Workgroup 2023 OKRs and Phenotype Phebruary Updates

OHDSI Community Call

Feb. 7, 2023 • 11 am ET

OHDSI is a GLOBAL community

OHDSI Data Partners

What does it take to be an OHDSI data partner? Anyone with access to observational data can standardize their database in the OMOP Common Data Model, apply OHDSI's open-source tools, and participate in collaborative research.

Who has already joined the journey and adopted the OMOP CDM? There are currently 453 databases, including 374 electronic health records, 34 registries and 30 administrative claims sources, that come from 41 different countries. Together, these databases represent more than 928 million unique patient records, approximately 12% of the world's population.

Australia (14)
ACA National Joint Replacement Registry
AU-eFIBRN (Australian Electronic practice based research network)
AUS Department of Veterans Affairs
Austin Health
iQVIA Australia LPD
Melbourne Children's Hospital
NPS MedicineWise
Pharmaceutical Benefits Scheme 10% extract
Primary Care GP data (Patron)
Royal Melbourne Hospital and Western Health Hospital Admissions
South Western Sydney LHD
Sydney Children's Hospital
Sydney Local Health District (LHD)
University of Queensland - Queensland Health

Austria (1)
Medical University of Vienna

Belgium (18)
AZ Damiaan Oostende
AZ Delta
AZ Klina
AZ Maria Middelares
Icometrix
IQVIA Belgium LPD
LynxCare
Mediahuis
Onze-Lieve-Vrouwziekenhuis Aalst-Asse-Ninove
THIN BE
Universitaire Ziekenhuizen KU Leuven
University Hospital Antwerp
University MS Center
University MS Center
UZ Brussel
UZ Leuven
VZW AZ Groeninge
Ziekenhuis Oost-Limburg

Bosnia and Herzegovina (1)
E-MEDIT D.O.O. & Hospital Travnik

Brazil (4)
Centro de Health Data and Knowledge Integration - Cidacs
DataSUS Ambulatory
Hospital Israelita Albert Einstein
IQVIA Brazil

Bulgaria (2)
National Scientific Programme "E-Health in Bulgaria"
SAT Health

Canada (3)
IQVIA Canada EMR
Provincial Health Services Authority (British Columbia)
The Hospital for Sick Children

China (7)
Beijing Anding Psychiatry Hospital
Beijing Shirendu Medical Science & Technology CO., Ltd.
Hebei Province Psychiatry Hospital
Jiangsu Province People's Hospital
Nanfang Hospital COVID-19 Research Database (NFHCHD)
Tianjin Anding Psychiatry Hospital
Wonders Information

Colombia (1)
Hospital Universidad del Norte

Croatia (7)
Bács-Kiskun Megyei Kórház a Szegedi Tudományegyetem Általános Orvostudományi Kar Oktató Kórháza
Clinical Hospital Dubrava
Hrasnička & University Hospital Centre Zagreb
IQEA d.o.o. & University Hospital Center Sestre milosrdnice
IN2 d.o.o. & Clinical Hospital Center Osijek
MCS Grupa d.o.o. & Health Care Center of Primorje-Gorski Kotar County
Szabolcs-Szatmár-Bereg Megyei Kórházak és Egyetemi Oktatókórház

Czechia (3)
Czech Myeloma Group
Institute of Rheumatology
OAKS Consulting s.r.o.

Denmark (2)
Aarhus University Hospital Database
Center for Surgical Science (CSS)

Estonia (2)
Estonian Genome Center at the University of Tartu (EGCUT)
University of Tartu

Finland (10)
Auria Clinical Informatics
BCB Medical Ltd.
Finnish Clinical Biobank Tampere
Finnish Hematology Registry/ HUS
Finnish Institute for Health and Welfare (THL)
Hospital District of Helsinki and Uusimaa
Hospital District of Southwest Finland
HUS Datalake eCare/forMe PCC
Pirkanmaa Hospital District
University of Turku (Prostate Cancer Registry of South West Finland)

France (13)
APHP-EDS
Assistance Publique - Hopitaux de Marseille
Assistance Publique - Hôpitaux de Paris (AP-HP)
Bordeaux University Hospital
CEGEDIM HEALTH DATA
Centre Hospitalier Universitaire de Lille
Centre Hospitalier Universitaire de Montpellier
Centre Hospitalier Universitaire de Toulouse
IQVIA France DA
IQVIA France LPD
Lille University Hospital
SNDS
THIN FR

Germany (11)
CancerDataNet GmbH
Charité - Universitätsmedizin Berlin
European Rare Kidney Disease Registry (ERKReg)
GermanOncology
Hanover Medical School, Germany
IQVIA Germany DA
Knebergstift Rheinland-Pfalz
MS Forschungs- und Projektentwickl-ungs-gGmbH
UKER
University Medicine Dresden
University of Ulm, ZBMT

Greece (4)
Diagnostic & Therapeutic Center Of Athens "Hygeia" Single Member Societe Anonyme
General Hospital of Kavala
Innovative Medical Research SA
Papageorgiou General Hospital

Hungary (2)
Semmelweis University
University of Pécs

India (1)
Buddhimed Technologies

Ireland (1)
Trinity St James's Cancer Institute, Dublin

Israel (4)
Hadassah OBGYN
Kinneret (Ministry of Health medical center network)
Lowese
The Directorate of Government Medical Centers at the Israeli Ministry Of Health

Italy (28)
Agenzia regionale di sanità della Toscana (ARS)
AO Card. G. Panico - Center for Neurodegenerative Diseases and Aging Brain
ASL Roma 1
ATS Bergamo
Azienda Ospedaliera SS Antonio e Biagio e Cesare Arrigo
Azienda Ospedaliera Universitaria Integrata Verona
AZIENDA OSPEDALIERO UNIVERSITARIA SAN LUIGI GONZAGA
Azienda Ospedaliero-Universitaria di Modena
Bambino Gesù Children's Hospital
Basilicata Cancer Registry
Casa di Cura Privata del Policlinico (CCPP)
Fondazione Casa Sollievo della Sofferenza
Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico
Fondazione IRCCS Istituto Neurologico Carlo Besta
Fondazione IRCCS Policlinico San Matteo
Fondazione Istituto Nazionale dei Tumori
Fondazione Policlinico Universitario Agostino Gemelli IRCCS
FONDAZIONE TOSCANA GABRIELE MONASTERIO PER LA RICERCA MEDICA E DI SANITÀ PUBBLICA (FTGM)
Grande Ospedale Metropolitano "Bianchi-Melacrino-Morelli"
IQVIA Italy LPD
IRCCS Azienda Ospedaliero-Universitaria di Bologna Policlinico di Sant'Orsola
IRCCS Policlinico San Donato
ISMETT
Modena Oncology Center - Azienda Ospedaliera Modena
Pediatrel
Società Italiana di Medicina Generale e delle cure Primarie (SIMG)
University Hospital of Parma

Japan (4)
IQVIA Japan Claims
IQVIA Japan HIS
Japan Medical Data Center (JMDC)
MDV (Medical Data Vision)

Luxembourg (1)
Registre National du Cancer du Luxembourg

Montenegro (1)
Clinical Center of Montenegro

Netherlands (12)
Amsterdam UMC
EBMT The European Society for Blood and Marrow Transplantation
European Clinical Research Alliance on Infectious Diseases (ECRAID) and University Medical Center Utrecht (UMCU)
Harm Stijger
IKNL
Integrated Primary Care Information (IPCI)
National Intensive Care Evaluation foundation
Netherlands Cancer Registry
Pharmo
Stichting VUmc
STIZON
VieCuri Medisch Centrum

Norway (2)
The Norwegian Cancer Registry
University Of Oslo

Portugal (11)
APDP
Centro Clínico Académico e Braga, Associação (2CA-Braga)
Centro Hospitalar Universitário de Coimbra (CHUC)
CUF
EGAS MONIZ HEALTH ALLIANCE
Hospital da Luz Learning Health
Hospital Distrital de Santarém (HDS)
Hospital do Espírito Santo de Évora
Instituto de Medicina Molecular
Registo Português de Doentes Reumáticos
Unidade Local de Saúde de Matosinhos

Republic of Korea (59)
Ajou University Hospital
Asan Medical Center
Bucheon Sejong Hospital
Catholic Kwandong University International ST. Mary's Hospital
Cha University Bundang Medical Center
Chonnam National University Hwasun Hospital
Chonnan National University Hospital
Chungnam National University Hospital
Chungnam National University Sejong Hospital
Daegu Catholic University Medical Center
Dankook University Hospital
Dongguk University Medical Center
Ewha Womans University Medical Center (Mokdong)
Ewha Womans University Medical Center (Seoul)
Gachon University Gil Medical CenterGachon University-Gil Medical Center
Gangnam Severance Hospital
Gangneung Asan Hospital
Gyeongsang National University Changwon Hospital
Gyeongsang National University Hospital
Hanyang University Seoul Hospital
Health Insurance Review & Assessment Service
Incheon Sejong Hospital
Inha University Hospital
Jeonbuk National University Hospital
Kangbuk Samsung Hospital
Kangdong Sacred Heart Hospital
Kangwon National University Hospital
Konkuk University Medical Center
Kyungpook National University Hospital
Korea Institute of Radiological & Medical Sciences
Korea University Anam Hospital
Korea University Ansan Hospital
Korea University Guro Hospital
Kyung Hee University Hospital At Gangdong
Kyung Hee University Medical Center
Kyungpook National University Chilgok Hospital
Kyungpook national university hospital
Myongji Hospital
Myongji Hospital (Jecheon)
National Cancer Center
National Health Insurance Service
National Health Insurance Service Ilsan Hospital
Pusan National University Hospital
Samsungmedical Center
Seoul National University Bundang Hospital
Seoul National University Hospital
Severance Hospital
Soonchunghyang University Hospital (Bucheon)
Soonchunghyang University Hospital (Cheonan)
Soonchunghyang University Hospital (Gumi)
Soonchunghyang University HospitalSeoul)
The Catholic University of Korea, Seoul ST. Mary's Hospital
The Catholic University of Korea, ST. Vincent's Hospital
The Catholic University of Korea, Uijeongbu ST. Mary's Hospital
The Catholic University of Korea, Yeouido ST. Mary's Hospital
Ulsan University Hospital
Wonju Severance Christian Hospital
Wonkwang University Hospital
Yongin Severance Hospital

Romania (1)
This Rs

Saudi Arabia (1)
Saudi Food and Drug Authority

Scotland (1)
HIC Dundee

Serbia (5)
Clinical-hospital center Zvezdara
Klinično-bolnični center Zvezdara (Clinical-hospital center Zvezdara)
Primary Healthcare Center Zemun
University Clinical Center of NiŠ
University Clinical Center of Serbia

Singapore (3)
Growing Up in Singapore Towards healthy Outcomes (GUSTO)
Khoo Teck Puat Hospital (SQ_KTPH)
National University Hospital Singapore

