existing documentation and remove outdated documentation
  - Document the STEM table and clarify its usage
  - Remove CDM v6.0 from website
  - Write down add-on, extension, expansion information
  - Write down our maturity model
CRAN Achieved!

CommonDataModel: OMOP CDM DDL and Documentation Generator

Generates the scripts required to create an Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) database and associated documentation for written in parameterized Structured Query Language (SQL) to the other supported dialects.

Version: 0.2.0
Depends: DatabaseConnector, SqlRender, rJava
Imports: rmarkdown, stringr, DBI, dplyr, readr
Suggests: knitr, testthat (≥ 3.0.0), RSQLite, withr
Published: 2024-02-07
Author: Clair Blacketer [aut, cre]
Maintainer: Clair Blacketer <mblacke@its.jnj.com>
License: Apache License 2.0
NeedsCompilation: no
Materials: README
CRAN checks: CommonDataModel results

Documentation:
Reference manual: CommonDataModel.pdf

Downloads:
Package source: CommonDataModel_0.2.0.tar.gz
Windows binaries: r-devel: CommonDataModel_0.2.0.zip, r-release: CommonDataModel_0.2.0.zip, r-oldrel: CommonDataModel_0.2.0.zip
macOS binaries: r-release (arm64): CommonDataModel_0.2.0.tar.gz, r-oldrel (arm64): CommonDataModel_0.2.0.tar.gz, r-release (x86_64): CommonDataModel_0.2.0.tar.gz

Linking:
Please use the canonical form https://CRAN.R-project.org/package=CommonDataModel to link to this page.
Network Data Quality Workgroup

2024 OKR Update
The Network Data Quality workgroup exists to recommend, enable, and develop best practices related to observational data quality at the level of a federated network.
2024 Objectives and Key Results

Objective 1: Improve Data Quality reporting for the OHDSI Community

• Complete information pages for all check types in the DQD – Q1
• Create a new Data Quality report that informs the user on how to interpret and remediate failing DQD checks
  • Draft of report in Q1,
  • Inform and work on refactor in Q2
• Create at least 1 new DQD check as identified by THEMIS
Objective 2: Support and collaborate with THEMIS and CDM working groups

• Specify the requirements necessary to make THEMIS conventions assessable and reportable as data quality checks
• Create at least 1 new DQD check as identified by THEMIS
Objective 3: Refine the approach for quantitatively assessing a OHDSI Network datasets' fitness for specific study questions.

- Define "fitness for use" in the context of an OHDSI network study
- Conduct an assessment of data diagnostics and provide a report of potential improvements (Q1)
- Prioritize improvements and implement the top prioritized features to data diagnostics (Q2 and beyond)
OHDSI APAC OKRs

Mui Van Zandt
2023 OHDSI APAC Key Results

**Research**

Build research expertise and collaboration amongst the different chapters through publication

- **10 Scientific Forums**
  - [JAMA Open](https://www.jamaopen.com)
  - [41 Publications](https://www.jamaopen.com)

**Training**

Create an APAC training program to expand reach to the general community

- **OHDSI SOS Challenge 2023**
  - Study overview: Is fluoroquinolone use associated with the development of aortic aneurysms and aortic dissections?
  - Chief Investigators:
    - Song Han, June Hove, Jing Li, Xinyu Qu, Limin Li - Fudan University
    - Jill Barrett, Guangzhou

**Communication**

Create collaboration activities that encourage collaborative generation and dissemination of evidence that promotes better health decisions and better care

- **10 Community Calls**
  - [Community Calls](https://www.ohdsi.org)
- **2 Titan Awards**
  - [Titan Awards](https://www.ohdsi.org)
- **4 Newsletters**
  - [Newsletters](https://www.ohdsi.org)
- **1 APAC Symposium**
  - [APAC Symposium](https://www.ohdsi.org)
2024 OHDSI APAC Goals

**Research**

Build research expertise and collaboration amongst the different chapters through publication

**Milestones**
- Conduct APAC SOS Challenge studies
- Replicate Cindy Kai’s SOS Challenge study

**Training**

Create an APAC training program to expand reach to the general community

**Milestones**
- Host at least 2 in-person trainings in APAC
- Train community through APAC SOS Challenge studies

**Communication**

Create collaboration activities that encourage collaborative generation and dissemination of evidence that promotes better health decisions and better care

**Milestones**
- Host APAC symposium
- Distribute quarterly newsletters
- Host monthly community calls and scientific forums
Join Us

Sign up for the OHDSI APAC WG!

