Week 3
Workgroup 2023 OKRs and Phenotype Phebruary Updates

OHDSI Community Call
Feb. 21, 2023 • 11 am ET

Upcoming OHDSI Community Calls

Date	Topic
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 the team of **Joel Swerdel, Darmendra Ramcharran and Jill Hardin** on the publication of **Using a data-driven approach for the development and evaluation of phenotype algorithms for systemic lupus erythematosus** in PLoS One.

PLOS ONE

OPEN ACCESS

Citation: Swerdel JN, Ramcharran D, Hardin J (2023) Using a data-driven approach for the development and evaluation of phenotype algorithms for systemic lupus erythematosus. PLoS ONE 18(2): e0281929. https://doi.org/10.1371/journal.pone.0281929

Editor: Luca Navarini, Universita Campus Bio-Medico di Roma, ITALY

Received: April 8, 2022

Accepted: February 4, 2023

Published: February 16, 2023

Peer Review History: PLOS recognizes the benefits of transparency in the peer review

RESEARCH ARTICLE

Using a data-driven approach for the development and evaluation of phenotype algorithms for systemic lupus erythematosus

Joel N. Swerdel[1,2]*, **Darmendra Ramcharran**[1¤], **Jill Hardin**[1,2]

1 Janssen Research and Development Epidemiology, Titusville, New Jersey, United States of America, **2** Observational Health Data Sciences and Informatics (OHDSI), New York, New York, United States of America

¤ Current address: Safety and Quantitative Innovation, GSK, Waltham, Massachusetts, United States of America
* jswerdel@its.jnj.com

Abstract

Background

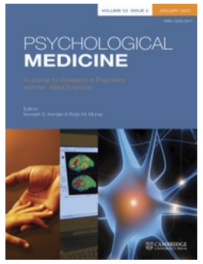
Systemic lupus erythematosus (SLE) is a chronic autoimmune disease of unknown origin. The objective of this research was to develop phenotype algorithms for SLE suitable for use in epidemiological studies using empirical evidence from observational databases.

Methods

We used a process for empirically determining and evaluating phenotype algorithms for health conditions to be analyzed in observational research. The process started with a literature search to discover prior algorithms used for SLE. We then used a set of Observational Health Data Sciences and Informatics (OHDSI) open-source tools to refine and validate the algorithms. These included tools to discover codes for SLE that may have been missed in prior studies and to determine possible low specificity and index date misclassification in algorithms for correction.

OHDSI Shoutouts!

Congratulations to the team of **Chungsoo Kim, Dong Yun Lee, Jimyung Park, Su-Jin Yang, Eng Hooi Tan, Daniel-Prieto Alhambra, Yo Han Lee, Sangha Lee, Seong-Ju Kim, Jeewon Lee, Rae Woong Park and Yunmi Shin** on the publication of **Safety outcomes of selective serotonin reuptake inhibitors in adolescent attention-deficit/hyperactivity disorder with comorbid depression: the ASSURE study** in Psychological Medicine.



Psychological Medicine

Article contents

Abstract

Footnotes

References

Safety outcomes of selective serotonin reuptake inhibitors in adolescent attention-deficit/hyperactivity disorder with comorbid depression: *the ASSURE study*

Published online by Cambridge University Press: **20 February 2023**

Chungsoo Kim, Dong Yun Lee, Jimyung Park, Su-Jin Yang, Eng Hooi Tan, Daniel-Prieto Alhambra, Yo Han Lee, Sangha Lee, Seong-Ju Kim and Jeewon Lee ...Show all authors Show author details

Article | Supplementary materials | Metrics

Get access | Share | Cite | Rights & Permissions

Abstract

Background

Attention deficit-hyperactivity disorder (ADHD) is related to depressive disorder, and adolescents with both present poor outcomes. However, evidence for the safety of concomitantly using a methylphenidate (MPH) and a selective serotonin reuptake inhibitor (SSRI) among adolescent ADHD patients is limited, a literature gap aimed to address through this investigation.

Methods

We conducted a new-user cohort study using a nationwide claims database in South Korea. We identified a study population as adolescents who were diagnosed both ADHD and depressive disorder. MPH-only users were compared with patients who prescribed both a SSRI and a MPH. Fluoxetine and escitalopram users were also compared to find a preferable treatment option. Thirteen outcomes including neuropsychiatric, gastrointestinal, and other events were

OHDSI Shoutouts!

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

Do you have anything you want to share? Please send to sachson@ohdsi.org so we can highlight 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	7 am	Medical Imaging
Wednesday	10 am	Surgery and Perioperative Medicine
Wednesday	11 am	Latin America
Thursday	9:30 am	Data Network Quality
Thursday	7 pm	Dentistry
Friday	9 am	GIS – Geographic Information System General
Friday	9 am	Phenotype Development and Evaluation
Friday	10 am	Education
Friday	11 am	Clinical Trials
Monday	10 am	Healthcare Systems Special Interest Group
Tuesday	9 am	OMOP CDM Oncology Genomic Subgroup

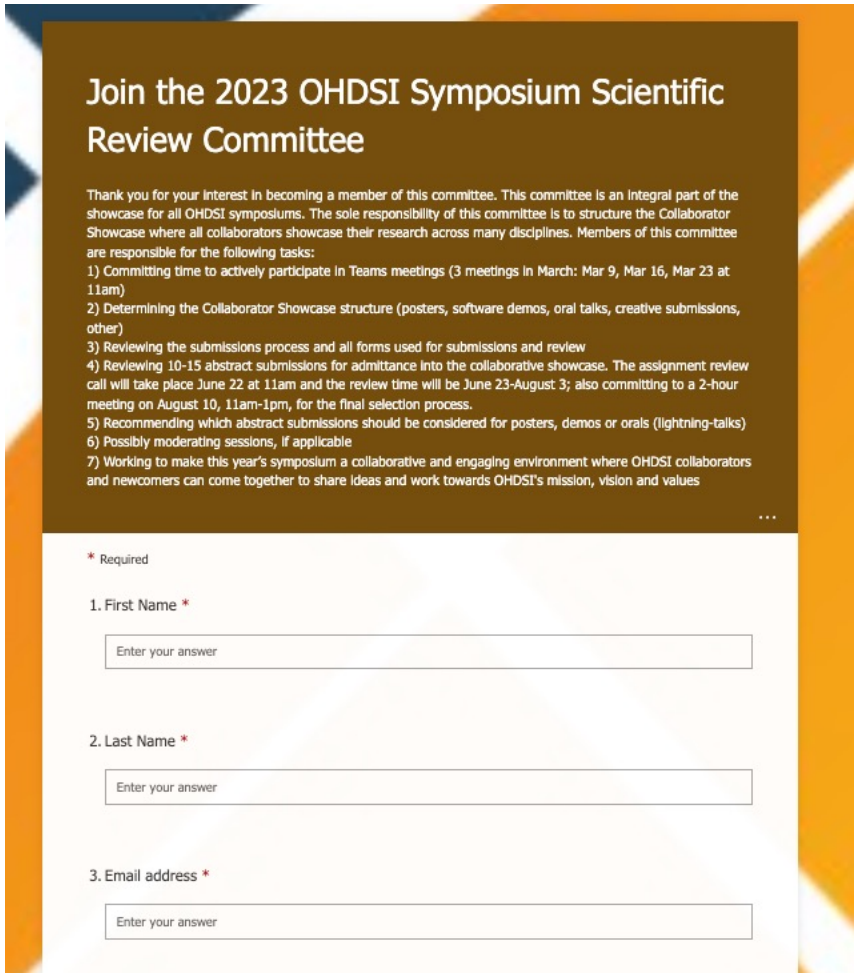
ohdsi.org/workgroups

Global Symposium: Oct. 20-22, 2023

Hilton East Brunswick Hotel & Executive Meeting Center • East Brunswick, N.J.