Spain (37)
Agencia Española de Medicamentos y Productos Sanitarios, AEMPS
BIOCRUCES BIZKAIA HEALTH RESEARCH INSTITUTE
Consellería de Sanidade
Consorci Corporació Sanitària Parc Taulí
Consorci Mar Parc de Salut de Barcelona (PSMAR)
CORPORACIÓ SANITARIA PARC TAULÍ
FISABIO-HSRU
Fundació Institut d'Investigació Sanitària Illes Balears
Fundació Institut d'Investigacions Mèdiques (FIMIM)
Fundacion de Investigacion Biomedica del Hospital Universitario 12 de Octubre
Fundación para la Investigación Biomedica (INCLIVA)
FUNDACION PARA LA INVESTIGACION DEL HOSPITAL UNIVERSITARIO LA FE DE LA COMUNIDAD VALENCIANA (HULAFE)
Fundación para la Investigación del Hospital Universitario La Fe de la Comunidad Valenciana (HULAFE)
Fundación para la Investigación e Innovación Biosanitaria en Atención Primaria (FIIBAP)
Healthcare Service of the Principality of Asturias
HM Hospitals
Hospital del Mar (HMAR)
Hospital Sant Joan de Déu
Hospital Universitario 12 de Octubre
INFOSANCO12
Information System of Parc de Salut Mar (IMASIS)
Institut Català d'Oncologia
Instituto Aragonés de Ciencias de la Salud (IACS)
IQVIA Spain LPD
Marina Salud (Hospital de Denia)
Parc Sanitari Sant Joan de Déu
Pedro Mallol
Research Institute - Hospital de la Santa Creu i Sant Pau
Riqa Salud
Servicio Cántabro de Salud and IDIVAL
Servicio Madrileño de Salud
Servicio Navarro de Salud Osasunbidea (SNS-O)
The Information System for Research in Primary Care
The Information System for Research in Primary Care – Hospitalization Linked Data (SIDIAP-H)
Vall d'Hebron Hospital Campus
Vall d'Hebron Institut Campus
Virgen Macarena University Hospital

Sweden (2)
MEB KI
Switeng

Switzerland (5)
CancerDataNet

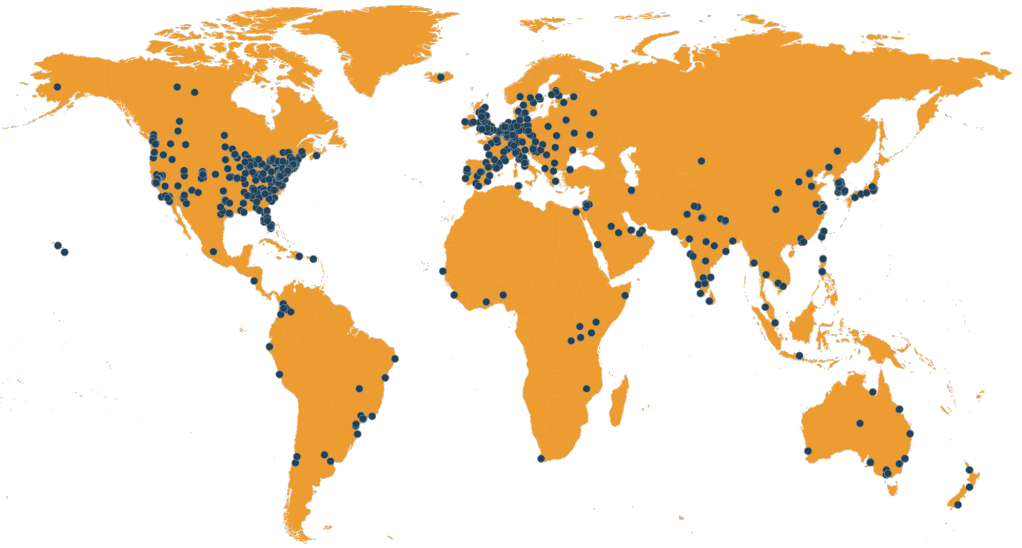
Data2time
Geneva Cancer Registry
HUG and SCOM
Institute of Social and Preventive Medicine, University of Bern

Taiwan (8)
NHIRD
Shuang Ho Hospital
Taipei Medical University Clinical Research Database (TMUCRD)
Taipei Medical University Hospital
Wanfang Hospital

Turkey (2)
Istanbul University Istanbul Faculty of Medicine
IUC Cerrahpaşa TIP Fakültesi

United Kingdom (26)
Aletheia Health
Barts Health NHS Trust
Clinical Practice Research Datalink (CPRD GOLD)
Clinical Practice Research Datalink Aurum (CPRD Aurum)
Connected Bradford
DataLoch
GOSH
Harvey Walsh Ltd
Kings College London
Leeds Teaching Hospitals
OPEN Health
Optimum Patient Care Limited
Queen Mary University of London
Royal College of General Practitioners Research and Surveillance Centre
SAIL Databank
THIN UK
UCL
UK Biobank
UK integrated Medical Record Database (iMRD) THIN
UK National Neonatal Research Database
UKCRIS
University College London CALIBER
University College London Hospitals
University College London Hospitals NHS Foundation Trust
University of Edinburgh
University of Edinburgh DataLoch

United States (136)
Tuji Health
Advocate Aurora Health & University of Madison Health Non-Muscle Invasive Bladder Cancer
Advocate Aurora Health COVID Database
All of Us Research Program
ALTAMED (University of Southern California)
Atrium – Wake Forest Baptist Health
Blue Health Intelligence
Boston Medical Center
Brown University - Rhode Island HIE
C-Path
Carilion Clinic
Case Western
Cerner HealthFacts
Cherokee Health Systems
Children's Hospital of Colorado
Children's Hospital of Los Angeles
Children's Hospital of Philadelphia
Children's National
Columbia University Irving Medical Center
CRHFEI
DARTnet Institute: CERD Study
Decision Resources Group (DRG)
Department of Health Services - Los Angeles
Duke University
Eau Claire Cooperative Health Center
Flatiron - OSCER
Geisinger Health System
George Washington University
Georgetown University ARIA
Georgia Tech Research Institute
GenOMOP
Harvard University Mass General Brigham
HealthVerity
IBM(R) MarketScan(R) Commercial Claims (CCAE)
IBM(R) MarketScan(R) Medicare Supplemental Database (MDCR)
IBM(R) MarketScan(R) Multi-State Medicaid Database (MDCD)
Icahn School of Medicine at Mount Sinai
Indiana University School of Medicine / Regenstrief Institute
Inova Health System
IQVIA US Ambulatory EMR
IQVIA US Hospital Charge Data Master (CDM)
IQVIA US Oncology EMR
IQVIA US Open Claims
IQVIA US PharMetrics Plus
Johns Hopkins University
Keck Medicine of University of Southern California
Loyola University New Orleans
Maine Medical Center
Mayo Clinic
Medical University of South Carolina
Medicare Research Identifiable Files
MedStar Health
Memorial Sloan Kettering Cancer Center
Mercer/Cox AD
Montefiore Medical Center (Albert Einstein College of Medicine)
NSC
Nemours Children's Health System
NorthShore University HealthSystem
Northwestern Medicine Enterprise Data Warehouse (NMEDW)
NYC-CDRH
NYU Langone
OCHIN (Oregon Community Health Information Network)
Ochsner Medical Center
Oklahoma University
One Fact Foundation Payless Health
OptumRI De-Identified Clinformatics(R) Data Mart Database - SES & DOD
OptumRI de-identified Electronic Health Record Dataset (PANTHER)
Oregon Health & Science University
Panalo Intelligence
PEDSnet
Penn State
Premier Healthcare Database
QuanaCare - Los Angeles
Reliant Medical Group
Rhode Island Quality Institute
Rush University Medical Center
Rutgers
Shriners Children's
Spectrum Health West Michigan
STAnford medicine Research data Repository (STARR)
Stony Brook
Surveillance, Epidemiology, and End Results Program (SEER) 8-Cad
TCC - Los Angeles
The Healthcare Cost and Utilization Project (HCUP), Nationwide Inpatient Sample (NIS)
The National Health and Nutrition Examination Survey (NHANES)
The Ohio State University Medical Center
TrialSpark
Tufts MC Research Data Warehouse (TRDW)
Tulane
UMass Memorial Medical Center
UNC Chapel Hill
University Medical Center New Orleans
University of Alabama at Birmingham
University of Arkansas
University of Buffalo
University of California Health
University of California, Davis
University of California, Irvine
University of California, Los Angeles
University of California, Riverside
University of California, San Diego
University of California, San Francisco
University of Chicago
University of Cincinnati
University of Colorado
University of Colorado, Anschutz Medical Center
University of Illinois Chicago
University of Iowa
University of Kentucky
University of Miami
University of Michigan
University of Minnesota
University of Mississippi Medical Center
University of Nebraska Medical Center
University of New Mexico Health Sciences Center
University of Pittsburgh
University of Pittsburgh - Banner
University of Rochester
University of Texas Houston
University of Texas Medical Branch
University of Texas Southwestern Medical Center
University of Utah
University of Virginia
University of Washington
University of Wisconsin Madison
US Department of Defense
US Department of Veterans Affairs
UTPhysicians
Vanderbilt University
Veradigm Health Insights Data - Allscripts
Veradigm Health Insights Data - Practice Fusion
Virginia Commonwealth University
Wake Forest University
WashU St Louis
Weill Cornell Medicine/NewYork-Presbyterian Hospital (East Campus)
West Virginia University
Winship Cancer Institute of Emory University



Upcoming OHDSI Community Calls

Date	Topic
Feb. 14	Phenotype Phebruary Weekly Update + Workgroup Plans for 2023
Feb. 21	Phenotype Phebruary Weekly Update + Workgroup Plans for 2023
Feb. 28	Phenotype Phebruary Weekly Update + Workgroup Plans for 2023
Mar. 7	Save Our Sisyphus (SOS) Research Idea Presentations
Mar. 14	OHDSI Debates
Mar. 21	Recent Publications
Mar. 28	SOS Week 1 Tutorial: Initiating A Network Study

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to **Sara Khalid** on recently being awarded the title of **Associate Professor at the University of Oxford**.

OHDSI Shoutouts!

Any shoutouts from the community? Please share and help promote and celebrate OHDSI work!

Have a study published? Please send to sachson@ohdsi.org so we can share during this call and on our social channels.
Let's work together to promote the collaborative work happening in OHDSI!

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	9 am	Patient-Level Prediction
Wednesday	9 am	Atlas
Wednesday	2 pm	Natural Language Processing
Wednesday	7 pm	Medical Imaging
Thursday	9:30 am	Data Quality Dashboard Development
Thursday	7 pm	Dentistry
Friday	9 am	Phenotype Development & Evaluation
Friday	9 am	GIS – Geographic Information System General
Friday	11 am	Clinical Trials
Friday	11 pm	China Chapter
Monday	10 am	Healthcare Special Interest Group
Monday	11 am	Early-Stage Researchers
Monday	4 pm	Eye Care & Vision Research
Tuesday	9 am	OMOP CDM Oncology Outreach/Research Subgroup

ohdsi.org/workgroups

Feb. 13 Speaker Series: Paul Nagy

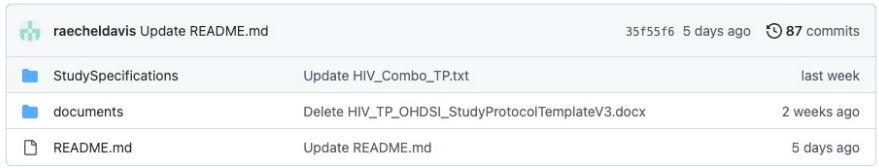
Paul Nagy will be the next guest for the Early-Stage Researchers Career Speaker Series, which will be held Monday, Feb. 13 (11 am-12 pm) inside Teams.

Learn more about Paul's path to OHDSI, his current work focusing on open-source development, medical imaging and more, and advice he has for young researchers.

bit.ly/OHDSILeaders

Join This OHDSI Network Study

Rachael Davis is leading a network study **to characterize and evaluate trends in pathways for antiretroviral therapy (ART) for individuals who have been diagnosed with Human Immunodeficiency Virus (HIV) and treated persistently over two years**. She is seeking collaborators and data partners for this study.



