OHDSI Speed Dating

OHDSI Community Call

Jan. 17, 2022 • 11 am ET

OHDSI
Observational Health Data Sciences and Informatics

Upcoming OHDSI Community Calls

Date	Topic
Jan. 24	Collaborations For Strategic Priorities
Jan. 31	Introduction to Phenotype Phebruary
Feb. 7	Phenotype Phebruary Weekly Update + Workgroup Plans for 2023
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

Upcoming OHDSI Community Calls

Date	Topic
Jan. 24	Collaborations For Strategic Priorities
Jan. 31	Introduction to Phenotype Phebruary
Feb. 7	Phenotype Phebruary Weekly Update + Workgroup Plans for 2023
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

Jan. 24:
Collaborations for Strategic Opportunities

Anna Ostropolets

Data Scientist, Odysseus Data Services, Inc.
PhD Graduate, Columbia University

Clair Blacketer

Director, Janssen Research and Development, Inc.

Patrick Ryan

Vice President, Observational Health Data Analytics, Janssen Research and Development, Inc.; Adjunct Assistant Professor, Columbia University

Three Stages of The Journey

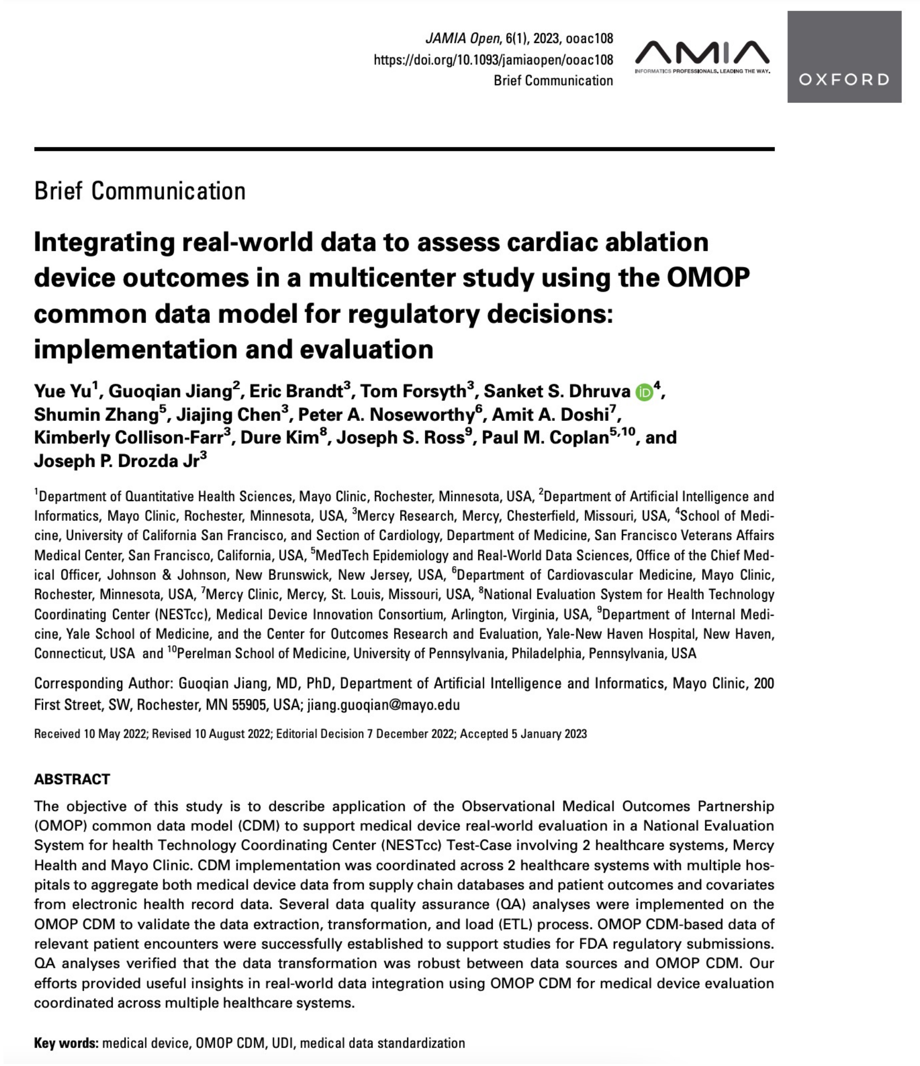
Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **Yue Yu, Guoqian Jiang, Eric Brandt, Tom Forsyth, Sanket Dhruva, Shumin Zhang, Jiajing Chen, Peter Noseworthy, Amit Doshi, Kimberly Collison-Farr, Dure Kim, Joseph Ross, Paul Coplan, and Joseph Drozda** on the publication of **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** in JAMIA Open.