Join The #OHDSI2023 Scientific Review Committee

We are looking for collaborators to join the OHDSI2023 scientific review committee. **Elisse Katzman** has opened the signup form to join the committee, and the first meeting is scheduled for March 9. **The deadline is Feb. 28.**



bit.ly/OHDSI2023ScientificReview

European Symposium: July 1-3, 2023

APAC Symposium: July 13-14, 2023

2023 APAC Symposium

July 13-14 • University of New South Wales • Sydney, Australia

We are excited to announce that the 2023 OHDSI APAC Symposium will be held in Sydney, Australia at the University of New South Wales! Agenda and registration details are coming soon so please stay tuned! Meanwhile, here are some important dates for you to save to your calendar:

Collaboration Showcase submissions open: Feb. 13

Collaboration Showcase submissions deadline: March 31

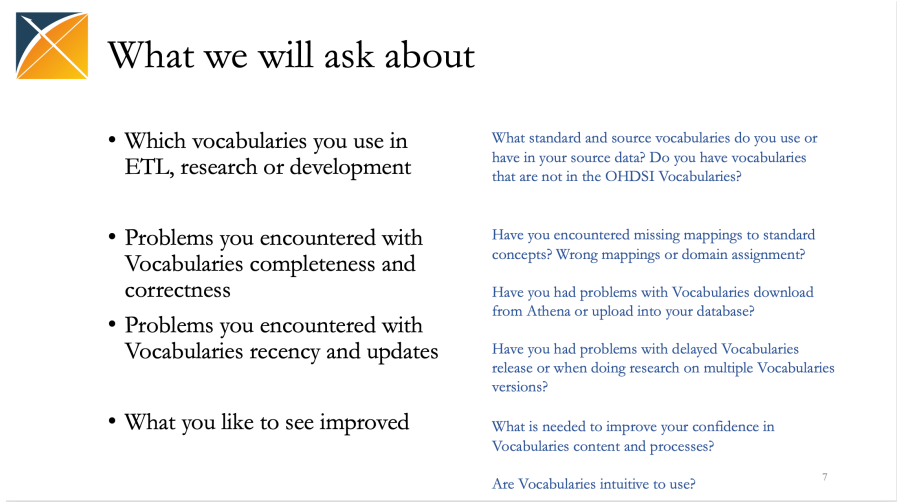
Symposium Day 1, main conference: July 13

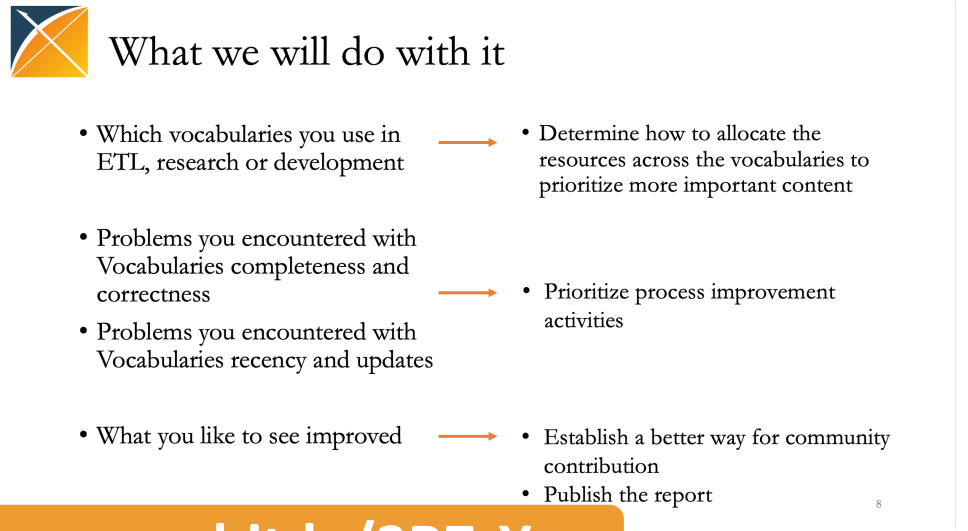
Symposium Day 2, tutorials: July 14

Vocabulary Landscape Assessment

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

The deadline is Feb. 23.





Main vocab: bit.ly/3iTnyco || ETL/Data owners: bit.ly/3R7rYcm

New Center for Advanced Healthcare Research Informatics at Tufts University

Mission

- Produce new modes and methods of using healthcare-generated data and other related data to support research.
- Deploy and use these methods at Tufts Medicine and in research communities that use healthcare data to generate evidence and improve care.
- Collaborate to develop and demonstrate novel and impactful solutions to challenges across the translational research spectrum with the other centers in Tufts ICRHPS and the CTSI, with Tufts University, and with multisite research networks.
- Promote interdisciplinary dialog on the application of advanced informatics and analytics for the improvement of health.

Welcome To Tufts University

Clark Evans

Kyrylo Simonov

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

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

OHDSI HADES releases: ResultModelManager 0.3.0

ResultModelManager 0.3.0 | Reference | Articles | Changelog | HADES

Upload Functionality

James P. Gilbert

2023-02-14

Contents

Introduction

This vignette describes the functionality for uploading results to a pre-created database schema. In the examples here, we assume the use of sqlite for simplicity. However, in principle any platform supported by the `DatabaseConnector` and `SqlRender` packages should work.

Creating a schema definition file

It is recommended that every analytics package that creates data output should contain a csv file. This is a requirement for packages that use the OHDSI `Strategus` library.

Schema definitions should conform to the following column headers:

```
table_name, column_name, data_type, is_required, primary_key
```

In addition, other packages may make use of additional fields such as `optional` or `empty_is_na`. These are not required for

OHDSI HADES releases: ROhdsiWebApi v1.3.2

ROhdsiWebApi 1.3.2 | Reference | Articles | Changelog | HADES

ROhdsiWebApi

R-CMD-check passing | codecov 59%

ROhdsiWebApi is part of HADES.

Introduction

ROhdsiWebApi is a R based interface to 'WebApi' (OHDSI RESTful services), and performs GET/PULL/POST/DELETE calls via the WebApi. All objects starting from R or output to R - are analysis ready R-objects like list and data.frame. The package handles the intermediary steps by converting R-objects to JSON and vice versa. To ensure r-objects are analysis ready, the objects are type converted where possible, e.g. date/date time are converted from string to POSIXct.

This package makes reproducible research easier, by offering ability to retrieve detailed study specifications, transport study specifications from one instance to another, programmatically invoke the generation of a sequence of steps that are part of a study, manage running studies in batch mode.

An example of a WebApi endpoint is "http://server.org:80/WebAPI".

Generic design pattern:

- Functions that go from WebApi to R: will return a R-object (e.g. list, data frame). The function will handle the conversion of non R-objects like JSON objects into R-objects in the background. The returned objects may be inspected in R. We have ensured that all returned objects are represented in the proper data types e.g. dates in are type date, instead of string, use of integer where

Links

Browse source code

Report a bug

Ask a question

License

Apache License 2.0

Citation

Citing ROhdsiWebApi

Developers

Gowtham Rao
Author, maintainer

Martijn Schuemie
Author

Jamie Gilbert
Author

More about authors...

OHDSI HADES releases: CohortExplorer 0.0.11

CohortExplorer 0.0.11 | Reference | Articles | Changelog | HADES

CohortExplorer

R-CMD-check passing | codecov 100%

CohortExplorer is part of HADES.

Introduction

An R package with a Shiny viewer to explore profiles of patients in a cohort. The output of this R-package is a self contained R shiny that contain person level data for review.

Warning

- Contains person level data. This package is not to be considered de-identified.
- Please do not share the output with others as it may violate protected health information.
- .RDS file in output contains PHI.

Features

- From an instantiated cohort, identifies specified number of random persons. It also allows for non random selection by specifying a set of personId as input.

Links

Browse source code

Report a bug

Ask a question

License

Apache License

Citation

Citing CohortExplorer

Developers

Gowtham Rao
Author, maintainer

More about authors...

New EHDEN Academy Course

EHDEN Academy | Courses | Home | Dashboard | My courses | About

One hour of your time: The Phenotyping Problem

Course Overview

Best practice in developing and evaluating a phenotype.