README.md

Characterizing Patterns in Antiretroviral Therapy for Individuals with Human Immunodeficiency Virus (HIV)

Study Status | Repo Created

- Analytics use case(s): **Characterization**
- Study type: **Clinical Application**
- Tags: **OHDSI, HIV, Antiretroviral Therapy**
- Study lead: **Raechel Davis**
- Study lead forums tag: **redavis**
- Study start date: **January 2023**
- Study end date: **-**
- Protocol: **HIV_TP**
- Publications: **-**
- Results explorer: **-**

How to run

Requirements:

- A Database in the Common Data Model (>= Version 5)
- Configured ATLAS Environment with capability to execute CohortPathways and Characterization Tools

1. Import Study Specifications

Import the following specifications for each analysis into ATLAS by copying and pasting json .txt provided into the Utilities, Import Tab for each analysis:

- Cohort Pathway 1: **Single Ingredient Antiretrovirals**
- Cohort Pathway 2: **Combination Antiretroviral Medications**

note: the minimum cell count default is 1

- Characterization: **Characterization**

2. Execute Cohort Pathway 1, Cohort Pathway 2, and Characterization on CDM Transformed Database

Facilitate the analysis under the Execution tab and click 'Generate'

3. Return csv Results Files

Share the following .csv results files with the study coordinator (Raechel):

- pathways
- EventCohortCounts
- DistinctEventCohorts
- All Characterization Results files

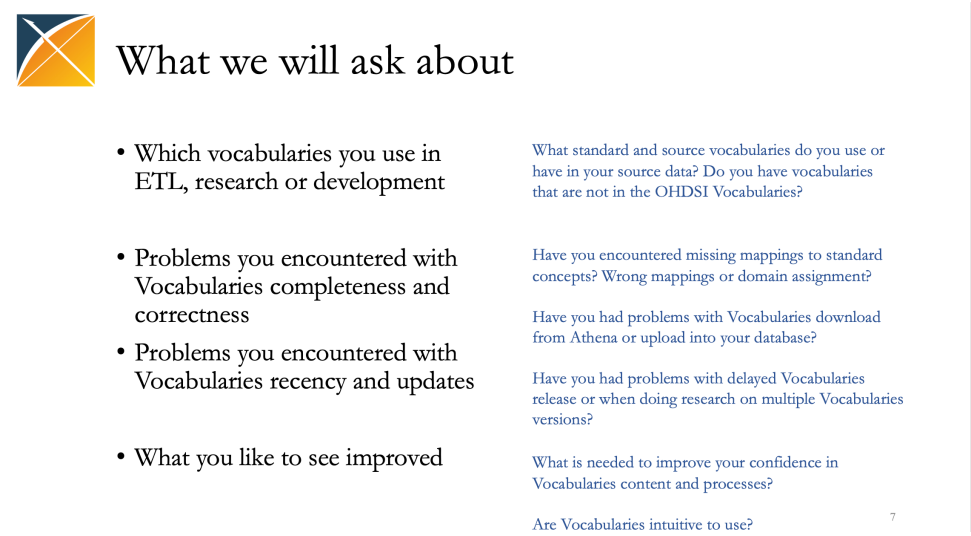
A secure file transfer protocol can be implemented to facilitate the sharing of the results files as requested by data partners- please reach out to raechel.davis@yale.edu

raechel.davis@yale.edu

Vocabulary Landscape Assessment

Anna Ostropolets introduced a vocabulary landscape assessment survey to directly inform which vocabularies and activities the vocabulary team prioritizes in 2023.

The deadline is Feb. 23.



bit.ly/3iTnyco

Save Our Sisyphus Challenge

Save Our Sisyphus Challenge

The task of taking a research study from idea through design through execution through publication can seem a daunting challenge, much like rolling a boulder up a hill. That task is all the more challenging when researchers try to go it alone, as each step requires a distinct set of skills. Observational study design requires epidemiologic understanding and statistical methodological expertise. Implementing a study design requires statistical programming ability. Interpreting and reporting results requires domain knowledge of the clinical problem.

But when you are part of the OHDSI community, you never have to go it alone. And as a team effort, what seems an arduous task can become an efficient and effective process.

We are seeking important research questions that you want to contribute and participate in to take from idea to publication. The OHDSI community will provide support through every step of the process, working with you to design an appropriate protocol, implement a network analysis package, execute across OHDSI data partners, and prepare a manuscript for publication. Our goal is to collaboratively complete this network study over the course of 8 weeks across April and May, using the open-source tools and process that OHDSI has

https://forms.gle/DySfETJPtmwgquKv9

Next CBER Best Seminar

The CBER BEST Seminar Series returns Wed., Feb. 8, at 11 am ET, as 2022 Titan Award recipient **Fan Bu** will provide a presentation on Bayesian Safety Surveillance with Adaptive Bias Correction.

Speaker: Dr. Fan Bu (UCLA)

Description: In this presentation, we will discuss a collaborative project with the FDA CBER BEST Initiative to improve on post-market vaccine safety surveillance procedures through Bayesian sequential analysis. Post-market surveillance on approved vaccine products is essential for addressing safety concerns. The goal is to detect rare or high-risk adverse events that often go undetected in clinical trials due to limited sample sizes. Collaborating with FDA CBER, we have developed a Bayesian alternative surveillance procedure that tackles these challenges in sequential analysis of observational data. The standard statistical approach for surveillance is Maximum Sequential Probability Ratio Test (MaxSPRT). Through comprehensive empirical evaluations on large-scale observational healthcare databases, we show that, compared to MaxSPRT, our Bayesian method offers more flexibility on the surveillance schedule, more transparency and interpretability in decision-making, and better error control through statistical correction of bias in observational data.

Latest Newsletter Is Available

OHDSI

The Journey Newsletter (Feb. 2023)

How can we strengthen OHDSI pillars like standardized vocabularies, data network & open-source tools? Learn more about these, as well as a pair of community activities (Phenotype Phebruary and the Save Our Sisyphus Challenge), and check out the recent publications and presentations from OHDSI in the latest edition of The Journey Newsletter! **#JoinTheJourney**

February Update Videocast

In the latest On The Journey, Patrick Ryan and Craig Sachson discuss how three OHDSI pillars can be strengthened through collaboration. They also discuss a pair of upcoming activities, Phenotype Phebruary and the Save Our Sisyphus Challenge.
(if video does not appear, click 'view this email in your browser')

Community Updates

Where Have We Been?

• Strengthening the OHDSI pillars of standardized vocabularies, a standardized data network and standardized open-source tools was one of the messages from **Patrick Ryan** during the first community call of 2023.
• **Anna Ostropolets** introduced a vocabulary landscape assessment survey to directly inform which vocabularies and activities the vocabulary team prioritizes in 2023. More information on this survey is posted below, and the deadline is Feb. 23.
• PatientLevelPrediction, a part of the HADES open-source tool library, is an R package for building and validating patient-level predictive models using data in the OMOP Common Data Model format. **Jenna Reps** recently created a series of demo videos to provide assistance with using v6 of the package, as well as the new Strategus tool.

Where Are We Now?

• Phenotype Phebruary begins today! **Gowtham Rao** and **Azza Shoaibi** are leading this community activity meant to both develop and evaluate phenotypes for health outcomes that could be investigated by the community. Check out more, including the introductory presentation, later in this newsletter.
• The Save Our Sisyphus Network Study event (learn more later in this newsletter) begins next month, but we are currently looking for an interesting research question to kick off this event. If you have one and would like to serve as P.I. for this global collaboration, please fill out this form.
• 2022 Titan Award recipient **Fan Bu** will discuss 'Bayesian Safety Surveillance with Adaptive Bias Correction' during the Feb. 8 edition of the CBER Best Seminar Series. This follows her OHDSI2022 talk on US FDA/CBER: Performance of vaccine safety surveillance methods. Registration is now open for this talk.

January Publications

Park K, Cho M, Song M, Yoo S, Baek H, Kim S, Kim K. Exploring the potential of OMOP common data model for process mining in healthcare. PLoS One. 2023 Jan 3;18(1):e0279641. doi: 10.1371/journal.pone.0279641. PMID: 36595527; PMCID: PMC9810199.

Yu Y, Jiang G, Brandt E, Forsyth T, Dhruva SS, Zhang S, Chen J, Noseworthy PA, Doshi AA, Collison-Farr K, Kim D, Ross JS, Coplan PM, Drozda JP Jr. Integrating real-world data to assess cardiac ablation device outcomes in a multicenter study using the OMOP common data model for regulatory decisions: implementation and evaluation. JAMIA Open. 2023 Jan 10;6(1):ooac108. doi: 10.1093/jamiaopen/ooac108. PMID: 36632328; PMCID: PMC9831049.

Schuemie MJ, Bu F, Nishimura A, Suchard MA. Adjusting for both sequential testing and systematic error in safety surveillance using observational data: Empirical calibration and MaxSPRT. Stat Med. 2023 Jan 15. doi: 10.1002/sim.9631. Epub ahead of print. PMID: 36642826.

Xie J, Brash JT, Turkmen C, Driessen S, Varrassi G, Argyriou G, Seager S, Reich C, Prieto-Alhambra D. Risk of COVID-19 Diagnosis and Hospitalisation in Patients with Osteoarthritis or Back Pain Treated with Ibuprofen Compared to Other NSAIDs or Paracetamol: A Network Cohort Study. Drugs. 2023 Jan 24:1–15. doi: 10.1007/s40265-022-01822-z. Epub ahead of print. PMID: 36692805; PMCID: PMC9872078.