JAMIA Open, 6(1), 2023, ooac108
https://doi.org/10.1093/jamiaopen/ooac108
Brief Communication

Brief Communication

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

Yue Yu[1], Guoqian Jiang[2], Eric Brandt[3], Tom Forsyth[3], Sanket S. Dhruva[4], Shumin Zhang[5], Jiajing Chen[3], Peter A. Noseworthy[6], Amit A. Doshi[7], Kimberly Collison-Farr[3], Dure Kim[8], Joseph S. Ross[9], Paul M. Coplan[5,10], and Joseph P. Drozda Jr[3]

[1]Department of Quantitative Health Sciences, Mayo Clinic, Rochester, Minnesota, USA, [2]Department of Artificial Intelligence and Informatics, Mayo Clinic, Rochester, Minnesota, USA, [3]Mercy Research, Mercy, Chesterfield, Missouri, USA, [4]School of Medicine, University of California San Francisco, and Section of Cardiology, Department of Medicine, San Francisco Veterans Affairs Medical Center, San Francisco, California, USA, [5]MedTech Epidemiology and Real-World Data Sciences, Office of the Chief Medical Officer, Johnson & Johnson, New Brunswick, New Jersey, USA, [6]Department of Cardiovascular Medicine, Mayo Clinic, Rochester, Minnesota, USA, [7]Mercy Clinic, Mercy, St. Louis, Missouri, USA, [8]National Evaluation System for Health Technology Coordinating Center (NESTcc), Medical Device Innovation Consortium, Arlington, Virginia, USA, [9]Department of Internal Medicine, Yale School of Medicine, and the Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, Connecticut, USA and [10]Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Corresponding Author: Guoqian Jiang, MD, PhD, Department of Artificial Intelligence and Informatics, Mayo Clinic, 200 First Street, SW, Rochester, MN 55905, USA; jiang.guoqian@mayo.edu

Received 10 May 2022; Revised 10 August 2022; Editorial Decision 7 December 2022; Accepted 5 January 2023

ABSTRACT

The objective of this study is to describe application of the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) to support medical device real-world evaluation in a National Evaluation System for health Technology Coordinating Center (NESTcc) Test-Case involving 2 healthcare systems, Mercy Health and Mayo Clinic. CDM implementation was coordinated across 2 healthcare systems with multiple hospitals to aggregate both medical device data from supply chain databases and patient outcomes and covariates from electronic health record data. Several data quality assurance (QA) analyses were implemented on the OMOP CDM to validate the data extraction, transformation, and load (ETL) process. OMOP CDM-based data of relevant patient encounters were successfully established to support studies for FDA regulatory submissions. QA analyses verified that the data transformation was robust between data sources and OMOP CDM. Our efforts provided useful insights in real-world data integration using OMOP CDM for medical device evaluation coordinated across multiple healthcare systems.

Key words: medical device, OMOP CDM, UDI, medical data standardization

OHDSI Shoutouts!

Congratulations to the team of **Martijn Schuemie, Fan Bu, Akihiko Nishimura, and Marc Suchard** on the publication of **Adjusting for both sequential testing and systematic error in safety surveillance using observational data: Empirical calibration and MaxSPRT** in Statistics in Medicine.

RESEARCH ARTICLE

Statistics in Medicine WILEY

Adjusting for both sequential testing and systematic error in safety surveillance using observational data: Empirical calibration and MaxSPRT

Martijn J. Schuemie[1,2] | **Fan Bu**[2,3] | **Akihiko Nishimura**[4] | **Marc A. Suchard**[2,3,5]

[1]Observational Health Data Analytics, Janssen Research & Development, Titusville, New Jersey,

[2]Department of Biostatistics, University of California, Los Angeles, California,

[3]Department of Human Genetics, University of California, Los Angeles, California,

[4]Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland,

[5]VA Informatics and Computing Infrastructure, US Department of Veterans Affairs, Salt Lake City, Utah,

Correspondence
Martijn J. Schuemie, Observational Health Data Analytics, Janssen Research & Development, Titusville, NJ, USA.
Email: schuemie@ohdsi.org

Funding information
US Food & Drug Administration, Grant/Award Number: 75F40120D00039

Post-approval safety surveillance of medical products using observational healthcare data can help identify safety issues beyond those found in pre-approval trials. When testing sequentially as data accrue, maximum sequential probability ratio testing (MaxSPRT) is a common approach to maintaining nominal type 1 error. However, the true type 1 error may still deviate from the specified one because of systematic error due to the observational nature of the analysis. This systematic error may persist even after controlling for known confounders. Here we propose to address this issue by combing MaxSPRT with empirical calibration. In empirical calibration, we assume uncertainty about the systematic error in our analysis, the source of uncertainty commonly overlooked in practice. We infer a probability distribution of systematic error by relying on a large set of negative controls: exposure-outcome pairs where no causal effect is believed to exist. Integrating this distribution into our test statistics has previously been shown to restore type 1 error to nominal. Here we show how we can calibrate the critical value central to MaxSPRT. We evaluate this novel approach using simulations and real electronic health records, using H1N1 vaccinations during the 2009–2010 season as an example. Results show that combining empirical calibration with MaxSPRT restores nominal type 1 error. In our real-world example, adjusting for systematic error using empirical calibration has a larger impact than, and hence is just as essential as, adjusting for sequential testing using MaxSPRT. We recommend performing both, using the method described here.