Course Content

Start Date: 3/02/23 Category: Methods

- General
  - Introduction to 1 hour of your time
- The phenotyping problem
- Developing a phenotype
- The Phenotyping Problem -Episode 3: Best Practice in Phenotyping
- Reference material: additional reading and useful links

Enrol Now

Last Updated On 17 February 2023

11 Students Enrolled

Course Features

Lectures	5
Start	3 Feb
Date	2023
Skill level	Beginner
Language	English

2023 DevCon: April 21

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

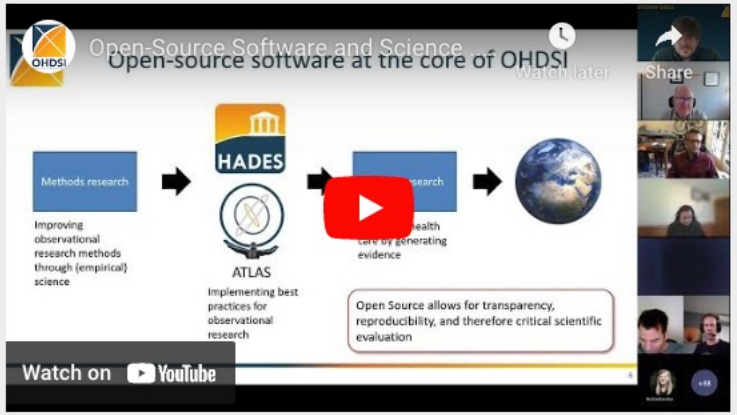
Watch All Eight Workshops, Talks & The Panel From DevCon Below

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.

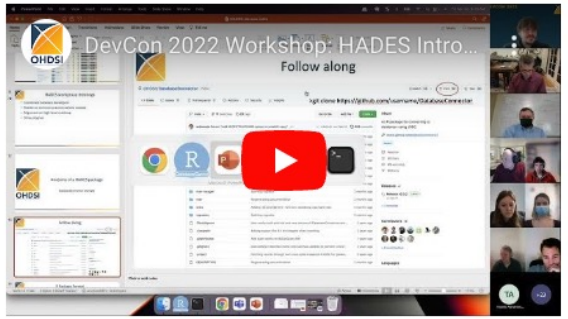
Click Here To Apply For The 2022-23 Khieron Contributor Cohort

Join The Open-Source Community Workgroup in MS Teams

Workshops

ATLAS
(Anthony Sena)

HADES Introduction
(Adam Black)



WebAPI
(Anthony Sena)

Cohort Diagnostics
(James Gilbert)

2023 AMIA Symposium Call For Participation

AMIA 2023 Annual Symposium Call for Participation

We invite you to contribute your best work for presentation at the AMIA 2023 Annual Symposium – the leading symposium for the science and practice of health and biomedical informatics. The AMIA 2023 Annual Symposium showcases submissions from scientists, clinicians, trainees, educators, policy makers, administrators, industry professionals, and technologists from around the world.

The AMIA 2023 Annual Symposium will consider submissions of the following types:

- Paper, Student Paper
- Podium Abstract
- Poster, Panel
- Informatics Debate
- Systems Demonstration
- Workshop

Proposals

Proposals are now being accepted.

Deadline: Mar. 8, 2023

Submit now

Do You Know Of A Collaboration Opportunity?

We are trying to keep the community updated on all collaboration opportunities, both inside AND outside of OHDSI activities. **Marty Alvarez** of Tufts University is doing a fantastic job of compiling them each week so we know what is on the horizon, and we are working on a format to post these for the community.

In the meantime, if you know of any upcoming opportunities (grants, conferences, calls for papers, etc.) that you think should be considered for this list, please send them to **Marta.Alvarez@tuftsmedicine.org**.

Thank you Marty!

Job Opening

Software Dev Analyst II - Res - G&C - CTSI

Job ID: REF9053H
Date posted: 2/20/2023

Employment Type: Full Time
Shift: Days
Location: Boston, MA

PRINCIPAL DUTIES AND ESSENTIAL FUNCTIONS:

Responsible for executing software development initiatives.

Implementation

- Collaborate with various stakeholders to understand requirements and design solutions
- Evaluate options and develop technical design
- Develop solution using appropriate programming language and/or technical tools
- Complete thorough testing of solution
- Provide input to the development of integrated test plan
- Execute integrated test plan
- Provide input to the development of LIVE plan
- Support LIVE activities

Ongoing Enhancements and Support

- Build enhancements to current functionality using appropriate programming language and/or technical tools
- Perform detailed testing of software updates and upgrades
- Communicate in a friendly and professional manner, share the ideas, solutions, the approach, risks, and impacts, set appropriate expectations for the development timeline
- Participate in after-hours on call support rotation for one or more applications which generate Incidents outside of business hours.
- Participate in cross-training, as a trainer and a learner, for personal development and to ensure adequate secondary coverage on all applications

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

Janssen R&D Summer Internships

General Administration

Epidemiology Graduate Intern

General Administration

OHDA Graduate Intern

General Administration

OHDA Undergraduate Intern

General Administration

Data Science RWE for R&D Summer Intern

General Administration

Data Science RWE DevCon Summer Intern

Summer Internships

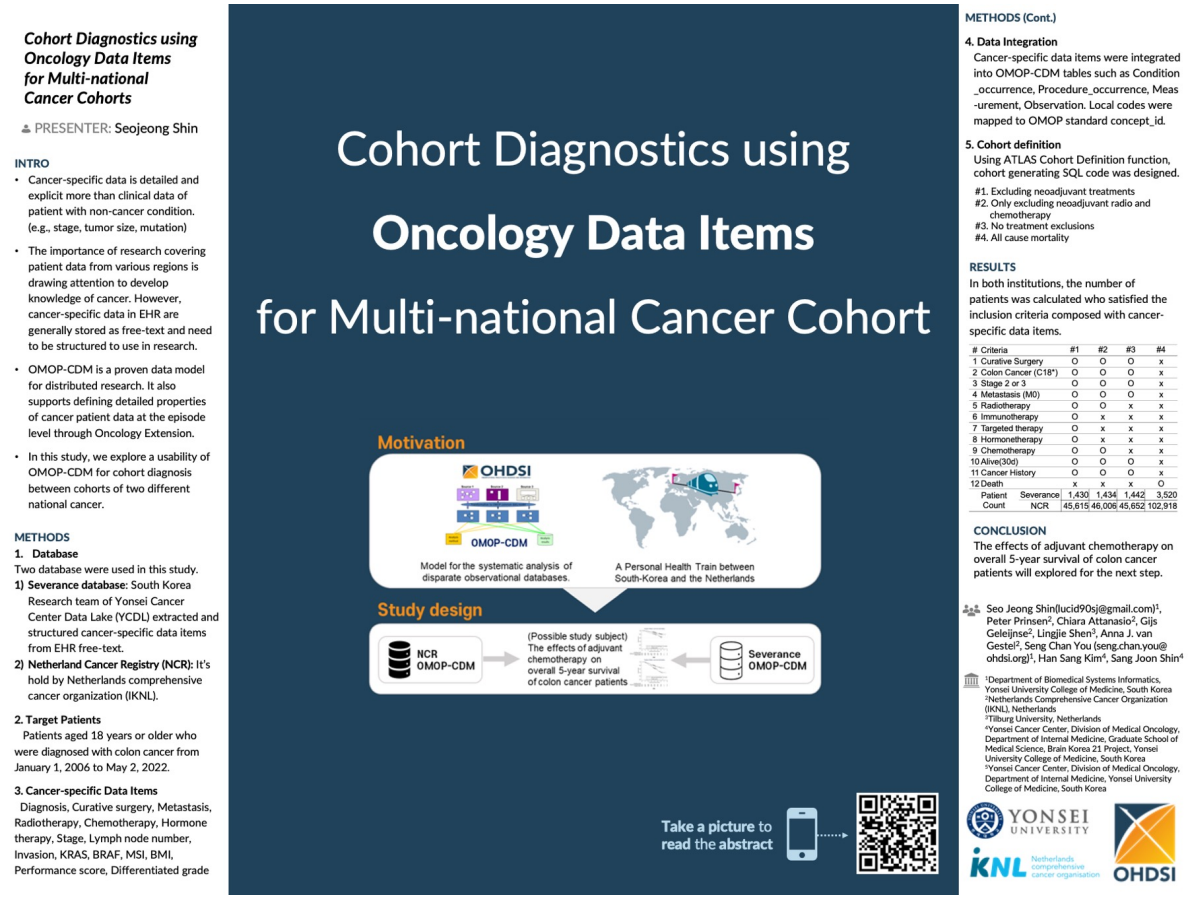
General Administration

Data Science RWE for R&D Summer Intern

General Administration

Data Science RWE DevCon Summer Intern

#OHDSISocialShowcase This Week



Cohort Diagnostics using Oncology Data Items for Multi-national Cancer Cohorts

PRESENTER: Seojeong Shin

INTRO

- Cancer-specific data is detailed and explicit more than clinical data of patient with non-cancer condition. (e.g., stage, tumor size, mutation)
- The importance of research covering patient data from various regions is drawing attention to develop knowledge of cancer. However, cancer-specific data in EHR are generally stored as free-text and need to be structured to use in research.
- OMOP-CDM is a proven data model for distributed research. It also supports defining detailed properties of cancer patient data at the episode level through Oncology Extension.
- In this study, we explore a usability of OMOP-CDM for cohort diagnosis between cohorts of two different national cancer.