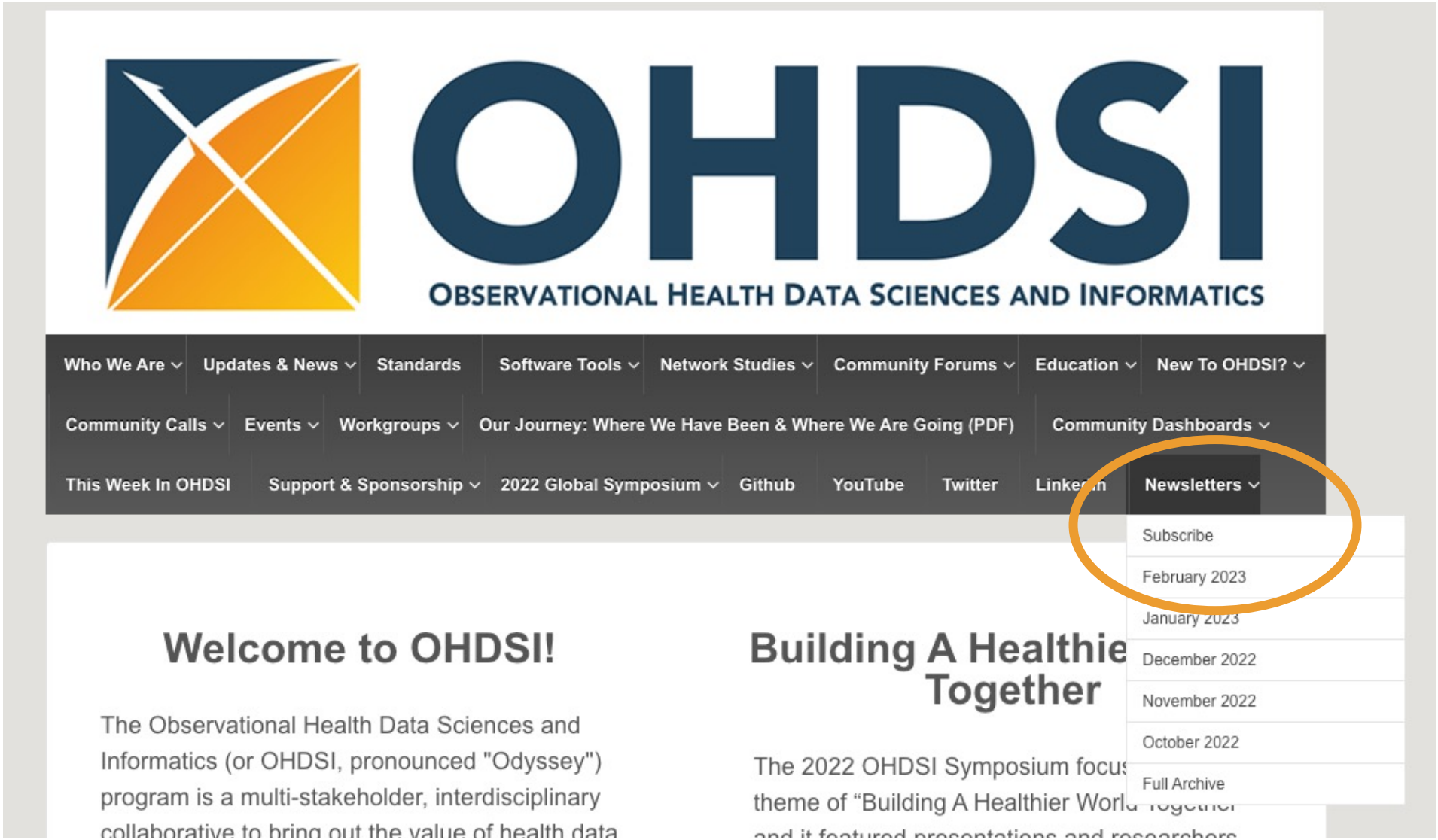
Reinecke I, Siebel J, Fuhrmann S, Fischer A, Sedlmayr M, Weidner J, Bathelt F. Assessment and Improvement of Drug Data Structuredness From Electronic Health Records: Algorithm Development and Validation. JMIR Med Inform. 2023 Jan 25;11:e40312. doi: 10.2196/40312. PMID: 36696159.

"Save Our Sisyphus" Network Study Challenge Seeks Research Topics By Feb. 27

The task of taking a research study from idea through design, execution and publication is a daunting challenge, and even more so when you attempt to drive it alone. Observational study design requires epidemiologic understanding and statistical methodological expertise. Implementing a study design requires statistical programming ability. Interpreting and reporting results requires domain knowledge of the clinical problem.

When you are part of the OHDSI community, you never have to go it alone. After all, we're all about Team Science!

Latest Newsletter Is Available



Save the Date! April 21: DevCon

OHDSI DevCon 2022 Welcomes & Mentors New Contributors To Our Open-Source Environment

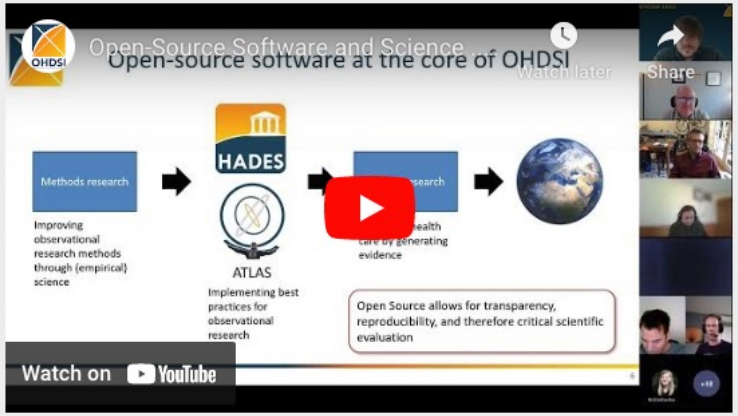
Watch All Eight Workshops, Talks & The Panel From DevCon Below

The Open-Source Community hosted the first Dev Con on Friday, April 22 as a way of accepting and mentoring new contributors to our environment. Organized by **Paul Nagy** and **Adam Black**, the event included eight workshops, talks and a panel discussion to both welcome and engage both current and future developers within OHDSI.

All videos from this session have or will be uploaded to this page. A big announcement from DevCon was the formation of the Khieron Contributor Cohort, which will help onboard and mentor open-source developers in the community. If you are interested in joining the effort, please fill out the application.

To learn more about the Khieron Contributor Cohort, please check out the State of the Open Source Community presentation below.

OHDSI DevCon Keynote

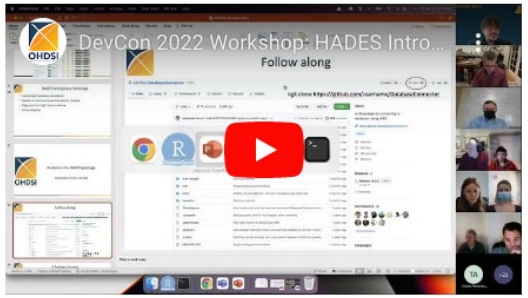


Martijn Schuemie provided the keynote address during DevCon 2022, entitled "Open-Source Software and Science … Obviously." His slides are available here.

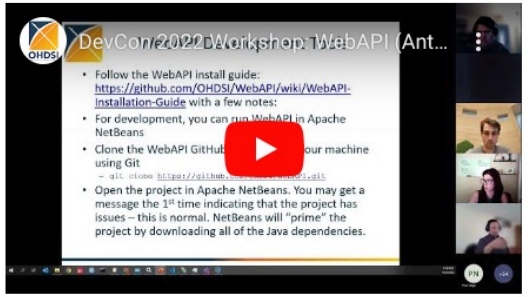
Workshops

ATLAS
(Anthony Sena)

HADES Introduction
(Adam Black)



WebAPI
(Anthony Sena)



Cohort Diagnostics
(James Gilbert)

White Rabbit
(Maxim Moinat)

Patient-Level Prediction
(Jenna Reps)

Teams invite will go out at a later date.

ICPE 2023 Abstract Deadline: Feb. 13

pharmacoepi.org/meetings/annual-conference/

2023 Health Data Science Black Internship Program

Dani Prieto-Alhambra announced that applications are open for the 2023 Health Data Science Black Internship Programme at the University of Oxford.

Closing date for registration is Feb. 27.

UNIVERSITY OF OXFORD

About the programme

The aim of our Black Internship Programme is to tackle the underrepresentation of Black people within the health data science sector. We are doing this by providing you with an opportunity to expand your knowledge around health data science and gain the experience you need to kickstart (or advance) your career in this field.

Our internships are a super way to gain hands-on experience, carrying out practical projects in the real world. It's a great way to find out about a rapidly advancing area of science, meet fellow interns, looks superb on a CV – and can open doors to new opportunities.

Planned for the summer of 2023, this internship programme will offer:

- 8-week paid internship
- Opportunity to join the wider 10,000 Black Interns community
- Certificate recognising intern achievements
- Opportunities across sectors in health data
- Customised learning pathway within HDR UK Futures
- Ongoing support post-programme from HDR UK Alumni Network and access to HDR UK Futures
- A real world data project developed by our host organisations
- Mentor and line manager
- Opportunity to shape next year's programme
- Cohort-building and training activities every Friday afternoon
- Team technical challenge and prize giving

Job Opening

COLUMBIA | COLUMBIA UNIVERSITY DEPARTMENT OF BIOMEDICAL INFORMATICS

DBMI Home | News & Events | Research | People | Prospective Students | Academics | Resources

Tenure Track Faculty

#105752

Description

The Department of Biomedical Informatics (DBMI) of Columbia University seeks exceptional junior-level faculty members in the tenure track.

The positions are open to researchers interested in developing and applying informatics theory and achieving tangible benefits to health care and biology. Three particular foci are (1) machine learning for healthcare and health-related data science, (2) health information technology-based interventions to improve health care and the health of individuals and populations, and (3) translational bioinformatics.

Job Opening

Job Details

Database Programmer

Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, Windmill Road, Oxford, OX3 7LD

We are seeking to appoint a highly qualified and dedicated Database Programmer to join the Health Data Sciences research group led by Professor Daniel Prieto-Alhambra at the Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), Oxford.

You will join an outstanding, multi-disciplinary and friendly Group of motivated and cutting-edge researchers and to contribute to clinical research by providing technical knowledge, software engineering expertise and data insight.

As a Database Programmer you will Develop new database applications for big clinical data to meet project requirements and deadlines, provide software feedback and carry out software improvement, extension, integration and further development on existing code. You will contribute to the harmonisation, curation, and processing of large clinical datasets and develop code to validate, test, document and maintain database applications. You will also represent the project, team, and the University in collaboration meetings, conferences and at external meetings.

You will have a Degree in computer science, software engineering, health informatics or an equivalent combination of training and professional experience. Proven understanding and experience in one or more RDBMSs and SQL dialects (e.g. PostgreSQL), excellent skills in at least one high level programming language (e.g. Python, C#, C++) and excellent analytical and problem-solving skills with great attention to detail are essential. Experience in common data models (CDMs) and in the extract, transform, and load (ETL) process, knowledge of R and/or RStudio and working experience in a research environment are desirable.

This is a full-time fixed-term appointment for 2 years.

The closing date for this position is 12 noon on Monday 27 February 2023. You will be required to upload a CV and supporting statement as part of your online application.

Contact Person :	HR Team, NDRMS	**Vacancy ID :**	163066
Contact Phone :		**Closing Date & Time :**	27-Feb-2023 12:00