KEYWORDS
empirical calibration, observational research, sequential testing

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
Tuesday	1 pm	Common Data Model
Wednesday	11 am	Open-Source Community
Wednesday	12 pm	Health Equity Journal Club
Thursday	12 pm	HADES
Thursday	1 pm	OMOP CDM Oncology – Vocabulary/Development
Thursday	7 pm	Dentistry
Friday	9 am	GIS – Geographic Information System Development
Friday	11 am	Clinical Trials
Monday	10 am	Africa Chapter
Monday	11 am	Early-Stage Researchers
Tuesday	9 am	OMOP CDM Oncology

ohdsi.org/workgroups

OHDSI HADES releases: SqlRender 1.11.1

SqlRender 1.11.1 Reference Articles SqlDeveloper Changelog HADES

SqlRender

R-CMD-check passing | codecov 80% | CRAN 1.11.1 | downloads 1990/month

SqlRender is part of HADES.

Introduction

This is an R package for rendering parameterized SQL, and translating it to different SQL dialects. SqlRender can also be used as a stand-alone Java library and a command-line executable.

Features

- Supports a simple markup syntax for making SQL parameterized, and renders parameterized SQL (containing the markup syntax) to executable SQL
- The syntax supports defining default parameter values
- The syntax supports if-then-else structures
- Has functions for translating SQL from one dialect (Microsoft SQL Server) to other dialects (Oracle, PostgreSQL, Amazon RedShift, Impala, IBM Netezza, Google BigQuery, Microsoft PDW, Snowflake, Azure Synapse, Apache Spark and SQLite)
- Can be used as R package, Java library, or as stand-alone executable through a command-line interface

Links

View on CRAN

Browse source code

Report a bug

Ask a question

License

Apache License 2.0

Citation

Citing SqlRender

Developers

Martijn Schuemie
Author, maintainer

Marc Suchard
Author

OHDSI HADES releases: EvidenceSynthesis 0.4.0

EvidenceSynthesis 0.4.0 | Reference | Articles | Changelog | HADES

EvidenceSynthesis

R-CMD-check passing | codecov 79% | CRAN 0.4.0 | downloads 271/month

EvidenceSynthesis is part of HADES.

Introduction

This R package contains routines for combining causal effect estimates and study diagnostics across multiple data sites in a distributed study. This includes functions for performing meta-analysis and forest plots.

Features

- Perform a traditional fixed-effects or random-effects meta-analysis, and create a forest plot.
- Use non-normal approximations of the per-data-site likelihood function to avoid bias when facing small and zero counts.

Example

Links

View on CRAN

Browse source code

Report a bug

Ask a question

License

Apache License 2.0

Citation

Citing EvidenceSynthesis

Developers

Martijn Schuemie
Author, maintainer

Marc A. Suchard
Author

More about authors...

OHDSI HADES releases: SelfControlledCaseSeries 4.1.0

SelfControlledCaseSeries 4.1.0 | Reference | Articles | Changelog | HADES

SelfControlledCaseSeries

R-CMD-check passing | codecov 87%

SelfControlledCaseSeries is part of HADES.

Introduction

SelfControlledCaseSeries is an R package for performing Self-Controlled Case Series (SCCS) analyses in an observational database in the OMOP Common Data Model.

Features

- Extracts the necessary data from a database in OMOP Common Data Model format.
- Optionally add seasonality using a spline function.
- Optionally add age using a spline function.
- Optionally add calendar time using a spline function.
- Optionally correct for event-dependent censoring of the observation period.
- Optionally add many covariates in one analysis (e.g. all drugs).
- Options for constructing different types of covariates and risk windows, including pre-exposure windows (to capture contra-indications).