METHODS

1. Database

Two database were used in this study.

1) **Severance database**: South Korea Research team of Yonsei Cancer Center Data Lake (YCDL) extracted and structured cancer-specific data items from EHR free-text.
2) **Netherland Cancer Registry (NCR)**: It's hold by Netherlands comprehensive cancer organization (IKNL).

2. Target Patients

Patients aged 18 years or older who were diagnosed with colon cancer from January 1, 2006 to May 2, 2022.

3. Cancer-specific Data Items

Diagnosis, Curative surgery, Metastasis, Radiotherapy, Chemotherapy, Hormone therapy, Stage, Lymph node number, Invasion, KRAS, BRAF, MSI, BMI, Performance score, Differentiated grade

Cohort Diagnostics using

Oncology Data Items

for Multi-national Cancer Cohort

Motivation

OHDSI

OMOP-CDM

Model for the systematic analysis of disparate observational databases.

A Personal Health Train between South-Korea and the Netherlands

Study design

NCR OMOP-CDM

(Possible study subject) The effects of adjuvant chemotherapy on overall 5-year survival of colon cancer patients

Severance OMOP-CDM

Take a picture to read the abstract

METHODS (Cont.)

4. Data Integration

Cancer-specific data items were integrated into OMOP-CDM tables such as Condition _occurrence, Procedure_occurrence, Meas -urement, Observation. Local codes were mapped to OMOP standard concept_id.

5. Cohort definition

Using ATLAS Cohort Definition function, cohort generating SQL code was designed.

#1. Excluding neoadjuvant treatments
#2. Only excluding neoadjuvant radio and chemotherapy
#3. No treatment exclusions
#4. All cause mortality

RESULTS

In both institutions, the number of patients was calculated who satisfied the inclusion criteria composed with cancer-specific data items.

#	Criteria		#1	#2	#3	#4
1	Curative Surgery		O	O	O	x
2	Colon Cancer (C18*)		O	O	O	x
3	Stage 2 or 3		O	O	O	x
4	Metastasis (M0)		O	O	O	x
5	Radiotherapy		O	O	x	x
6	Immunotherapy		O	x	x	x
7	Targeted therapy		O	x	x	x
8	Hormonetherapy		O	x	x	x
9	Chemotherapy		O	O	x	x
10	Alive(30d)		O	O	O	x
11	Cancer History		O	O	O	x
12	Death		x	x	x	O
	Patient Count	Severance	1,430	1,434	1,442	3,520
		NCR	45,615	46,006	45,652	102,918

CONCLUSION

The effects of adjuvant chemotherapy on overall 5-year survival of colon cancer patients will explored for the next step.

Seo Jeong Shin(lucid90sj@gmail.com)[1], Peter Prinsen[2], Chiara Attanasio[2], Gijs Geleijnse[2], Lingjie Shen[3], Anna J. van Gestel[2], Seng Chan You (seng.chan.you@ohdsi.org)[1], Han Sang Kim[4], Sang Joon Shin[4]

[1]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, South Korea
[2]Netherlands Comprehensive Cancer Organization (IKNL), Netherlands
[3]Tilburg University, Netherlands
[4]Yonsei Cancer Center, Division of Medical Oncology, Department of Internal Medicine, Graduate School of Medical Science, Brain Korea 21 Project, Yonsei University College of Medicine, South Korea
[5]Yonsei Cancer Center, Division of Medical Oncology, Department of Internal Medicine, Yonsei University College of Medicine, South Korea

YONSEI UNIVERSITY

KNL Netherlands comprehensive cancer organisation

OHDSI

MONDAY

Cohort Diagnostics using Oncology Data Items for Multi-national Cancer Cohort (Seo Jeong Shin, Peter Prinsen, Chiara Attanasio, Gijs Geleijnse, Lingjie Shen, Anna J. van Gestel, Seng Chan You, Han Sang Kim, SangJoon Shin)

#OHDSISocialShowcase This Week

Framework for Assessing the Reproducibility of Observational Comparative Effectiveness Research: DOACs and Ischemic or Hemorrhagic Events Case Study

PRESENTER: Asieh Golozar

INTRO

- Value in producing real-world evidence on risk and benefits associated with medications
- Concerns over reliability, feasibility and generalizability of produced evidence Mission of Reproducibility Service at OHDSI Center at Northeastern University: to provide a framework for systematic assessment of reliability of evidence generated from real-world data

Objective: Evaluate the reliability of Ray et al: association of anticoagulants (rivaroxaban and apixaban) with major ischemic or hemorrhagic events in patients with atrial fibrillation (1).

METHODS

- **Study Design:** New user active comparator cohort design where confounders adjusted using 1:1 propensity score matching
- **Data:** US claims (IQVIA Open Claims), US Hospital data (IQVIA Hospital CDM) and US electronic health records (Stanford)
- **Study Population:**
  - **Target:** New users of rivoraxaban with prior atrial fibrillation
  - **Comparator:** New users of apixaban with prior atrial fibrillation
- **Outcomes:** Hemorrhagic stroke, ischemic stroke, intracranial bleed and extracranial bleed
- **Sensitivity Analysis:** we changed the definition of the outcomes, altered the exposure groups by timeframe and age, and changed the time at risk.

RESULTS

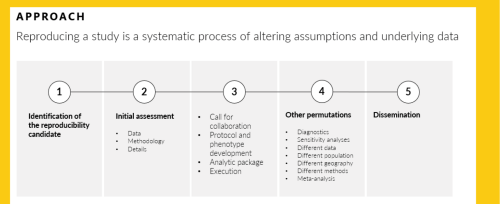
Using Ray's definitions for exposure cohorts, we identified 852,413 patients with atrial fibrillation who were new users of rivaroxaban or apixaban between 2013 and 2018 with similar comorbidities and age distribution in the IQVIA Open Claims database used for replication.

DISCUSSION

- We observed similar effect estimates to the original paper with overlapping confidence intervals, therefore we were able to reproduce the results.
- The evidence from the reproducing data sources suggest no increased risk of the adverse events for rivaroxaban compared to apixaban, which is also exhibited through the sensitivity analysis.
- A systematic approach to reproducible evidence generation requires assessment of the clinical question using multiple permutations of design and data in order to confirm the association of clinical idea in question

OHDSI methodologies and **toolkits** are well equipped to **replicate** and **reproduce research** that might otherwise be **limited** in reliability and generalizability

Figure 1. Systematic and sequential steps in replicating and reproducing a published study



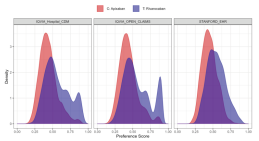


Figure 2: Preference Score Density Plot. From left to right this displays IQVIA Hospital, IQVIA Open Claims and Stanford EHR

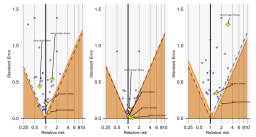


Figure 3: Negative Control Calibration Plot. From left to right this displays IQVIA Hospital, IQVIA Open Claims and Stanford EHR. Only Extracranial Bleed shows positive association across databases

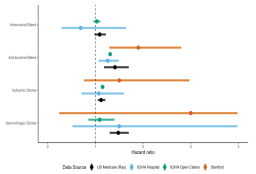


Figure 4: Forest Plot Hazard ratio (HR) (95% confidence intervals (CI)) for rivaroxaban versus apixaban on selected outcomes. The results in black are the original study and the results in other colors are the reproduced results from IQVIA Hospital (sky blue), IQVIA Open Claims (green) and Stanford HER (vermillion). Only Extracranial Bleed has a positive association in all database. Note the HR for Intracranial bleed in Stanford could not be calculated due to small outcome counts