Summer Internships

General Administration

Epidemiology Graduate Intern

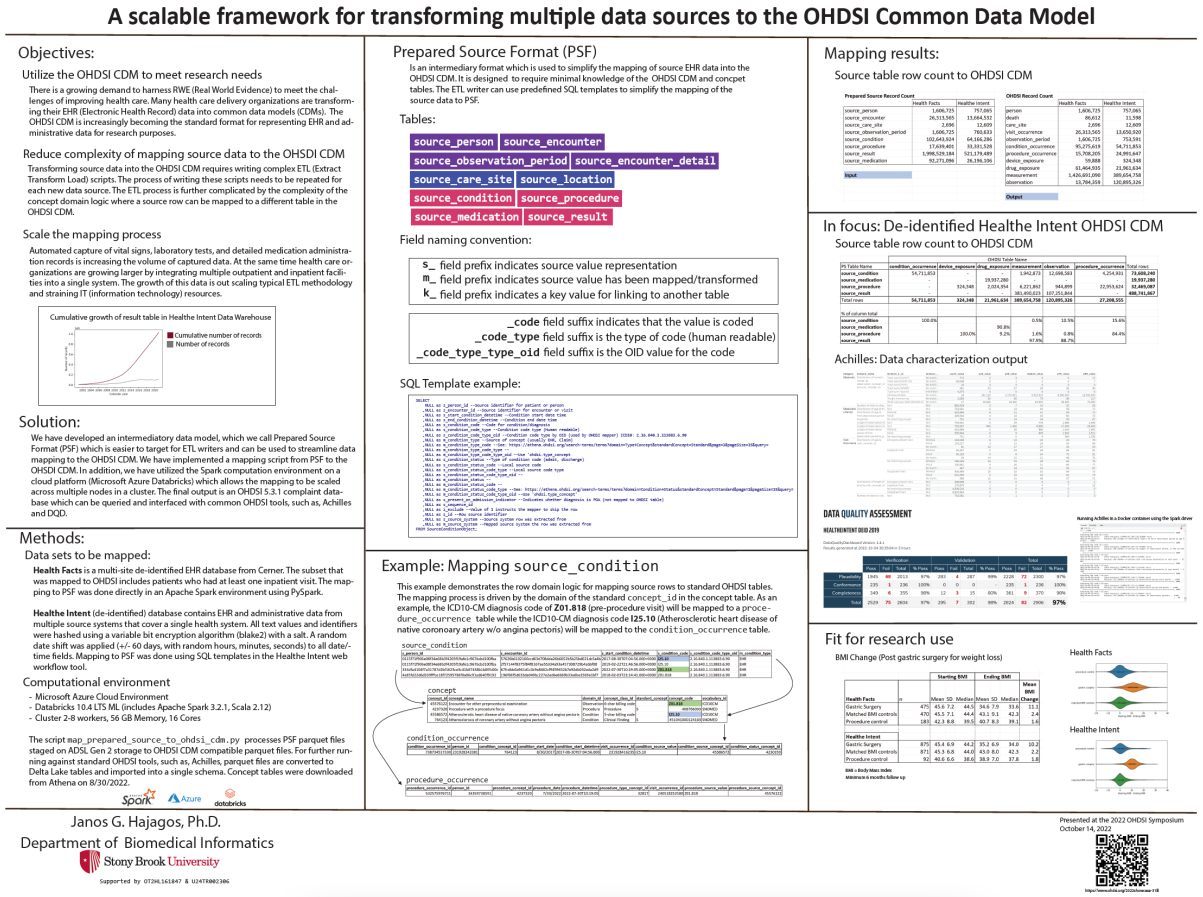
General Administration

OHDA Graduate Intern

General Administration

OHDA Undergraduate Intern

#OHDSISocialShowcase This Week



A scalable framework for transforming multiple data sources to the OHDSI Common Data Model

Objectives:

Utilize the OHDSI CDM to meet research needs

There is a growing demand to harness RWE (Real World Evidence) to meet the challenges of improving health care. Many health care delivery organizations are transforming their EHR (Electronic Health Record) data into common data models (CDMs). The OHDSI CDM is increasingly becoming the standard format for representing EHR and administrative data for research purposes.

Reduce complexity of mapping source data to the OHSDI CDM

Transforming source data into the OHDSI CDM requires writing complex ETL (Extract Transform Load) scripts. The process of writing these scripts needs to be repeated for each new data source. The ETL process is further complicated by the complexity of the concept domain logic where a source row can be mapped to a different table in the OHDSI CDM.

Scale the mapping process

Automated capture of vital signs, laboratory tests, and detailed medication administration records is increasing the volume of captured data. At the same time health care organizations are growing larger by integrating multiple outpatient and inpatient facilities into a single system. The growth of this data is out scaling typical ETL methodology and straining IT (information technology) resources.

Cumulative growth of result table in Healthe Intent Data Warehouse

Cumulative number of records
Number of records

Solution:

We have developed an intermediatory data model, which we call Prepared Source Format (PSF) which is easier to target for ETL writers and can be used to streamline data mapping to the OHDSI CDM. We have implemented a mapping script from PSF to the OHSDI CDM. In addition, we have utilized the Spark computation environment on a cloud platform (Microsoft Azure Databricks) which allows the mapping to be scaled across multiple nodes in a cluster. The final output is an OHDSI 5.3.1 complaint database which can be queried and interfaced with common OHDSI tools, such as, Achilles and DQD.

Methods:

Data sets to be mapped:

Health Facts is a multi-site de-identified EHR database from Cerner. The subset that was mapped to OHDSI includes patients who had at least one inpatient visit. The mapping to PSF was done directly in an Apache Spark environment using PySpark.

Healthe Intent (de-identified) database contains EHR and administrative data from multiple source systems that cover a single health system. All text values and identifiers were hashed using a variable bit encryption algorithm (blake2) with a salt. A random date shift was applied (+/- 60 days, with random hours, minutes, seconds) to all date/-time fields. Mapping to PSF was done using SQL templates in the Healthe Intent web workflow tool.

Computational environment

- Microsoft Azure Cloud Environment
- Databricks 10.4 LTS ML (includes Apache Spark 3.2.1, Scala 2.12)
- Cluster 2-8 workers, 56 GB Memory, 16 Cores

The script `map_prepared_source_to_ohdsi_cdm.py` processes PSF parquet files staged on ADSL Gen 2 storage to OHDSI CDM compatible parquet files. For further running against standard OHDSI tools, such as, Achilles, parquet files are converted to Delta Lake tables and imported into a single schema. Concept tables were downloaded from Athena on 8/30/2022.

Spark Azure databricks

Janos G. Hajagos, Ph.D.
Department of Biomedical Informatics
Stony Brook University

Supported by OT2HL161847 & U24TR002306

Prepared Source Format (PSF)

Is an intermediary format which is used to simplify the mapping of source EHR data into the OHDSI CDM. It is designed to require minimal knowledge of the OHDSI CDM and concpet tables. The ETL writer can use predefined SQL templates to simplify the mapping of the source data to PSF.

Tables:

- source_person
- source_encounter
- source_observation_period
- source_encounter_detail
- source_care_site
- source_location
- source_condition
- source_procedure
- source_medication
- source_result

Field naming convention:

`s_` field prefix indicates source value representation
`m_` field prefix indicates source value has been mapped/transformed
`k_` field prefix indicates a key value for linking to another table

`_code` field suffix indicates that the value is coded
`_code_type` field suffix is the type of code (human readable)
`_code_type_type_oid` field suffix is the OID value for the code

SQL Template example:

```
SELECT
    NULL as i_person_id --Source identifier for patient or person
    ,NULL as i_encounter_id --Source identifier for encounter or visit
    ,NULL as s_start_condition_datetime --Condition onset date time
    ,NULL as s_end_condition_datetime --Condition end date time
    ,NULL as s_condition_code --Code for condition/diagnosis
    ,NULL as s_condition_code_name --Condition code name (human readable)
    ,NULL as s_condition_code_type_oid --Condition code type by OID (used by OHDSI mapper) ICD10: 2.16.840.1.113883.6.90
    ,NULL as s_condition_type --Source of concept (usually EHR, Claim)
    ,NULL as s_condition_type_type --Type: http://athena.ohdsi.org/search-terms/terms?domain=Type+Concept&standardConcept=Standard&page=1&pageSize=15&query=
    ,NULL as s_condition_type_code_type
    ,NULL as s_condition_type_code_type_oid --Use 'OHDSI_type_concept'
    ,NULL as s_condition_status --Status of condition (admit, discharge)
    ,NULL as s_condition_status_code --Local source code
    ,NULL as s_condition_status_code_type --Local source code type
    ,NULL as s_condition_status_code_type_oid --
    ,NULL as s_condition_status_code
    ,NULL as s_condition_status_code_type --Doc: https://athena.ohdsi.org/search-terms/terms?domain=Condition+Status&standardConcept=Standard&page=1&pageSize=15&query=
    ,NULL as s_condition_status_code_type_oid --Use 'OHDSI_type_concept'
    ,NULL as s_condition_source_indicator --Indicates whether diagnosis is POA (not mapped to OHDSI table)
    ,NULL as i_sequence_id
    ,NULL as i_suitable --Value of 1 instructs the mapper to skip the row
    ,NULL as s_load_id --Load identifier
    ,NULL as i_source_system --Source system row was extracted from
    ,NULL as m_source_system --Mapped source system the row was extracted from
FROM SourceEditionObject;
```

Example: Mapping `source_condition`

This example demonstrates the row domain logic for mapping source rows to standard OHDSI tables. The mapping process is driven by the domain of the standard `concept_id` in the concept table. As an example, the ICD10-CM diagnosis code of **Z01.818** (pre-procedure visit) will be mapped to a `procedure_occurrence` table while the ICD10-CM diagnosis code **I25.10** (Atherosclerotic heart disease of native coronoary artery w/o angina pectoris) will be mapped to the `condition_occurrence` table.

`source_condition`

`concept`

`condition_occurrence`

`procedure_occurrence`

Mapping results:

Source table row count to OHDSI CDM

Prepared Source Record Count

	Health Facts	Healthe Intent
source_person	1,606,725	757,065
source_encounter	26,313,565	13,694,532
source_care_site	2,696	12,699
source_observation_period	1,606,725	760,633
source_condition	102,643,924	64,166,286
source_procedure	17,639,401	33,331,520
source_result	1,998,529,184	521,179,489
source_medication	92,271,096	26,196,106

Input

OHDSI Record Count

	Health Facts	Healthe Intent
person	1,606,725	757,065
death	86,612	11,598
care_site	2,696	12,699
visit_occurrence	26,313,565	13,650,920
observation_period	1,606,725	753,591
condition_occurrence	95,275,619	54,715,853
procedure_occurrence	15,708,205	34,991,647
device_exposure	59,888	324,348
drug_exposure	61,464,935	21,961,634
measurement	1,426,691,096	389,654,758
observation	13,794,159	120,895,326

Output

In focus: De-identified Healthe Intent OHDSI CDM

Source table row count to OHDSI CDM

	OHDSI Table Name					
Pt Table Name	condition_occurrence	device_exposure	drug_exposure	measurement	observation	procedure_occurrence	Total rows
source_condition	54,711,853	-	-	1,942,873	12,698,583	4,254,931	73,608,240
source_medications	-	-	19,937,280	-	-	-	19,937,280
source_procedure	-	324,348	2,024,954	6,221,802	944,899	22,953,624	32,469,087
source_result	-	-	-	381,450,023	107,251,864	-	488,701,887
Total rows	54,711,853	324,348	21,961,634	389,654,758	120,895,326	27,208,555	
% of columns total							
source_condition	100.0%			0.5%	10.5%		15.6%
source_medication			90.8%				
source_procedure		100.0%	9.2%	1.6%	0.8%		64.4%
source_result				97.9%	88.7%		

Achilles: Data characterization output

DATA QUALITY ASSESSMENT

HEALTHEINTENT DEID 2019

	Verification				Validation				Total			
	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass	Pass	Fail	Total	% Pass
Plausibility	1345	98	2013	87%	283	4	287	99%	2228	72	2300	97%
Conformance	235	1	236	100%	0	0	0	-	235	1	236	100%
Completeness	343	6	355	98%	12	3	15	80%	351	9	370	99%
Total	2529	75	2604	97%	295	7	302	98%	2814	82	2896	97%

Fit for research use

BMI Change (Post gastric surgery for weight loss)

		Starting BMI		Ending BMI		
Health Facts	n	Mean	SD	Median	Mean	SD	Median	Mean BMI Change
Gastric Surgery	475	45.6	7.2	44.5	34.6	7.9	33.6	11.1
Matched BMI controls	470	45.5	7.1	44.4	43.1	9.1	42.3	2.4
Procedure control	183	42.3	8.8	39.5	40.7	8.3	39.1	1.6
Healthe Intent								
Gastric Surgery	875	45.4	6.9	44.2	35.2	6.9	34.0	10.2
Matched BMI controls	871	45.3	6.8	44.0	45.0	8.0	42.3	2.2
Procedure control	92	40.6	6.6	38.6	38.9	7.0	37.8	1.8