Links

Browse source code

Report a bug

Ask a question

License

Apache License 2.0

Citation

Citing SelfControlledCaseSeries

Developers

Martijn Schuemie
Author, maintainer

Patrick Ryan
Author

Trevor Shaddox
Author

Marc Suchard
Author

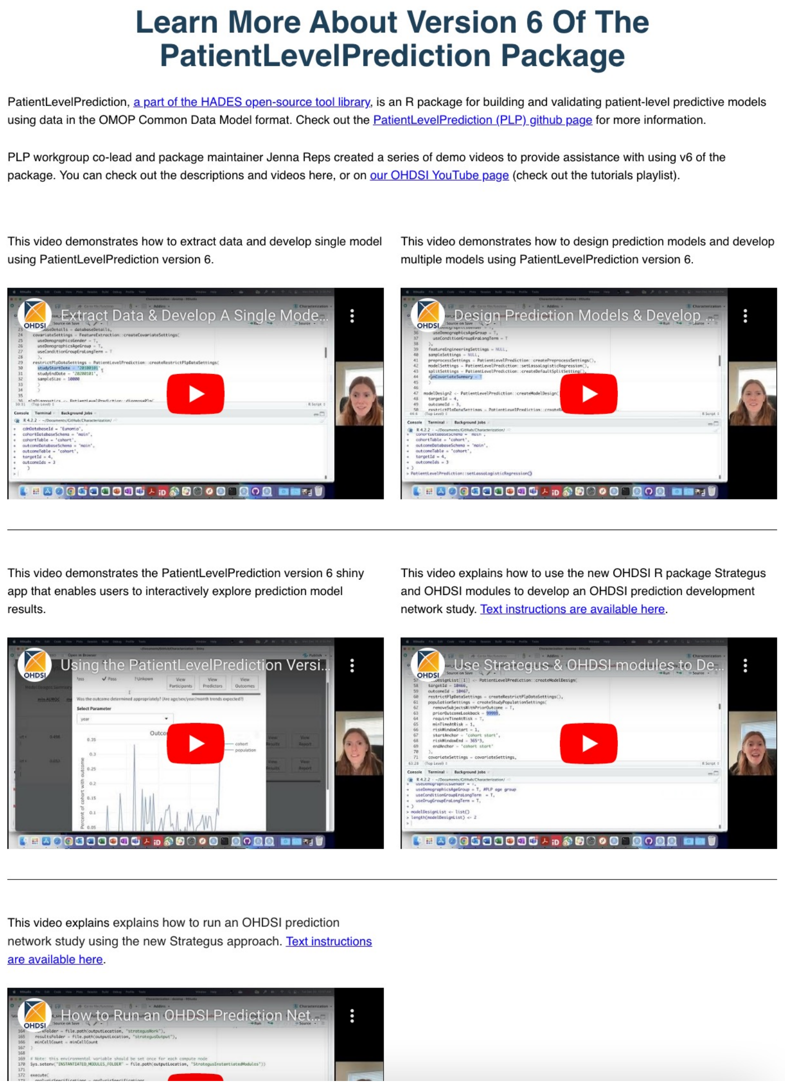
New Demos: PatientLevelPrediction v6/Strategus

Jenna Reps, co-lead of the PLP workgroup, recently shared several video tutorials of version 6 of the PatientLevelPrediction tool. The demos are available on both our website and our YouTube page.

Videos

- how to extract data and develop single model using PLP v6
- how to design prediction models and develop multiple models using PLP v6
- demonstrating the PLP v6 shiny app that enables users to interactively explore prediction model results
- how to use the new OHDSI R package Strategus and OHDSI modules to develop an OHDSI prediction development network study
- how to run an OHDSI prediction network study using the new Strategus approach

ohdsi.org/plp-v6-demos/



Learn More About Version 6 Of The PatientLevelPrediction Package

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. Check out the PatientLevelPrediction (PLP) github page for more information.

PLP workgroup co-lead and package maintainer Jenna Reps created a series of demo videos to provide assistance with using v6 of the package. You can check out the descriptions and videos here, or on our OHDSI YouTube page (check out the tutorials playlist).

This video demonstrates how to extract data and develop single model using PatientLevelPrediction version 6.

Extract Data & Develop A Single Mode...

This video demonstrates how to design prediction models and develop multiple models using PatientLevelPrediction version 6.

Design Prediction Models & Develop ...

This video demonstrates the PatientLevelPrediction version 6 shiny app that enables users to interactively explore prediction model results.

Using the PatientLevelPrediction Versi...

This video explains how to use the new OHDSI R package Strategus and OHDSI modules to develop an OHDSI prediction development network study. Text instructions are available here.

Use Strategus & OHDSI modules to De...

This video explains explains how to run an OHDSI prediction network study using the new Strategus approach. Text instructions are available here.

How to Run an OHDSI Prediction Net...

Oxford Real World Evidence Summer School

Oxford Summer School 2023: Real World Evidence using the OMOP Common Data Model

COURSE DIRECTORS

Daniel Prieto-Alhambra

Professor of Pharmaco- and Device Epidemiology

COURSE ADMINISTRATOR

Mahkameh Mafi

Personal Assistant to Professor Prieto-Alhambra

OTHER COURSES

Statistics: Designing clinical research and biostatistics

Brief Description:

Our Real World Evidence Summer School will provide participants with the tools and concepts necessary to plan and execute Real World Evidence studies, with a focus on the use of the OMOP common data model. The course will have morning lectures followed by afternoon practicals where concepts discussed in the morning will be put in practice with hands-on sessions. Practical sessions will have two tracks: a) for those interested in the design of studies and use of existing analytical and data curation tools; and b) for more advanced data scientists and programmers interested in the development or modification of analytical code using R.