• APAC Community Calls
  • Every third Thursday, 12 p.m. Korea time

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 18</td>
<td>APAC 2024 Kickoff</td>
</tr>
<tr>
<td></td>
<td>Training Session #7 by Japan</td>
</tr>
<tr>
<td></td>
<td>Eye Care and Vision Research WG Intro</td>
</tr>
<tr>
<td>Feb 15</td>
<td>New WG Intro: Evidence Translation, Industry</td>
</tr>
<tr>
<td>Mar 21</td>
<td>Vocabulary Contribution by Korea</td>
</tr>
<tr>
<td></td>
<td>OHDSI Evidence Network</td>
</tr>
<tr>
<td>Apr 18</td>
<td>Newcomers Session</td>
</tr>
</tbody>
</table>

• APAC Scientific Forum
  • Every first Thursday, 12 p.m. Korea time

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb 1</td>
<td>Survey Results and Plans for 2024</td>
</tr>
<tr>
<td></td>
<td>APAC Study Updates</td>
</tr>
<tr>
<td>Mar 7</td>
<td>Perseus Intro &amp; Demo</td>
</tr>
<tr>
<td>Apr 4</td>
<td>Genomic Data Mapping</td>
</tr>
</tbody>
</table>

Direct link to community calls

Direct link to scientific forums
OHDSI Industry Working Group
OKRs
OHDSI Industry Work Group

Who should attend?

This is an open meeting with a focus on those members of the OHDSI community who have ties and affiliations with the Pharma and Biotech industries and would like to work together to represent those interests more broadly within OHDSI.

- Foster a stronger collaboration between the life science, pharma, and biotech industries, and the OHDSI community.

- Identify and develop strategies to encourage the active participation of these industries in OHDSI studies and initiatives.

- Facilitate knowledge transfer, sharing industry expertise and learnings with the broader OHDSI community.

- Identify opportunities for mutual support, leveraging industry resources and capabilities to advance OHDSI's goals.

- Increase the visibility and understanding of OHDSI's initiatives within these industries, promoting active involvement and commitment.
2024 OHDSI Industry Working Group Goals

Data Marketplace
- Develop OMOP data marketplace and supporting framework

Go-To-Market
- OHDSI advocacy group

Collaboration
- Use case identification and OHDSI working group partnerships

Sign up for the OHDSI Industry WG!
2024 OHDSI Industry Working Group Goals

**Data Marketplace**

- Develop OMOP data marketplace and supporting framework

  **Milestones**
  - Create a catalogue of OMOP datasets that are open for industry sponsored studies
  - Develop framework for interacting with the marketplace

**Go-To-Market**

- Design structure and purpose of OHDSI advocacy group and ensure interoperability

  **Milestones**
  - Design structure and purpose of OHDSI ‘advocacy group’
  - Create recommendations for OHDSI/OMOP models to be regulatory/governmentally aligned – ISO standards etc

**Collaboration**

- Collaboration with other working groups to build use cases specific to industry

  **Milestones**
  - Identify 2-3 use cases
  - Identify work group partnerships for each use case
Eye Care and Vision Research Workgroup

- Workgroup purpose: The purpose of the Eye Care and Vision Research Workgroup is to advance the development and implementation of data standards in ophthalmology, optometry, and the vision sciences, and to support studies using observational ophthalmic data for generating insights to improve health and vision outcomes.