BMI = Body Mass Index
Minimum 6 months follow up

Health Facts

Healthe Intent

Presented at the 2022 OHDSI Symposium
October 14, 2022

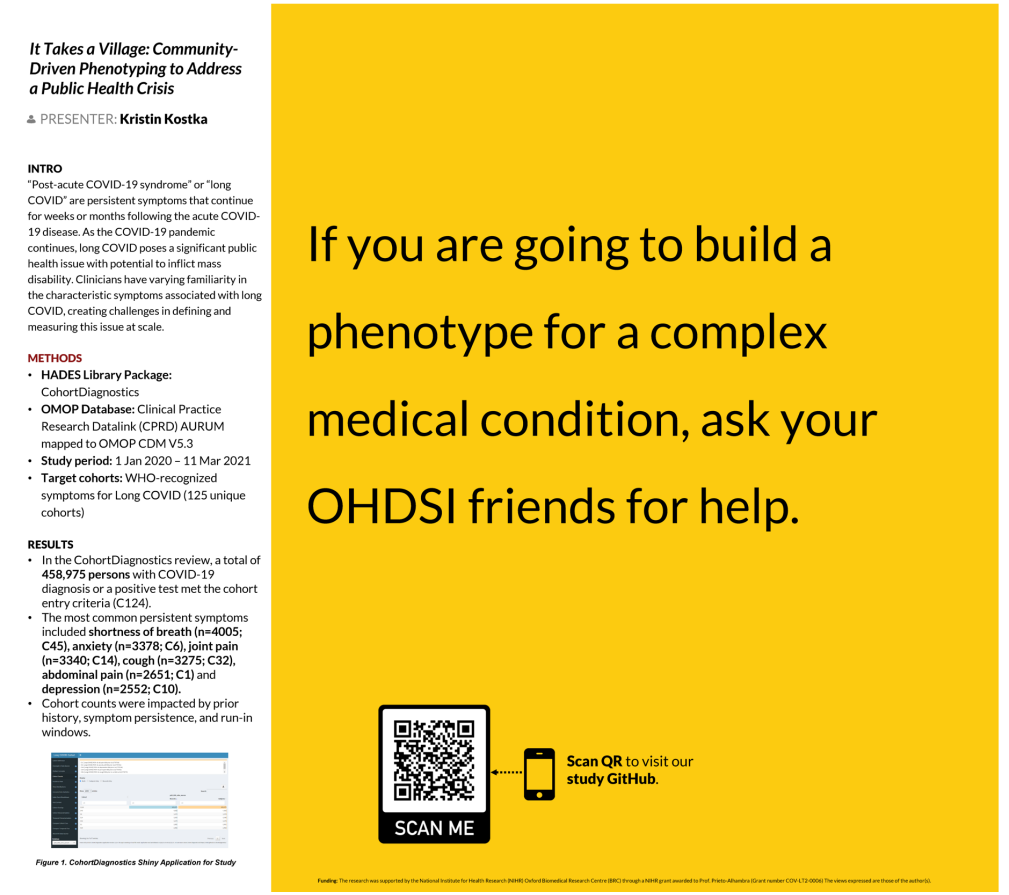
MONDAY A scalable framework for transforming multiple data sources to the OHDSI Common Data Model (Janos Hajagos)

#OHDSISocialShowcase This Week

TUESDAY

PDA-OTA: Privacy-preserving Distributed Algorithms Over the Air, an OHDSI journey (Yong Chen, Jiayi Tong, Chongliang Luo, Lu Li, Yiwen Lu, Hai-Shuo Shu)

#OHDSISocialShowcase This Week



It Takes a Village: Community-Driven Phenotyping to Address a Public Health Crisis

PRESENTER: **Kristin Kostka**

INTRO

"Post-acute COVID-19 syndrome" or "long COVID" are persistent symptoms that continue for weeks or months following the acute COVID-19 disease. As the COVID-19 pandemic continues, long COVID poses a significant public health issue with potential to inflict mass disability. Clinicians have varying familiarity in the characteristic symptoms associated with long COVID, creating challenges in defining and measuring this issue at scale.

METHODS

- **HADES Library Package:** CohortDiagnostics
- **OMOP Database:** Clinical Practice Research Datalink (CPRD) AURUM mapped to OMOP CDM V5.3
- **Study period:** 1 Jan 2020 – 11 Mar 2021
- **Target cohorts:** WHO-recognized symptoms for Long COVID (125 unique cohorts)

RESULTS

- In the CohortDiagnostics review, a total of **458,975 persons** with COVID-19 diagnosis or a positive test met the cohort entry criteria (C124).
- The most common persistent symptoms included **shortness of breath (n=4005; C45), anxiety (n=3378; C6), joint pain (n=3340; C14), cough (n=3275; C32), abdominal pain (n=2651; C1)** and **depression (n=2552; C10).**
- Cohort counts were impacted by prior history, symptom persistence, and run-in windows.

Figure 1. CohortDiagnostics Shiny Application for Study

If you are going to build a phenotype for a complex medical condition, ask your OHDSI friends for help.

SCAN ME

Scan QR to visit our **study GitHub.**

Funding: The research was supported by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC) through a NIHR grant awarded to Prof. Prieto-Alhambra (Grant number COV-LT2-0006) The views expressed are those of the author(s).

WORK POWERED BY THE OHDSI COMMUNITY

We hosted a Long COVID phenotyping hackathon on December 7, 2021.

Figure 2. Actual footage of community in action

In the hackathon, we used the World Health Organization (WHO) Delphi consensus of the clinical case definition of post COVID-19 condition. We assembled concept sets for the 25 individual symptoms using OHDSI best practice (Figure 2).

Together, the community produced:

- **7 final clinical symptom** concept set expressions meeting the OHDSI best practices
- **9 drafted clinical symptom** concept set expression for further review with OHDSI diagnostics
- **9 clinical symptom** concept set expressions to be developed.
- **(… and Gowtham broke PHOEBE. ☺)**

Iterative results are stored in the OHDSI Phenotype WG Long COVID channel on OHDSI MS Teams.

Kristin Kostka[1,2], Evan Minty[3], Antonella Delmestri[1], Barrack Omondi[1], Martí Català[1], Elena Roel[5], Edward Burn[1,5], Daniel Prieto-Alhambra[1,4], Annika M. Jödicke[1]

Affiliations: [1]Pharmaco- and Device Epidemiology, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences, University of Oxford, OX3 7LD, UK; [2]The OHDSI Center at the Roux Institute, Northeastern University, Portland, ME, US; [3]O'Brien Centre for Population Health, Faculty of Medicine, University of Calgary, CA; [4]Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, NL; [5]Fundació Institut Universitari per a la recerca a l'Atenció Primària, Barcelona, ES

NDORMS OHDSI

WEDNESDAY

It Takes a Village: Community-Driven Phenotyping to Address a Public Health Crisis (Kristin Kostka, Evan Minty, Antonella Delmestri, Barrack Omondi, MartiCatala, Elena Roel, Edward Burn, Daniel Prieto-Alhambra, Annika M. Jadicke)

#OHDSISocialShowcase This Week

Reduce, Reuse, & Recycle: Going "Green" with Atlas Reusables

PRESENTER: Ajit Londhe

INTRO:

- Atlas offers a variety of criteria logic and attributes that are the building blocks of cohort analyses.
- CIRCE is the library that provides the cohort logic framework, from entry events to inclusion criteria to subgroups and censoring events.
- When tasked with defining logic – simple or complex – multiple times, the user experience in Atlas can be challenging, particularly with multiple nested criteria and attribute selections.
- We developed a new Atlas module named "Reusables," which is now available in Atlas versions 2.11 and above.

METHODS:

- We observed that all CIRCE criteria and attribute types would need to be available to import cleanly into Atlas study assets.
- Besides these standard CIRCE features, we observed that the design of a Reusable could benefit from being able to import other Reusables.
- We examined the context with which CIRCE logic is applied across analytic modules in Atlas, aiming to align Reusables with standard study design patterns.
- In cohort definitions, cohort entry and censoring are possible through the union of pathways rather than the intersection of those pathways.
- But inclusion criteria can involve the union or intersection of criteria
- We architected the Reusable import process to pass the CIRCE logic by value rather than by reference to ensure backwards compatibility with previous Atlas versions.

Reusable Atlas Logic Prevents Repeated Mistakes Resulting in Reputable Results

Build

Import

Reuse

Poster QR Code

https://github.com/OHDSI/Atlas
https://github.com/OHDSI/WebAPI

RESULTS:

- The Reusables designer interface allows for parameterization of concept sets and all the CIRCE-powered criteria types found throughout Atlas.
- As inclusion criteria can involve the union or intersection of criteria., the Reusables designer presents the user with explicit choice on whether the Reusable will be used as (1) an initial or censoring event, or (2) a criteria group-based event.
- The import process is simple, with "From Reusable" as a new event type button available in Cohort Entry Events, Inclusion Criteria, Censoring Events, Characterization custom features or subgroups, and Incidence Rate strata
- Importing the Reusable allows the user to fulfill any parameters defined in the Reusable's design before completing the import into the Atlas study asset.
- A Reusable can consist of 1 or more Reusables, leading to a new ability for users to craft multi-part algorithms with more modularity.

CONCLUSIONS:

- Atlas has been an important platform for OHDSI studies over the years, allowing for robust cohort design in a standardized method without programming.
- Reusables provide a powerful new method in Atlas for ensuring streamlined design and reproducibility of algorithms both within and between OHDSI sites.

Ajit Londhe[1,4], Brad Rechkemmer[2,4], Gregory Klebanov[3,4], Alexey Manoylenko[3,4], Anton Abushkevich[3,4], Sergey Suvorov[3,4]

1) At development time: Amgen, Inc. Thousand Oaks, CA, USA; Currently: Boehringer Ingelheim, Ridgefield, CT, USA
2) SimulStat, Inc. Solana Beach, CA, USA
3) Odysseus Data Services, Inc., Cambridge, MA, USA
4) Observational Health Data Sciences and Informatics (OHDSI), New York, NY, USA

OHDSI

THURSDAY Reduce, Reuse, & Recycle: Going Green with Atlas Reusables (Ajit Londhe (presenter), Brad Rechkemmer, Gregory Klebanov, Alexey Manoylenko, Anton Abushkevich, Sergey Suvorov)

#OHDSISocialShowcase This Week

Comparison of Biopsy and Diagnosis Code-Based Breast Cancer Phenotypes

Matthew Spotnitz, Thomas Falconer, Maura Beaton, Karthik Natarajan

Department of Biomedical Informatics, Columbia University Irving Medical Center

Background

Gold standard phenotyping of cancer can improve the rigor and reproducibility of research about the disease. There are multiple candidate phenotypes for the same kind of cancer, which can have different strengths. Specifically, some phenotypes may have higher sensitivities and others may have higher antecedent data frequencies. We used ATLAS to characterize the antecedent data frequencies for biopsy and diagnosis code-based breast cancer phenotypes.

Methods

The criteria for each phenotype are shown in Table 1. We compared the one-year antecedent data frequencies of both phenotypes on multiple databases, using ATLAS characterization plots.

Phenotype	Biopsy-Based	Diagnosis Code-Based
Index Event	Biopsy Procedure Code	Breast Cancer Diagnosis Code
Additional Criteria	•Breast cancer diagnosis within 90 days •No history of breast cancer at least 30 days prior •Women •18 to 80 years old	•No history of breast cancer at least 30 days prior •Women •18 to 80 years old
Observation Period	•90+ days prior	•90+ days prior
Chronological Period	•01/01/2000 or later	•01/01/2000 or later

Table 1: Biopsy and diagnosis code-based breast cancer phenotype criteria.