Registration: It is now open

Venue: Lady Margaret Hall Talbot Hall Theatre, Norham Gardens, Oxford OX2 6QA

Date: 19th- 23rd June 2023

For booking please use **Booking information**

Please see the Preliminary Programme here

AUDIENCE:

Pharmacists, clinicians, academics (including statisticians, epidemiologists, and related MSc/PhD students); Industry (pharmacy or device) or Regulatory staff with an interest in the use of routinely collected data for research.

LEARNING GOALS:

OHDSI Calendar of collaboration opportunities

January

SU	MO	TU	WE	TH	FR	SA
1	2	3	4	5	6	7
8	9	10	11	12	13	14
	ICHPS, Scottsdale AZ			ISPOR due		
		2023 kickoff				
15	16	17	18	19	20	21
		speed dating				
22	23	24	25	26	27	28
		2023 priorities			WG OKR due	
29	30	31				
		phenotype				

February

SU	MO	TU	WE	TH	FR	SA
			1	2	3	4
			Phenotype Phebruary			
			JSM due			
5	6	7	8	9	10	11
		Phenotype Phebruary				
12	13	14	15	16	17	18
		Phenotype Phebruary				
	ICPE due					
19	20	21	22	23	24	25
		Phenotype Phebruary				
26	27	28				
Phenotype Phebruary						

March

SU	MO	TU	WE	TH	FR	SA
			1	2	3	4
					WG Leader sum	
5	6	7	8	9	10	11
		SOS vote opens				
12	13	14	15	16	17	18
	AMIA Summit, Seattle WA US					
19	20	21	22	23	24	25
26	27	28	29	30	31	
		SOS challenge kickoff				

April

SU	MO	TU	WE	TH	FR	SA
						1
2	3	4	5	6	7	8
SOS challenge: data diagnostics						
9	10	11	12	13	14	15
SOS challenge: phenotype development						
			MLHC due			
16	17	18	19	20	21	22
SOS challenge: phenotype evaluation						
				DevCon tbd		
23	24	25	26	27	28	29
SOS challenge: analysis design						
30						

May

SU	MO	TU	WE	TH	FR	SA
	1	2	3	4	5	6
SOS challenge: network execution						
7	8	9	10	11	12	13
ISPOR, Boston MA US						
SOS challenge: diagnostics, evidence synthesis						
14	15	16	17	18	19	20
SOS challenge: results interpretation, dissemination						
21	22	23	24	25	26	27
			ACIC, Austin TX US			
	Rstudio, Dallas Tx					
28	29	30	31			

June

SU	MO	TU	WE	TH	FR	SA
				1	2	3
					OHDSI EU	
4	5	6	7	8	9	10
	AIPM, Portland ME US					
					Rotterdam, NL	
11	12	13	14	15	16	17
		SER, Portland OR US			FR-SUN in June	
18	19	20	21	22	23	24
					Date TBD	
25	26	27	28	29	30	
	DIA, Boston MA US					

July

SU	MO	TU	WE	TH	FR	SA
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
MedInfo, Sydney Australia					OHDSI APAC	
16	17	18	19	20	21	22
Sydney						
23	24	25	26	27	28	29
30	31					

August

SU	MO	TU	WE	TH	FR	SA
		1	2	3	4	5
6	7	8	9	10	11	12
JSM, Toronto, ON, CA					ML4HC, NY US	
13	14	15	16	17	18	19
20	21	22	23	24	25	26
		ISPE, Halifax, Nova Scotia, CA				
27	28	29	30	31		

September

SU	MO	TU	WE	TH	FR	SA
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

October

SU	MO	TU	WE	TH	FR	SA
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
					OHDSI Global	
22	23	24	25	26	27	28
tbd						
29	30	31				

November

SU	MO	TU	WE	TH	FR	SA
			1	2	3	4
5	6	7	8	9	10	11
	ISOP, Bali, Indonesia					
12	13	14	15	16	17	18
AMIA, New Orleans, US						
19	20	21	22	23	24	25
26	27	28	29	30		

December

SU	MO	TU	WE	TH	FR	SA
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
		2023 year in review				
17	18	19	20	21	22	23
		holiday party				
24	25	26	27	28	29	30
31						

OHDSI community calls
OHDSI community events
OHDSI collaboration activities

External conferences
Deadlines

ICPE 2023 Abstract Deadline: Feb. 13

August 23 - 27

HALIFAX, NOVA SCOTIA, CANADA
HALIFAX CONVENTION CENTRE

ispe

pharmacoepi.org
#ICPE23 | @IntPharmacoEpi

ICPE 2023 Call for Abstracts
Submission Deadline: February 13, 2023

Abstract submissions for the 39th International Conference on Pharmacoepidemiology and Therapeutic Risk Management (ICPE 2023) are now being accepted online

Call for Abstracts
ICPE 2023 will be a live event held at the Halifax Convention Centre, Halifax, Nova Scotia, Canada, August 23-27, 2023. Virtual presentations are not permitted for the event; all presentations must be delivered in person. If you submit an abstract, it is with the intention that you will physically attend the conference to present it.