References

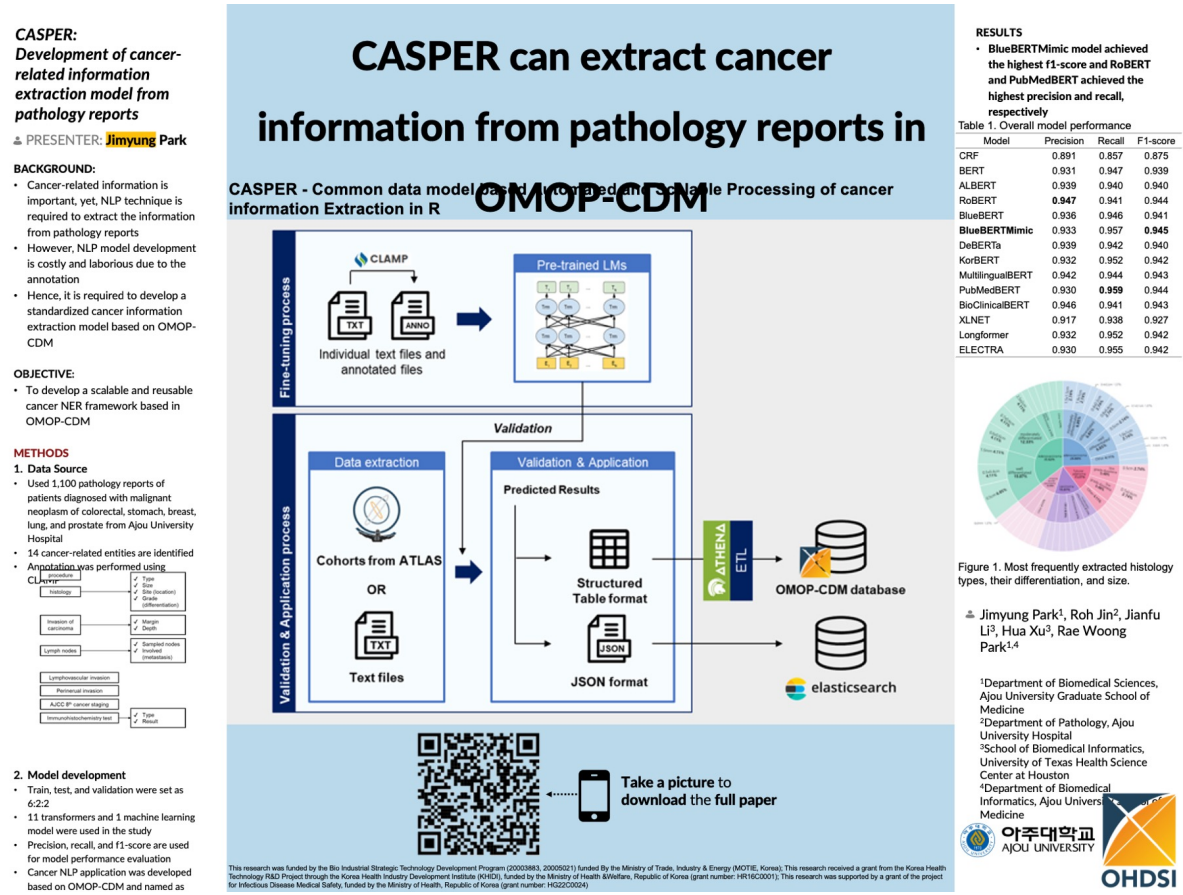
Anna Ostropolets, Martin Lavallee, Adam Black, Ines Neves, Priya Desai, Christian Reich

TUESDAY

Framework for Assessing the Reproducibility of Observational Comparative Effectiveness Research: DOACs and Ischemic or Hemorrhagic Events Case Study

(Asieh Golozar, Anna Ostropolets, Martin LaVallee, Adam Black, Ines Neves, Christian Reich)

#OHDSISocialShowcase This Week



CASPER:
Development of cancer-related information extraction model from pathology reports

PRESENTER: **Jimyung Park**

BACKGROUND:
- Cancer-related information is important, yet, NLP technique is required to extract the information from pathology reports
- However, NLP model development is costly and laborious due to the annotation
- Hence, it is required to develop a standardized cancer information extraction model based on OMOP-CDM

OBJECTIVE:
- To develop a scalable and reusable cancer NER framework based in OMOP-CDM

METHODS

1. Data Source
- Used 1,100 pathology reports of patients diagnosed with malignant neoplasm of colorectal, stomach, breast, lung, and prostate from Ajou University Hospital
- 14 cancer-related entities are identified
- Annotation was performed using CLAMP

2. Model development
- Train, test, and validation were set as 6:2:2
- 11 transformers and 1 machine learning model were used in the study
- Precision, recall, and f1-score are used for model performance evaluation
- Cancer NLP application was developed based on OMOP-CDM and named as

CASPER can extract cancer information from pathology reports in OMOP-CDM

CASPER - Common data model based Automated and Scalable Processing of cancer information Extraction in R

Fine-tuning process: CLAMP, TXT, ANNO, Individual text files and annotated files, Pre-trained LMs

Validation

Validation & Application process: Data extraction, Cohorts from ATLAS, OR, TXT, Text files, Validation & Application, Predicted Results, Structured Table format, JSON format, ATHENA ETL, OMOP-CDM database, elasticsearch

Take a picture to download the full paper

This research was funded by the Bio Industrial Strategic Technology Development Program (20003883, 20005021) funded By the Ministry of Trade, Industry & Energy (MOTIE, Korea); This research received a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health &Welfare, Republic of Korea (grant number: HR16C0001); This research was supported by a grant of the project for Infectious Disease Medical Safety, funded by the Ministry of Health, Republic of Korea (grant number: HG22C0024)

RESULTS
- **BlueBERTMimic model achieved the highest f1-score and RoBERT and PubMedBERT achieved the highest precision and recall, respectively**

Table 1. Overall model performance

Model	Precision	Recall	F1-score
CRF	0.891	0.857	0.875
BERT	0.931	0.947	0.939
ALBERT	0.939	0.940	0.940
RoBERT	**0.947**	0.941	0.944
BlueBERT	0.936	0.946	0.941
BlueBERTMimic	0.933	0.957	**0.945**
DeBERTa	0.939	0.942	0.940
KorBERT	0.932	0.952	0.942
MultilingualBERT	0.942	0.944	0.943
PubMedBERT	0.930	**0.959**	0.944
BioClinicalBERT	0.946	0.941	0.943
XLNET	0.917	0.938	0.927
Longformer	0.932	0.952	0.942
ELECTRA	0.930	0.955	0.942

Figure 1. Most frequently extracted histology types, their differentiation, and size.

Jimyung Park[1], Roh Jin[2], Jianfu Li[3], Hua Xu[3], Rae Woong Park[1,4]

[1]Department of Biomedical Sciences, Ajou University Graduate School of Medicine
[2]Department of Pathology, Ajou University Hospital
[3]School of Biomedical Informatics, University of Texas Health Science Center at Houston
[4]Department of Biomedical Informatics, Ajou Univers... Medicine

아주대학교 AJOU UNIVERSITY
OHDSI

WEDNESDAY

Development of cancer-related information extraction model from pathology reports using transfer learning (Jimyung Park, Roh Jin, Jianfu Li, Hua Xu, Rae Woong Park)

#OHDSISocialShowcase This Week

Secure Federated Hospital Profiling via dGEM Algorithm using Real-World Data

Jiayi Tong[a], Jenna Reps[b,c], Chongliang Luo[d], Jiang Bian[e], Milou Brand[f], Zhaoyi Chen[e], Scott DuVall[g], Thomas Falconer[h], Kevin He[i], Miguel Angel Mayer[j], Bhavnisha Patel[g], Di Wang[f], Hua Xu[k], Yujia Zhou[k], Haitao Chu[l], Martijn Schuemie[b,c], Patrick Ryan[b,c], Rachel M.Werner[m,n], David A. Asch[m,n], Yong Chen[a,n]

a.Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA
b. Janssen Research and Development, Titusville, NJ, USA
c. Observational Health Data Sciences and Informatics (OHDSI), New York, NY
d. Division of Public Health Sciences, Department of Surgery, Washington University in St. Louis, St. Louis, MO, USA
e. College of Medicine, Health Outcomes & Biomedical Informatics, University of Florida, Gainesville, FL, USA
f. IQVIA, England, United Kingdom
g. George E. Whalen Department of Veterans Affairs Health Care System, 500 Foothill Drive, Salt Lake City, UT, USA
h. Department of Biomedical Informatics, Columbia University, New York, NY, USA
i. Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA
j. Hospital del Mar Medical Research Institute (IMIM), Barcelona, Spain
k. School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, Texas, USA
l. Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, MN
m. Division of General Internal Medicine, University of Pennsylvania, Philadelphia, PA, USA
n. Leonard Davis Institute of Health Economics, Philadelphia, PA, USA