Data Sources

Our data sources were the Columbia University Irving Medical Center (CUIMC) Electronic Health Record, the IBM MarketScan Commercial Claims and Encounters (CCAE), and the IBM MarketScan Medicare Supplemental Beneficiaries (MDCR) Databases.

Results: Antecedent Data Frequencies

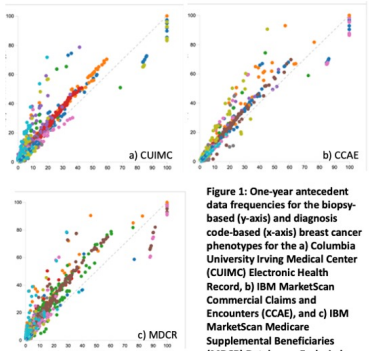


Figure 1: One-year antecedent data frequencies for the biopsy-based (y-axis) and diagnosis code-based (x-axis) breast cancer phenotypes for the a) Columbia University Irving Medical Center (CUIMC) Electronic Health Record, b) IBM MarketScan Commercial Claims and Encounters (CCAE), and c) IBM MarketScan Medicare Supplemental Beneficiaries (MDCR) Databases. Each circle corresponds with a covariate and each color with a data domain.

Results: Baseline Characteristics

Database	Biopsy-Based	Diagnosis Code-Based
CUIMC	5792	18759
CCAE	198575	408440
MDCR	23986	145231

Table 2: Phenotype counts by database.

Database	Biopsy-Based	Diagnosis Code-Based	Std. Diff.
CUIMC	58 [49-68]	59 [49-69]	0.01
CCAE	54 [47-60]	55 [48-60]	0.02
MDCR	71 [68-75]	71 [67-75]	-0.04

Table 3: Median age (with interquartile range) by database and phenotype. Standardized difference: Std. Diff.

Database	Biopsy-Based	Diagnosis Code-Based	Std. Diff.
CUIMC	3672 [1439-6330]	3203 [1271-5570]	-0.10
CCAE	721 [334-1405]	415 [180-1050]	-0.21
MDCR	918 [414-1692]	337[182-809]	-0.52

Table 4: Prior observation time (with interquartile range) by database and phenotype. Standardized difference: Std. Diff.

Conclusions

On multiple databases, the biopsy-based breast cancer phenotype had higher one-year antecedent data frequencies and longer prior observation times, suggesting more complete patient records. When choosing a cohort definition, reviewing such plots may help the investigators make data-driven decisions.

Contact: mes2165@cumc.columbia.edu

FRIDAY

Comparison of Biopsy and Diagnosis Code Based Breast Cancer Phenotypes

(Matthew Spotnitz, Thomas Falconer, Maura Beaton, Karthik Natarajan)

Where Are We Going?

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

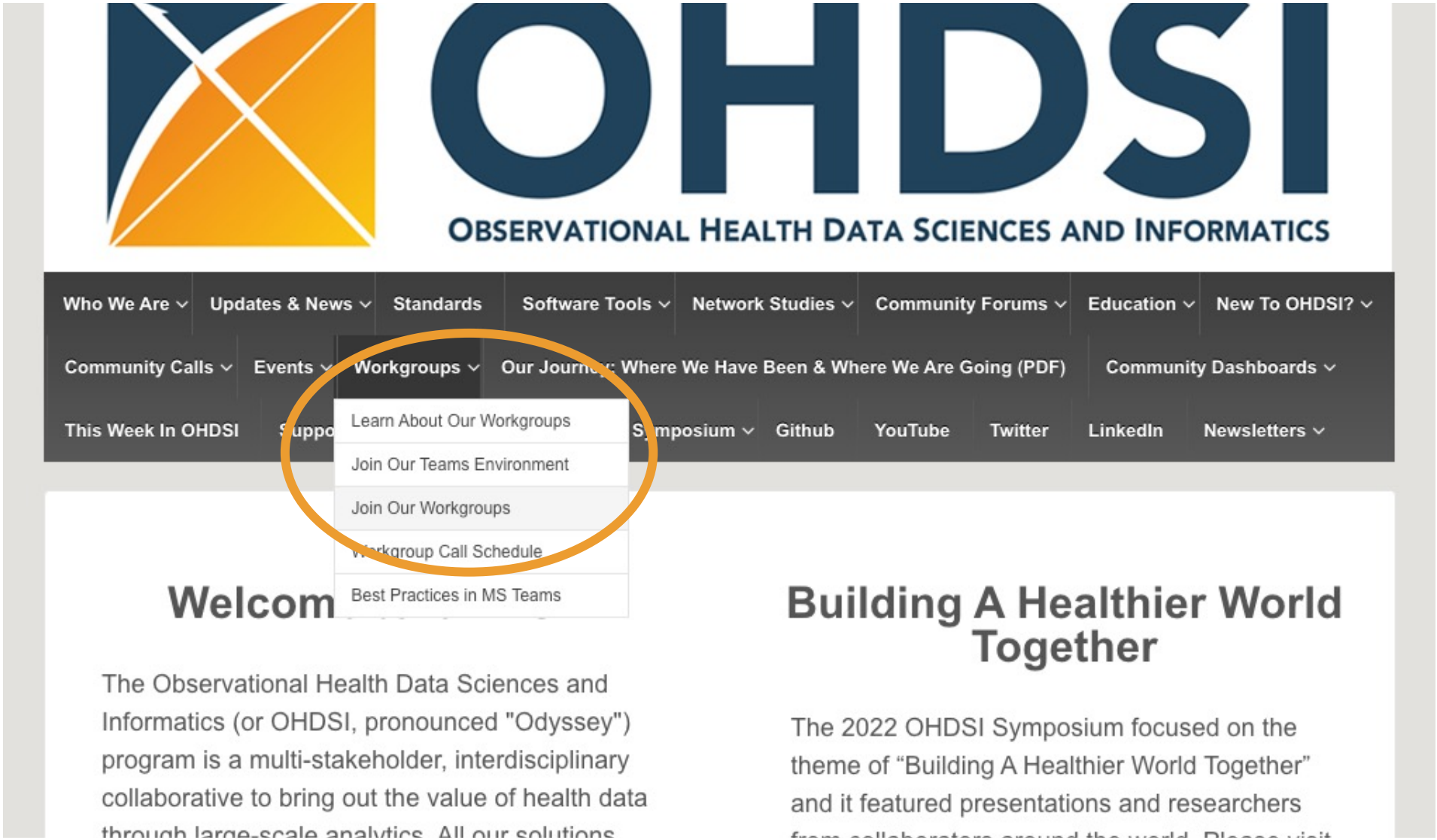
Three Stages of The Journey

Where Have We Been?

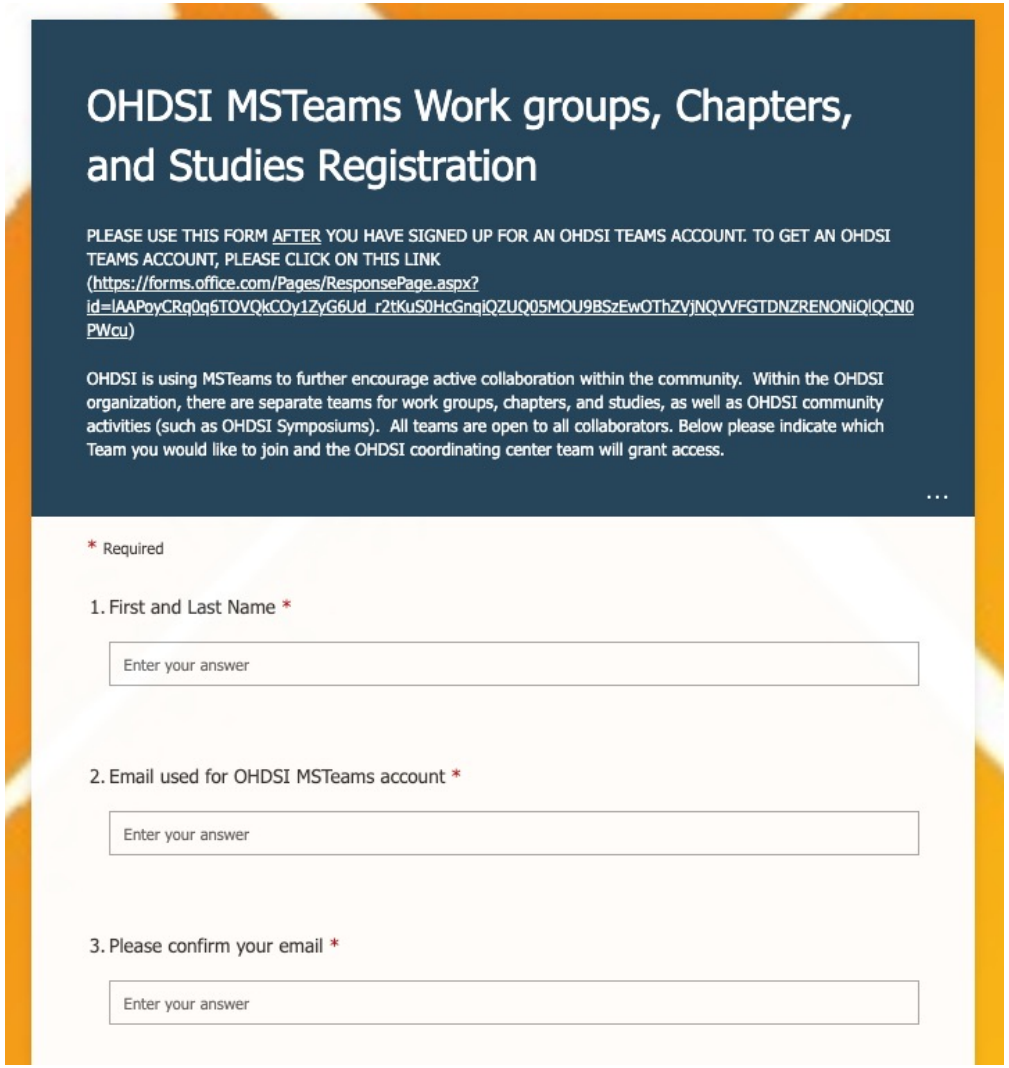
Where Are We Now?

Where Are We Going?

How To Join The Workgroups



How To Join The Workgroups



OHDSI MSTeams Work groups, Chapters, and Studies Registration

PLEASE USE THIS FORM AFTER YOU HAVE SIGNED UP FOR AN OHDSI TEAMS ACCOUNT. TO GET AN OHDSI TEAMS ACCOUNT, PLEASE CLICK ON THIS LINK (https://forms.office.com/Pages/ResponsePage.aspx?id=lAAPoyCRq0q6TOVQkCOy1ZyG6Ud_r2tKuS0HcGnqiQZUQ05MOU9BSzEwOThZVjNQVVFGTDNZRENONiQlQCN0PWcu)

OHDSI is using MSTeams to further encourage active collaboration within the community. Within the OHDSI organization, there are separate teams for work groups, chapters, and studies, as well as OHDSI community activities (such as OHDSI Symposiums). All teams are open to all collaborators. Below please indicate which Team you would like to join and the OHDSI coordinating center team will grant access.

...