The ICPE 2023 is a unique forum for the exchange of scientific information from the fields of pharmacoepidemiology and therapeutic risk management among those in the pharmaceutical industry, government, academia, service

pharmacoepi.org/meetings/annual-conference/

#OHDSISocialShowcase This Week

Extending the OMOP Standard Vocabulary to Include Botanical Natural Products

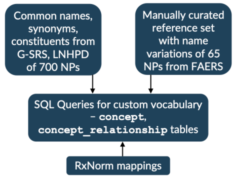
Sanya B. Taneja, Mary F. Paine, Sandra L. Kane-Gill, Richard D. Boyce

INTRODUCTION

OBJECTIVE: extend the OMOP vocabulary to include natural products, their synonyms, phytoconstituents, and name variations to standardize the natural product reports in spontaneous reporting systems.

- Increase in consumption of natural products and/or dietary supplements has led to adverse event concerns.
- Spontaneous reporting systems (e.g., FAERS) can be used for natural product pharmacovigilance by identifying reports with natural products.
- Lack of interoperability in natural product data sources, coverage of synonyms, scientific names and common names, and ambiguity in natural product names are major challenges.

METHODS



RESULTS

- 303 unique natural product Latin binomials
- 2,289 unique concepts in `concept` table
- 2,772 manually curated name variations for 65 natural products from FAERS
- Relationships: *napdi_pt*, *napdi_is_pt_of*, *napdi_has_const*, *napdi_is_const_of*, *napdi_spell_vr*, *napdi_is_spell_vr_of*, *napdi_np_maps_to*, *napdi_const_maps_to*
- 47,601 reports matched to natural product names, 60,223 reports matched to natural product names & name variations, & 100,522 reports matched to natural product constituents.

303 botanical natural products, 2,289 concepts, and **2,772 name variations** added to **extended OHDSI vocabulary**.

160,745 adverse event reports identified using terms for 65 natural products from the extended vocabulary.

Includes relationships to natural product constituents and RxNorm concepts.

Take a picture to download the full paper

EXTENDED VOCABULARY

Table 1: `concept` table with green tea concepts.

concept_id	concept_name	vocab_id	concept_class_id
-7000189	Black tea [Camellia sinensis]	NAPDI	Green tea
-7000190	Green tea [Camellia sinensis]	NAPDI	Green tea
-7000191	Oolong tea [Camellia sinensis]	NAPDI	Green tea
-7000192	Tea [Camellia sinensis]	NAPDI	Green tea
-7000193	White Tea [Camellia sinensis]	NAPDI	Green tea
-7000293	Camellia sinensis [Camellia sinensis]	NAPDI	Green tea

Table 2: `concept` table with green tea constituents.

concept_name	constituent_name	concept_id
Green tea	EPICATECHIN	-7001895
Green tea	EPICATECHIN GALLATE	-7002175
Green tea	EPIGALLOCATECHIN	-7001785
Green tea	EPIGALLOCATECHIN GALLATE	-7002248
Green tea	GALLOCATECHIN	-7002061
Green tea	GALLOCATECHIN GALLATE	-7001793

Table 3: `concept` table with green tea name variations.

concept_name	name_variation	concept_id
Green tea	GUARANA GREEN TEA	-7004112
Green tea	CAMELLIA SINENSIS/PANAX GINSENG EXTRACT	-7004069
Green tea	APPLE CIDER VINEGAR + GREEN TEA SUPPLEMENT	-7003800
Green tea	UNSPECIFIED GREEN TEA EXTRACT SUPPLEMENT	-7002714
Green tea	TEA, GREEN (TEA, GREEN)	-7002713

Table 4: Green tea concepts mapped to RxNorm terms.

rxnorm_id	napdi_concept_id	rxnorm_concept	rxnorm_class
19121499	-7001008	GREEN TEA PREPARATION 25 MG	Clinical Drug Comp
1304239	-7001008	GREEN TEA LEAF EXTRACT	Ingredient
1304273	-7001008	GREEN TEA LEAF EXTRACT 1000 MG ORAL TABLET	Clinical Drug
1396861	-7001008	GREEN TEA EXTRACT 315 MG ORAL CAPSULE	Clinical Drug