Background

- **Hospital profiling** evaluates how much patient outcomes are influenced by the hospital and allows for a quantitative comparison of healthcare providers' quality of care for certain clinical outcomes
- **Generalized linear mixed effects model (GLMM)** models the hospital-specific effects and patient-level effects simultaneously, toward the investigation of hospital profiling
- **Challenges** in utilizing GLMM if the data are stored in a distributed format across sites: patient privacy regulations & computational complexity of GLMM
- **Our goal:**
  - To develop a **decentralized** and **non-iterative** GLMM model which requires only **aggregated data** from each clinical site
  - To provide **accurate** estimated fixed and random effects while **accounting for heterogeneity and hospital-level calibration**
  - To use the dGEM method to study the hospital profiling on the mortality rates of COVID-19 patients with two use cases: distributed OHDSI network and centralized kidney registry data

Methods – dGEM and federated counterfactual modeling

- **Proposed method**: dGEM: Decentralized algorithm for Generalized mixed Effect Models for hospital profiling (example: using COVID-19 mortality status as outcome)
  - *Step 1*: Each site fits a logistic regression model and shares estimated fixed effect
  - *Step 2*: With the meta-estimator obtained with the initial estimates from Step 1, each site estimates the hospital-effects. A meta-regression is used to calibrate the hospital-level characteristics (e.g., volume, number of ICU, local COVID-19 burden, location, etc.)
  - *Step 3*: With the calibrated hospital-effects, each site calculates the counterfactual rates. Hospital ranking based on the measure -- directly standardized mortality rate, which is calculated by adopting the counterfactual modeling idea.
- **Counterfactual modeling** for hospital profiling: through estimating hospital-specific effects, can estimate patient-specific mortality risk as if patient (counterfactually) attended hospital different from the one truly attended
- **Patient-level encryption**: only aggregated data are shared across sites
- **Hospital-level encryption**: each hospital only has the access to a single column of the whole matrix, shared via a 3rd party.

Contact: ychen123@pennmedicine.upenn.edu

Data Application & Results

- **Use case 1: National Kidney Transplant Registry Data from** U.S. Organ Procurement and Transplantation Network (OPTN)
  - Cohort: 44,428 adult deceased donor recipients who experienced a kidney transplant between 01/01/2008 and 12/31/2012 from 201 transplant centers.
- **Use case 2: dGEM-Covid within OHDSI network**
  - Database: 12 sites from US, Spain, Netherlands, and UK
  - Cohort: More than 190,000 patients (365+ days of observation time prior to index date; aged 18+) who have an inpatient visit with a diagnosis of COVID-19 on or during the visit or a positive test for COVID-19 on or during the visits
  - Temporal trend of hospital profiling: two different time periods corresponding to the Wave A and Wave B per data set.
- **Implementation of dGEM-Covid study**
  - Software: PDA-OTA (Privacy-preserving Distributed Algorithm Over the Air)
  - Patient-level data protection: achieved by dGEM algorithm which only requires minimal aggregated data
  - Hospital-level encryption: directly standardized mortality rate based on counterfactual modeling and mask of hospital true ID by assigning a temporal ID (i.e., one-way hash)

THURSDAY

dGEM: Decentralized algorithm for Generalized mixed Effect Models with the Application in Hospital Profiling (Jiayi Tong, Chongliang Luo, Jiang Bian, Milou Brand, Zhaoyi Chen, Scott DuVall, Thomas Falconer, Mengchun Gong, Kevin He, Chung-Soo Kim, Miguel Angel Mayer, Bhavnisha Patel, Di Wang, Hua Xu, Guanjin Yin, Yujia Zhou, David A. Asch, Yong Chen)

#OHDSISocialShowcase This Week

Federated Patient-Level Prediction

1) integrated a common data model and federated learning process
2) federated prediction models with a PLP package

Byungjin Choi *, Dong Yun Lee †, Chungsoo Kim *, Jimyung Park *, Rae Woong Park*, †

INTRO

- Classic machine learning can learn only from centralized data, so the multi-institutional data use is limited.
- Federated learning is a method of developing a multi-institutional model by sharing only weights without sharing data.
- No common pipeline for feature extraction, so researchers at individual institutions had to manually extract
- Our aim is creating framework that unifies clinical data extraction and federated learning

METHODS

1. Training phase

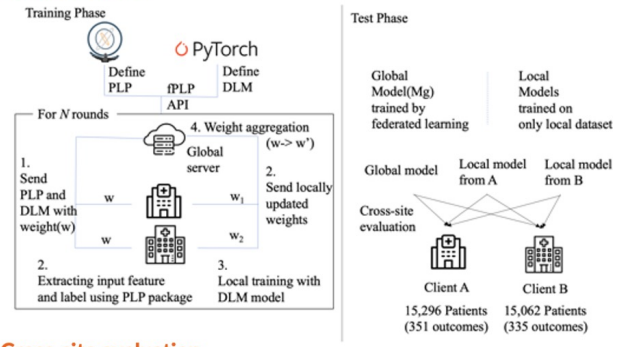
- Specifying cohort and feature settings using ATLAS. Settings are sent to client server
- Using defined PLP settings, labelled dataset were extracted from local OMOP CDM database
- Federated learning started. In each round, the server sends the global model weight to individual clients, global model weights are trained with local cohort and features. Then, only updated weights are sent to the global server. Global server aggregates client weights and makes up new global weights. This process is repeated for a predefined number of rounds by the Flower package.

2. Test Phase

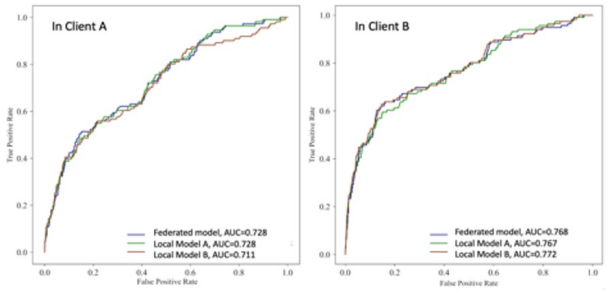
- In the test phase, the global model created through federated learning and all client models trained only with individual client data is distributed to every client server
- In each local server, the global model and all local models are validated on the each client test dataset. This process is called cross-site evaluation.

Federated Patient-Level Prediction : Federated learning library integrated with Patient-level prediction

- **Study Overview**



- **Cross-site evaluation**



3. Proof-of-concept

- For proof-of-concept, we simulated two clients with the Ajou University School of Medicine(AUSOM) CDM database
- Develops federated prediction model of acute kidney injury after coronary intervention .

RESULTS

- In the test dataset of client A, the global model, local model A, local model B showed AUC 0.728, 0.728, 0.711. And in the test dataset of client B, global model, local model A, local model B showed AUC 0.768, 0.767, 0.772
- The local model showed an average AUC reduction of 0.011 in cross-site evaluation. In contrast, the federated learning model shows an performance decrease of AUC 0.002
- This suggests that the federated model has generalizability while showing performance very close to the models trained and tested in same clients
- In further, we plan to develop and validate a federated model in a real-world multi-institutional environment

QR code to **Github page**

*Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, Gyeonggi-do, Republic of Korea

†Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Gyeonggi-do, Republic of Korea

FRIDAY **Federated Patient-Level Prediction (Byungjin Choi, Dong Yun Lee, Chungsoo Kim, Jimyung Park, Rae Woong Park)**

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?

Early-Stage Researchers

Co-leads: Faaizah Arshad & Ross Williams

faaizaha23@g.ucla.edu

ross.williams@ohdsi.org

Meetings

Monthly meet the expert meetings

Second Monday of the month at 17:00 CET

Events at both the OHDSI Europe and OHDSI Global symposiums

WG Mission

To create an inviting venue for junior OHDSI community members early in their careers to navigate OHDSI's resources, ask questions, present their research, find mentorship through networking, and seek insight on their career trajectories.

OKR 1

Objective 1: Facilitate mentorship and networking by holding one career speaker event per month, while making sure to diversify speaker background

Key Results:

Co-leads to hold a planning meeting ahead of each advertised event.

All career speakers are scheduled with a date, meeting link, and flyer, at least 2 weeks before each event.

Greater than 30 participants at each event.