- Workgroup past accomplishments:
  - Published a gap analysis of two large, well-known EHR systems for eye care (Epic and Cerner).
  - In addition to standing monthly meetings, organized in-person meetings at major conferences.
  - Organized additional subgroups: retina, glaucoma, pediatrics, uveitis, imaging and ETLs.
  - Collaborated with Verana Health for OMOP transformation of the AAO IRIS Registry.
  - Partnered with the NIH Bridge2AI AI-READI project to map ophthalmic data elements; pilot public release in spring 2024.
  - Submitted retinal condition codes to SNOMED International.
  - Submitted glaucoma examination codes to SNOMED International.
  - Submitted uveitis phenotypes to HowOften.
  - Supported SOS Challenge project examining the risk of kidney injury associated with anti-VEGF.
  - Engaged with LOINC to develop framework for representing visual acuity data.
  - Started working on ETLs of ophthalmic data at several participating sites.
Eye Care and Vision Research Workgroup OKRs

**Objective 1: Continue advancing data standards development around specific use cases**

Key Result 1: Build upon prior success with developing tonometry-related concepts with the glaucoma subgroup and advance representation of additional concepts relevant for glaucoma research, including gonioscopy-related concepts and visual field concepts. Timeline – end of Q2 2024.

Key Result 2: Submit diabetic retinopathy phenotype-related concepts from the retina subgroup to SNOMED for subsequent incorporation into the CDM. Timeline – end of Q1 2024.

Key Result 3: Contribute to public release of pilot data from the AI-READI Bridge2AI project, which includes ophthalmic data element mapped to standard OMOP concepts. Timeline – end of Q2 2024.

**Objective 2: Map common ophthalmic data elements at multiple institutions**

Key Result 1: Submit visual acuity codes to LOINC using panel approach. Timeline – end of Q1 2024.

Key Result 2: Trial ETL processes at 3 institutions for visual acuity data. Timeline – end of Q4 2024.

Key Result 3: Trial ETL processes at 3 institutions for IOP data. Timeline – end of Q2 2024.

**Objective 3: Develop long-term sustainability to workgroup efforts.**

Key Result 1: Organize grant-writing committee to plan proposals. Timeline – end of Q2 2024.

Key Result 2: Submit grant for funding data network. Timeline – end of Q4 2024.
Surgery and Perioperative WG

Objectives and key results 2024
Feb 13 2024
OKR Themes

• WG Growth and Processes
• Cohorts & Characterization
• Community Events / Evidence Generation
• Perioperative Prediction
OKR Theme: WG Growth and Processes

• Strategic WG Growth
• Key Results
  • Involvement of (at least) 3 new members from (at least 3 different) surgical / perioperative science focused lab groups
  • Presentation of WG supported work in at least one surgical / perioperative medicine conference.
  • Presentation of broader OHDS mission and capabilities within at least one surgical / perioperative conference.
  • Establishing 2 strategic collaborations with other WGs: 2 joint meetings during 2024
Theme: Cohorts & Characterization

• Completion of HowOften Incidence Rate Characterization Studies
  • Key Results: Submission for publication
    • Surgical cohorts against post operative outcomes of interest
    • CRC against post operative outcomes of interest
    • Surgical cohorts with post op afib, ischemic stroke.

• Completion of Fragility Fracture Study thon
  • Key Result: Submission for publication
Theme: Cohorts & Characterization

- Exploration of proxies for pre / post operative functional outcomes within the OHDSI network
  - Key Results:
    - Compile list of existing proxies within the vocabulary, determine use in the network
    - Survey WG / OHDSI network with respect to existence of proxy data at their site.
Theme: Community Events
(Cohorts & Characterization)

• Support one surgical cohort hack-a-thon
  • Key Results:
    • Complete the extension of HowOften Surgical Cohorts into standard OHDSI vocabulary representations
    • Creation of at least 3 denovo surgical cohorts during the hackathon.
    • Initiate 3 members new to cohort building in OHDSI into the cohort Building process
Theme: Perioperative Prediction

• Execution of a Perioperative Prediction study (jointly with PLP workgroup) using the Major Non-Cardiac Surgery cohort; other surgical risk cohorts, and outcomes of post operative interest.

  • Key Results:
    • Generation of study design, and github page ready to launch study
    • Execution of PLP network study on at least 3 OHDSI data sources.
    • Preparation of at least 1 draft manuscript (Q4 2024)
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