Details: github.com/dbmi-pitt/np-terminology-imports

FAERS: FDA Adverse Event Reporting System

University of Pittsburgh

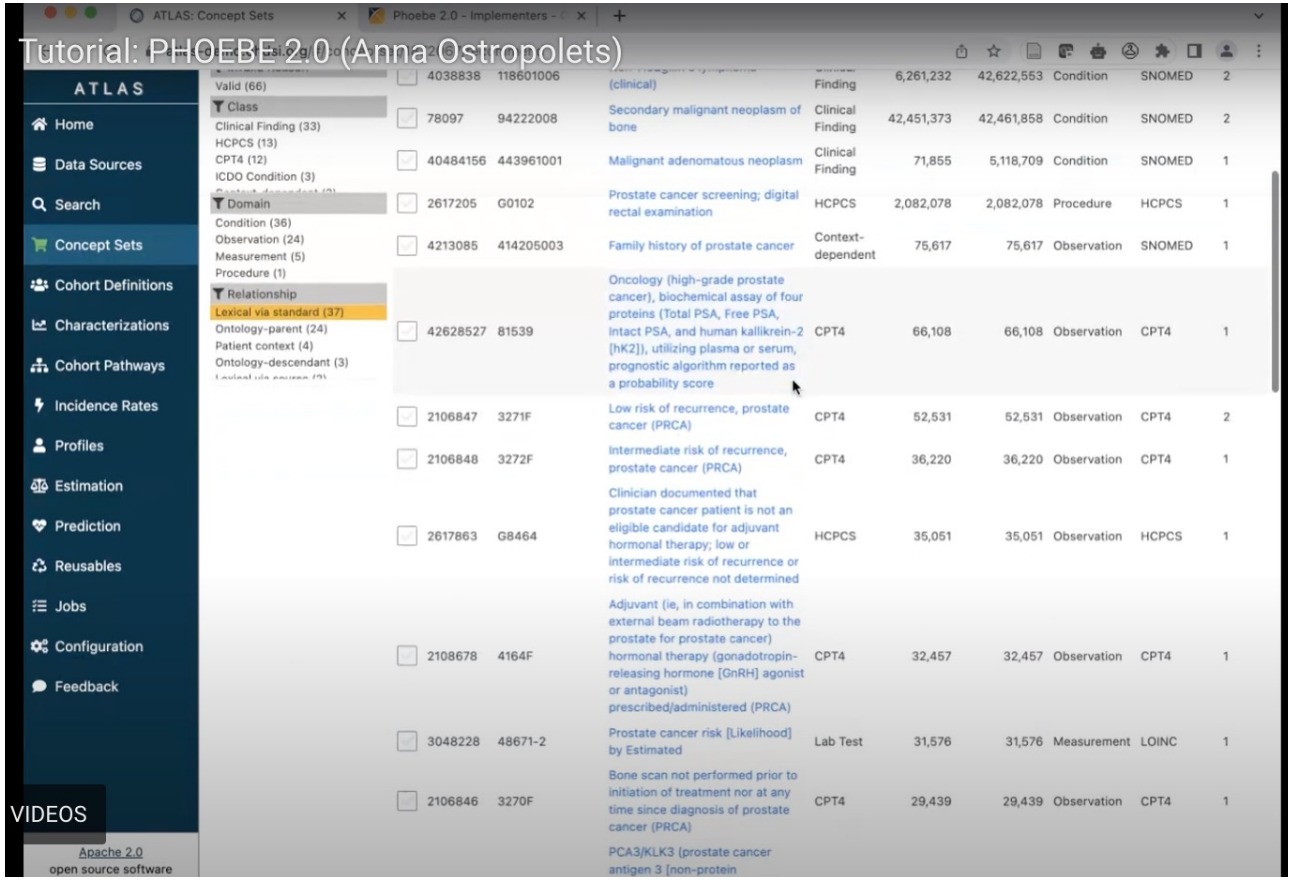
NaPDI Center of Excellence for Natural Product-Drug Interaction Research

OHDSI

TUESDAY

Extending the OMOP Standard Vocabulary to Include Botanical Natural Products (Sanya B. Taneja, Mary F. Paine, Sandra L. Kane-Gill, Richard D. Boyce)

#OHDSISocialShowcase This Week



WEDNESDAY

PHOEBE 2.0: selecting the right concept sets for the right patients using lexical, semantic, and data-driven recommendations (Anna Ostropolets, George Hripcsak, Christopher Knoll, Patrick Ryan)

#OHDSISocialShowcase This Week

Adaptation and Validation of the Charlson Comorbidity Index In Administrative Claims Data Using the SNOMED CT Standardized Vocabulary

Stephen P Fortin[1], Jenna Reps[1], Patrick Ryan[1]
[1]Janssen R&D, LLC, Raritan, NJ, USA

Background

- The Charlson Comorbidity Index (CCI) is commonplace in epidemiological research providing an aggregate measure of patient comorbidity
- Multiple adaptations of the CCI have emerged for application to administrative claims data using the International Classification of Diseases, Ninth and Tenth Revisions (ICD-9/10) and their clinical modifications (ICD-9/10-CM)
- No prior literature exists describing the development and validation of the current implementation of the CCI used by the OHDSI community; and prior research has shown large discrepancies in the measurement of comorbidities between the OHDSI and Quan adaptations of the CCI.