OKR 2

Objective 2: Improve career visibility for members

Key Results:

Create onboarding roadmap for students along different disciplines and career paths (Informatics, Epidemiology, Medicine, Statistics, etc.)

Have at least 5 posters/presentations at the OHDSI symposium from Early-Stage Researchers members.

Hold Meet the Mentor event at OHDSI Global & European symposiums in 2023

Come and join us!

OHDSI Workgroup
Objectives and Key Results (OKR)

Healthcare Systems Interest Group
(formerly the Electronic Health Record Working Group)

Workgroup Name: Healthcare Systems
Workgroup lead: Melanie Philofsky

Objective 1: To provide support for transforming source EHR data to the CDM

2023 Key Results:

1. “Office hours” style agendas during our regularly scheduled meetings 10 times a year.
2. Monitor & answer questions on the CDM Builders, Implementers, and Uncategorized forums related to source data or the ETL process.

Workgroup Name: Healthcare Systems

Objective 2: Support healthcare systems with building the business case to utilize the OMOP CDM

2023 Key Results:

1. Finish survey, review results and publish “Why OMOP”

Workgroup Name: Healthcare Systems

Objective 3: Support health systems in using their CDM to deliver value for their organizations

2023 Key Results:

1. Identify additional analytic use cases from healthcare systems
2. Discuss, describe, and share solutions to use cases brought by community members. This will be recorded and added to the Teams file.
3. Support OHDSI network research by identifying at least 5 OHDSI network research projects and highlighting them to the health system community on the HSIG calls

Eye Care and Vision Research Workgroup

2/21/23

Purpose

- Advance development and implementation of data standards in ophthalmology, optometry, and the vision sciences
- Support studies using observational ophthalmic data for generating insights to improve health and vision outcomes

Accomplishments

- Conducted gap analysis of two large, well-known EHR systems for eye care (Epic and Cerner) to examine where OMOP standards are lacking for commonly used data elements
- Organized in-person meetings at major conferences, including annual meetings of the American Academy of Ophthalmology (AAO) and the Association for Research in Vision and Ophthalmology (ARVO)
- Organized two subgroups focused on subspecialty domain areas (glaucoma and retina)
- Began collaboration with Verana Health for OMOP transformation of the AAO Intelligent Research In Sight (IRIS) Registry, a large nationwide registry of eye care data
- Partnered with the NIH Bridge2AI Initiative to map data elements that will be included in a new Data Generation Project with ophthalmic components (AI-READI)

OKRs for 2023

Objective 1: Formulate data standards development around specific use cases with clear applications for eye care practitioners and vision scientists

- Key Result 1: Solicit a defined use case from the glaucoma subgroup (Q1 2023)
- Key Result 2: Solicit a defined use case from the retina subgroup (Q1 2023)
- Key Result 3: Map ophthalmic data collected for AI-READI Bridge2AI project (Q1 2023)

Objective 2: Develop a clear plan for how visual acuity data should be structured, standardized, and converted to OMOP

- Key Result 1: Engage workgroup members to summarize information about visual acuity and key challenges/considerations (Q1 2023)
- Key Result 2: Present information from Key Result 1 and generate questions for the CDM workgroup to discuss options for concept mapping (i.e. use of an extension table?) and conventions for transforming recorded visual acuity values into standardized measurements (Q1 2023)
- Key Result 3: Submit requests for updated visual acuity standards based on the discussion with the CDM workgroup (Q2 2023)
- Key Result 4: Trial ETL processes at 3 institutions for visual acuity data. (Q4 2023)

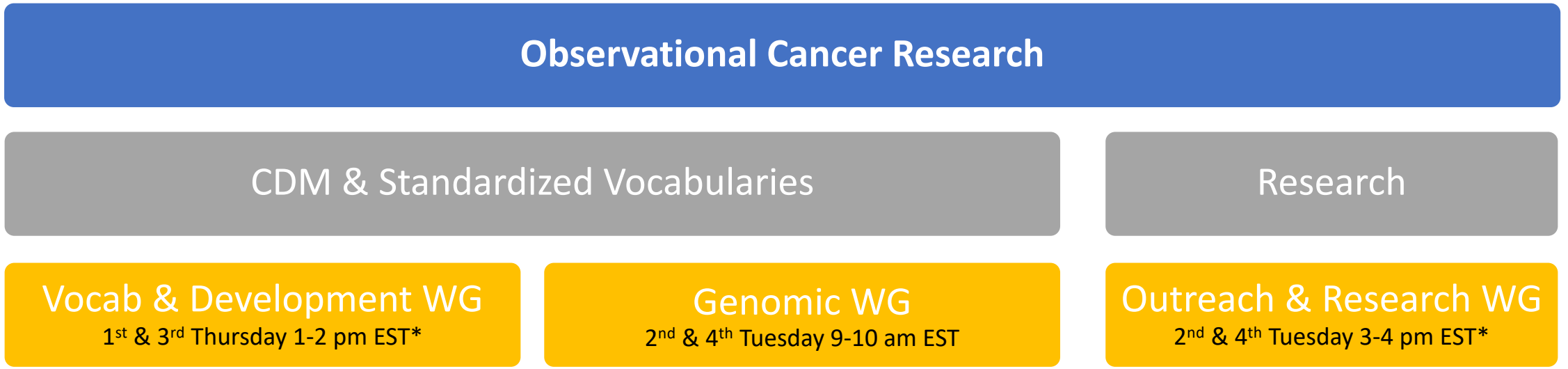
Oncology Working Group- 2023 OKR

Feb 21, 2023

Purpose

- The OMOP Oncology module aims to provide a foundation for representing cancer data at the levels of granularity and abstraction required to support observational cancer research

Oncology WG Structure



* New meeting schedule will be announced soon.

What have we done in 2022?

- **Vocabulary & Development**
  - Cancer modifiers: finalized representation and mapping of metastasis, stage, grade and dimension

- **Genomic**
  - OMOP Genomic Vocabularies: update and addition of new source vocabularies
  - KOIOS – ETL tool to support mapping genomic data

- **Research & Outreach**
  - 360 review and reprioritization of effort to support oncology phenotype development and ongoing studies
  - Initiated new oncology network studies and research collaborations

What is our plan in 2023?

- Focused to make it effective
- Release a stable version of the Oncology module ready for implementation
- Create an opportunity for broader engagement and collaboration

1. Stabilize the vocabularies
2. Fix genomics
3. Run studies
4. Onboarding onRamps

Oncology WG- 2023 Version

- Real-time documentation of the ETL conventions and best practices: Conventions
- Creation of a discussion page to ensure broader engagement of the WG members
- Adopting a use-case driven approach and prioritizing the ongoing effort based on the ongoing and/or future oncology studies. All tracked here: Oncology Release 2.0 project page

Vocab & Development WG OKR-Q1 2023

1. **Improve and finalize ETL from registry data to OMOP**
    - Finish vocabulary mapping for prioritized NAACCR modifier
    - Improve and publish ETL package for mapping tumor registry to regular OMOP Tables
        - Publish a stable release
        - Update NAACCR ETL to use the new conventions
        - Publish and regularly update the documentation on the ETL convention

Oncology Genomic WG OKR-Q1 2023

2. **Continue validation of the vocabulary and model**
   - Assess and publish the completeness of the new vocab release in covering somatic variations in 1) commercial panels and 2) FDA list of approved targets and CAP data items*
   - Publish a list of somatic variants from the FDA and CAP list expected to be in the vocabulary

3. **Improve KOIOS functionality**
   - Collect feedback on KOIOS performance from 4+ sites with available genomic data and propose a plan and timeline for improvement

4. **Publish the documentation of the OMOP Genomic vocabulary**
   - Draft/V1.0 of the genomic vocab documentation

* contingent on the vocab priority and resources

Oncology Research & Outreach WG OKR-Q1 2023

5. **Dissemination and adoption of the Oncology Module**
   - Finalize the outline for Onc Module OnRamps
6. **Support data partners and investigators in the application of the Oncology Module for oncology research**
   - Finalize SWT analytic package
   - Onboard data partners for NSCLC-liver mets study
   - Develop outlines for best practices for oncology phenotype development

Call for Community Engagement in Development & Research

- Data: US tumor registry, non-US tumor registry, EHR, Claims
- Research questions: High impact use cases
- Domain modelers and vocab developers: radiology, surgery, precision medicine
- ETL developers
- Methodologists: Support of best practices

1. **Vocab & Development WG:** 1st & 4rd Thursday 1-2 pm EST*
2. **Genomic WG:** 2nd & 4th Tuesday 9-10 am EST
3. **Outreach & Research WG:** 2nd & 4th Tuesday 3-4 pm EST*

https://github.com/OHDSI/OncologyWG/wiki

Questions: golozar@ohdsi.org

Latin America Workgroup

Jose Posada

Latin America Workgroup Purpose

The OHDSI LATAM WG empowers the Latin American community to generate high-quality evidence with data where our more than 600 million lives are better represented. We are a common touch point across organizations and countries and serve as guides and facilitators towards resources that make the journey possible.