* Required

1. First and Last Name *

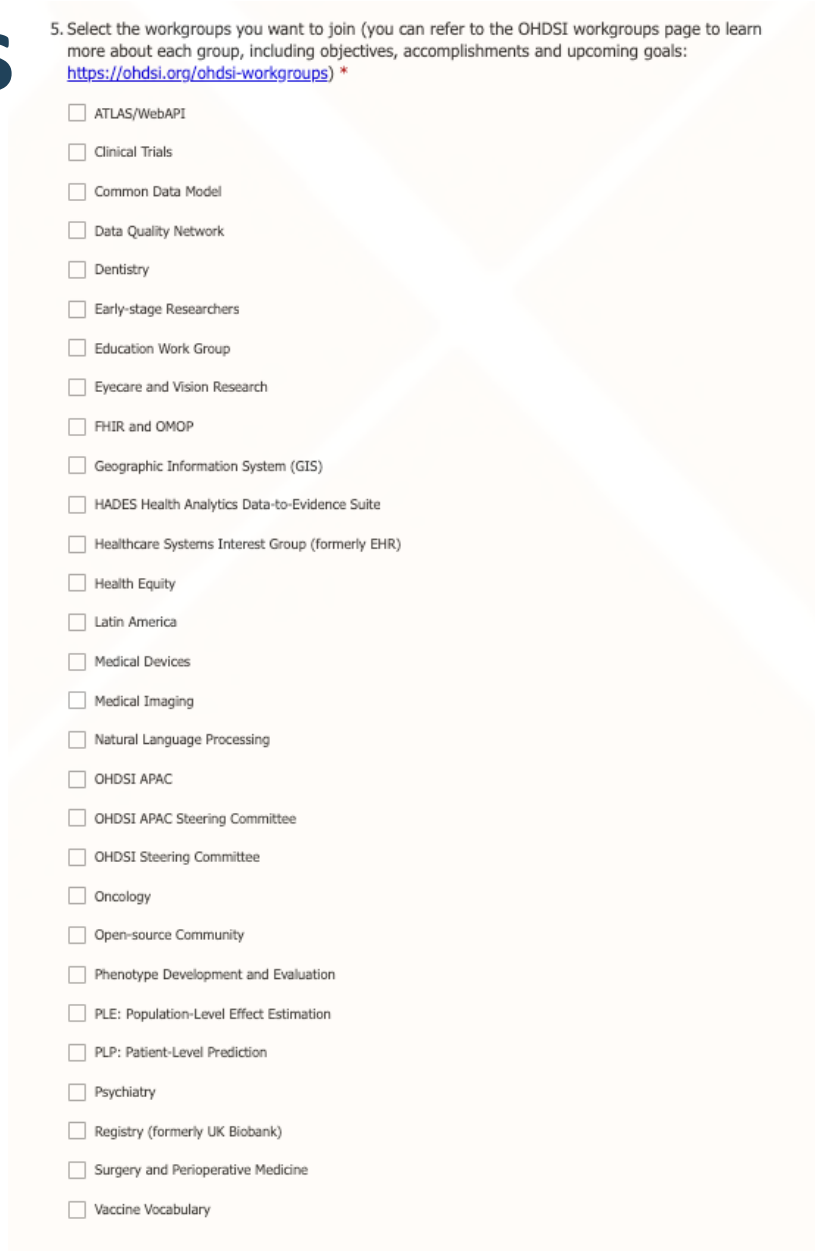
Enter your answer

2. Email used for OHDSI MSTeams account *

Enter your answer

3. Please confirm your email *

Enter your answer



5. Select the workgroups you want to join (you can refer to the OHDSI workgroups page to learn more about each group, including objectives, accomplishments and upcoming goals: https://ohdsi.org/ohdsi-workgroups) *

- [ ] ATLAS/WebAPI
- [ ] Clinical Trials
- [ ] Common Data Model
- [ ] Data Quality Network
- [ ] Dentistry
- [ ] Early-stage Researchers
- [ ] Education Work Group
- [ ] Eyecare and Vision Research
- [ ] FHIR and OMOP
- [ ] Geographic Information System (GIS)
- [ ] HADES Health Analytics Data-to-Evidence Suite
- [ ] Healthcare Systems Interest Group (formerly EHR)
- [ ] Health Equity
- [ ] Latin America
- [ ] Medical Devices
- [ ] Medical Imaging
- [ ] Natural Language Processing
- [ ] OHDSI APAC
- [ ] OHDSI APAC Steering Committee
- [ ] OHDSI Steering Committee
- [ ] Oncology
- [ ] Open-source Community
- [ ] Phenotype Development and Evaluation
- [ ] PLE: Population-Level Effect Estimation
- [ ] PLP: Patient-Level Prediction
- [ ] Psychiatry
- [ ] Registry (formerly UK Biobank)
- [ ] Surgery and Perioperative Medicine
- [ ] Vaccine Vocabulary

Common Data Model Workgroup

Clair Blacketer

Common Data Model WG Purpose

The CDM workgroup exists to curate and maintain the OMOP Common Data Model which in turn enables the community to develop standardized methods, software, and generate evidence that promotes better health decisions and better care.

CDM WG 2023 Objectives and Key Results

- Harmonize the vocabulary mapping work and model expansion efforts currently being conducted throughout the community by various working/interest groups.
- Conduct at least 10 meetings with a presentation from a working group or model expansion owner
- Design a maturity model and repository for OMOP CDM expansions

Data Network Quality Workgroup

Clair Blacketer

Data Network Quality WG Purpose

The Data Network Quality WG exists to develop the OHDSI Standardized Data Network and to recommend, enable, and develop best practices related to data quality. The higher quality data in the network, the higher quality evidence we will produce in support of OHDSI's mission.

Data Quality Network WG
2023 Objectives and Key Results

a. Create an OHDSI data network catalog to encourage network studies across interested partners and promote data quality practices
   - Get at least 20 individual databases to contribute data to the network.
   - Produce data diagnostic results across the contributors to the data network for at least 5 potential studies.

b. Continued support, maintenance, and development of the DataQualityDashboard package in relation to the Data Network initiative.
   - Add at least 3 new data quality checks.
   - Determine the subset of existing data quality checks that should be included in the DbProfile extract that network contributors will be expected to run.

Surgery & Perioperative Medicine WG

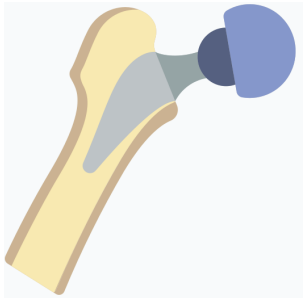
To collaboratively generate the observational health evidence needed to better understand the role of surgical and perioperative interventions in medical care

Evan Minty

Jenny Lane

Why Surgery and the Peri-Operative period is cool

- THE data science problem of our time
- It happened ☺
- There is a risk window
- The clinical problem has exciting devices & procedures



Aims

- Refining Surgical and PeriOp cohorts for the library
- Improving procedural vocabularies
- Project evaluation and prioritization
- Characterisations, PLEs & PLPs in Surgery & PeriOp research

Collaborations with other WGs

OKRs for 2023

1. Recruitment! Establish MDT membership, esp those new to OHDSI
2. Major Non-Cardiac Surgery phenotype + Characterisation
3. Collab with PLP WG for Major Non-Cardiac Surgery
4. Fragility Fracture Studyathon

How do I join?

Meeting every 4 weeks based upon time zones of members

MS teams environment for shared items

First meeting: 22nd February 3pm GMT / 10am EST / 8am MST

Minty@ohdsi.org or Lane@ohdsi.org

HADES

HEALTH ANALYTICS DATA-TO-EVIDENCE SUITE

HADES objective

Enable the OHDSI community to **perform observational research** following **OHDSI best practices** for characterization, population-level estimation, and patient-level prediction by providing a **cohesive set of open-source analytic software**.

Key results

- **Regular HADES-wide releases**
  - HADES-wide snapshots every 6 months
- **More user involvement**
  - Get feedback from at least 20 users
  - Have at least 5 champions / evangelist
  - Have at least 3 presentations by HADES users in HADES meetings
- **Have roadmaps, design specifications**
  - Establish a HADES-wide roadmap per HADES release
  - Have design specs for all new HADES packages / major changes
  - Align on input and output conventions
- **Stability**
  - Have test servers for all supported DB platforms
  - At least 50% of HADES packages use 50% of testing servers
  - Have and enforce a HADES dependency white list

(Population-Level Estimation) Methods Research Workgroup

Methods Workgroup objective

Perform the **methods research** to support OHDSI's analytics use cases

Goals

- Create awareness of who is researching what methods in OHDSI (anyone using OMOP)
- A sounding board for methods research in progress
- Identify new methods research questions
- Find collaborators

Please come share your ongoing / future methods research at a next WG meeting! (e.g. March 2)


OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS
Open Source Community
Adam Black & Paul Nagy

Workgroup purpose

The OHDSI Open Source Community exists to promote the health and sustainability of the OHDSI open source software ecosystem.

Pillars

EDUCATIONAL MATERIALS AND COMMUNITY EVENTS **FOR DEVELOPERS**

MATURE GOVERNANCE USING BEST PRACTICES FROM OTHER OPEN SOURCE COMMUNITES

DOCUMENTATION OF STANDARD INTERFACES AND PROCESSES

MONITOR METRICS OF OPEN-SOURCE COMMUNITY HEALTH AND SUSTAINABILITY

SUPPORT MAINTIANERS WITH INFRASTRUCTURE NEEDS

T1 2023 Key Results

Objective: Provide educational and community events for OHDSI software developers

Key result: Host OHDSI devcon 2023 in April

Objective: Train casual contributors to become active contributors/maintainers (Kheiron Program)

Key results: Graduate 2022 Kheiron cohort program and start 2023 Kheiron cohort

Objective: Comprehensive testing of OHDSI software all supported database platforms

Key Result: Document OHDSI test servers and how to request access

Objective: Formation of a Technical Advisory Board (TAB)

This a group who coordinates what software and what version of software is part of a release (distribution). Priorities are security, stability, and a consistent release for network.

Key result: Define and launch this group Q1 2023 (April)

OHDSI Steering Workgroup

Patrick Ryan, George Hripcsak

On behalf of the Steering WG

Steering WG Purpose

Steering WG exists to support the community and its leaders in collaboratively generating the evidence that promotes better health decisions and better care, by identifying, organizing, and guiding collaborative activities, facilitating communications across the community, providing input to operations of the OHDSI Central Coordinating Center, and building consensus on the vision for where the OHDSI community should go together.

Steering WG 2023 Objectives and Key Results

Objective 1: Empower workgroups to contribute to collaboratively generating the evidence that promotes better health decisions and better care

Key results:

1. 100% of active workgroups have defined purpose and 2023 OKRs that are communicated to broader community to promote focus and encourage contributions; Timeline: 1Q2023
2. 1 Workgroup Leader Summit convened to ensure appropriate communication across workgroups; Timeline: 1Q2023

Objective 2: Create collaboration activities that encourage collaborative generation and dissemination of the evidence that promotes better health decisions and better care

Key results:

1. OHDSI2023 Global Symposium scheduled with location/dates announced; Timeline: 1Q2023
2. 2 community activities with >30 collaborators participating: 1- Phenotype Phebruary , timeline: Feb2023; 2- Sisyphus Challenge, timeline: Mar-April2023
3. Community resource published and updated once per month with updated listed of conferences, grants, RFI, abstract/paper submission, and call for papers deadlines

Objective 3: Draft a strategy for a general purpose OHDSI asset library

Key results:

1. Document that lays out strategy with pluses and minuses to approach and resource requirements to build it; Timeline: 1Q2023