Latin America Workgroup 2023 OKRs

Objective 1: Increase the number of "OHDSI talks" in conferences organized by different scientific organizations in LATAM

Key results

- Four (4) submissions during 2023 by members of the OHDSI LATAM WG on different events mostly taking place in LATAM

Objective 2: Increase the number of new participants in monthly calls

Key results

- Five to ten new persons are invited to every call
- Reach out to 3 to 4 fellow OHDSI members for invited talks
- Send a monthly newsletter.

Latin America Workgroup 2023 OKRs

Objective 3: Facilitate conversions by working with the Vocabulary WG to ensure the inclusion of relevant local vocabularies into the main OHDSi vocabulary.

Key results

- Inclusion of two relevant Colombian vocabularies
- Inclusion of two relevant Brazilian vocabularies
- Inventory list of required vocabulary for Chile and Mexico

Psychiatry Workgroup

Dmitry Dymshyts & Andrew Williams

Purpose

- Expand the vocabulary for representing neuropsychiatric scales and scores to enable OHDSI studies that capture the detailed phenomology of neurologic, psychiatric and neuropsychological conditions.

- Use data representing the phenomena in OHDSI studies that improves understanding and care for patients with these conditions. Provide feedback and assistance to community members who want to conduct OHDSI studies on populations with these conditions.

OHDSI Psychiatry WG previous accomplishments

- Built a first version of neuropsychiatric scales vocabulary.
- Established collaboration with SNOMED on continuing of scales and scores vocabulary work
- Supported the ETL of the psychiatric data to OMOP CDM.
- Supported several network studies related to psychiatry, e.g. CERVELLO.

Psychiatry WG
2023 Objectives and Key Results

- Validate prior concepts created to represent neuropsychiatric scales and scores by assessing the proposed grouping of related concepts in relation to an etiologic- or physiologic-based grouping strategy.
- Support study on health equity relating to patients with psychiatric conditions.
- Clarify the relationship between extended OMOP vocabulary for neuropsychiatric scales and scores and planned restructuring of SNOMED using observable entity approach and a process-and function-based hierarchy instead of a site- and morphology-based hierarchy
- Plan for a large convening activity of experts to contribute to the restructuring of SNOMED to a process-and function-based hierarchy for psychiatric and neuropsychological conditions.

Health Equity Workgroup

Jake Gillberg

Health Equity Workgroup Purpose

The Health Equity Workgroup exists to promote and enable collaboration on health equity work within the OHDSI community.

Health Equity Workgroup 2023 OKRs

Objective 1: Enable collaboration on health equity work, 1 health equity focused study with at least 5 collaborators Q12023, 3 presentations from community members sharing their work Q12023, 3 sessions with external stakeholders to explore opportunities for collaboration on Social Drivers, Risk Factors, and Needs representation in OMOP

Objective 2: Promote health equity work 3 health-equity focused journal articles presented and discussed Q12023

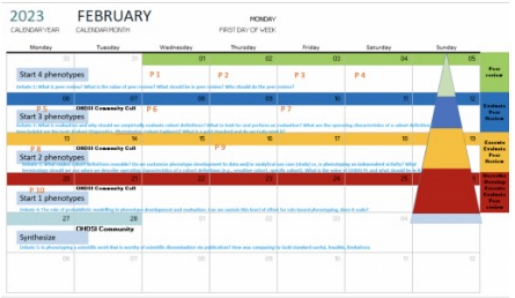
Join Our Workgroups

OHDSI workgroups are always seeking new collaborators. If you are interested in **Joining The Journey** with any of our workgroups, **please visit our sign-up page (see link below or QR code)** and join our global collaborators in the mission to generate the real-world evidence that promotes better health decisions and better care.

ohdsi.org/workgroups

New Phenotype Phebruary Homepage

Phenotype Phebruary 2023: How To Join The Effort

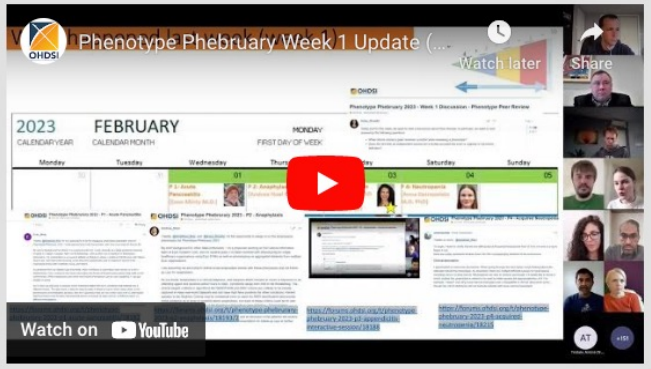


The schedule to the left lists the phenotypes that will be investigated throughout the month, along with the respective leads and reviewers. Check for updates to this graphic as more people join the effort. The graphic to the right highlights the four debates/discussions around phenotyping that are happening this month. Please use the forum links below to join any of these activities.

"Phenotype Phebruary" is a community-wide initiative to both develop and evaluate phenotypes for health outcomes that could be investigated by the community.

This is the second year of Phenotype Phebruary in the OHDSI community (look back at Year 1 here). It was introduced during the Jan. 31 community call (watch here), and will go on throughout the month. This year, the leadership team of **Gowtham Rao** and **Azza Shoaibi** helped identify 10 phenotypes that are being investigated throughout the month. If you would like to join the discussions around any of the phenotypes, please visit the appropriate links below, which will take you to the proper threads on the OHDSI forums.

Week 1 Update



Join Our Community Efforts Around Any Of These Phenotypes

(when phenotype threads get initiated, they will be added to the chart below)

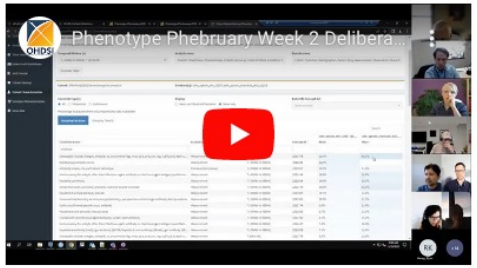
Announcements and Meeting/Workshop Links	Acute Pancreatitis	Anaphylaxis	Appendicitis
Acquired Neutropenia	Systemic Lupus Erythematosus	Acute Hepatic Failure	Idiopathic Inflammatory Myopathies
Parkinson's Disease	ST Elevation Myocardial Infarction	Neonatal Hypoxic Ischemic Encephalopathy	Neurofibromatosis type 1 with Optical Pathway Glioma

Join Our Community Discussions Around These Phenotype Phebruary Topics

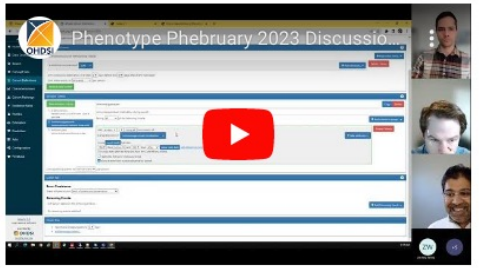
(when phenotype threads get initiated, they will be added to the chart below)

Phenotype Peer Review	Chart review gold standard validation vs innovative methods like PheValuator
What makes cohort definitions reusable, and what is the value of the OHDSI Phenotype Library? What should be in it?	The role of probabilistic modeling in phenotype development and evaluation

Phenotype Phebruary Videos



(Feb. 10) Week 2 of Phenotype Phebruary concluded with this OHDSI Phenotype Development and Evaluation workgroup meeting. In this session, the workgroup assigned leads to each phenotype that are



(Feb. 8) Christopher Mecoli, MD, and team demonstrated progress in the development of a cohort definition for Inflammatory Dermatomyositis at Johns Hopkins University. The team discussed

ohdsi.org/phenotype-phebruary-2023