Study Objectives: To address the limitations of the OHDSI adaptation, the current study adapted and validated a new coding algorithm for the CCI using the SNOMED CT standardized vocabulary, henceforth referred to as the SNOMED adaptation.

Methods

Development of SNOMED Adaptation: Adapted through direct translation of Quan coding algorithms followed by manual curation by clinical subject matter experts
Code Mapping Diagnostics: For transparency, all discrepant codes between the SNOMED and Quan adaptations of the CCI were identified and classified into the following categories:

- **Multiple ICD codes to single SNOMED CT code**
  - **Information Gain**: Discrepant codes are clinically relevant
  - **Added Noise**: Discrepant codes are not clinically relevant
- **Deprecated ICD code unmapped to SNOMED CT code**
- **Specificity of ICD code mapping to SNOMED CT code**

Validation of SNOMED Adaptation:
Study Design: Descriptive study
Data Source: Data were from two U.S. administrative claims databases:

1. IBM® MarketScan® Multi-State Medicaid Database (MDCD)
2. Optum® De-Identified Clinformatics Data Mart Database – Date of Death (DOD)

Study Population: Patients aged ≥18 years with an inpatient visit during the calendar years of 2013 or 2018 with at least 365 days of prior observation (index = first inpatient visit)
Covariates: The CCI and each comorbid condition comprising the CCI were measured based on all observed diagnosis codes recorded at or any time prior to index using the SNOMED vs. Quan adaptations.
Statistical Analysis

- All analyses were stratified by database and calendar year.
- Descriptive statistics were produced for each study covariate. Differences between study comparison groups were assessed using standardized mean differences (SMD; SMD >0.10 considered imbalanced).
- For each comorbid condition, the overlap in patient capture between the SNOMED and Quan adaptations between study comparison groups was measured.
- Logistic regression was used to predict 1-year mortality using the CCI as the independent variable. The predictive performance of each vocabulary was assessed using the c statistic, measured as the area under the curve of the receiver operating characteristics curve.

Results

As shown in Table 1, among 5,343 codes mapping to ICD-9/10-CM from either the SNOMED or Quan adaptation, 695 (13.0%) were discrepant codes. The mapping of multiple ICD codes to a single SNOMED CT code was the most common cause of discrepant codes (n=560; 80.6%), which resulted in the additional capture of clinically relevant codes (i.e., information gain) in 24.6% (n=138) of cases.

Table 1. Summary of findings from code mapping diagnostics

	Concordant Codes	Multiple ICD Codes to Single SNOMED CT Code		Deprecated ICD code unmapped to SNOMED CT Code	Specificity of ICD code mapping to SNOMED CT Code
		Information gain	Added Noise		
N	4,648 (87.0%)	138 (2.6%)	422 (7.9%)	123 (2.3%)	12 (0.2%)

A total of 328,740 (MDCD, 2013), 804,707 (DOD, 2013), 491,311 (MDCD, 2018), and 1,109,389 (DOD, 2018) met the study criteria for each database-calendar year combination.

Figures 1-4 show the overlap in patient capture for each comorbidity for each database-calendar year combination. The degree of patient overlap in patient capture was >97.6% across all comorbid conditions. There was no significant difference in the SMD of comorbid conditions between study comparison groups.

As shown in **Table 2**, the performance of the SNOMED and Quan adaptations in predicting one-year mortality was similar across all analyses.

Table 2. Performance of the SNOMED versus Quan adaptations of the CCI in predicting one-year mortality

Year	Adaptation of CCI	MDCD, c-statistic (95% CI)	DOD, c-statistic (95% CI)
2013	SNOMED	0.725 (0.721, 0.728)	0.789 (0.787, 0.79)
	Quan	0.723 (0.72, 0.726)	0.787 (0.786, 0.789)
2018	SNOMED	0.754 (0.751, 0.757)	0.757 (0.755, 0.758)
	Quan	0.752 (0.749, 0.754)	0.757 (0.756, 0.758)

Conclusions

The SNOMED adaptation had similar performance to the Quan adaptation in terms of measuring the overall CCI, frequency of individual comorbidities comprising the CCI, and predicting one-year mortality among hospitalized patients. Given the SNOMED adaptation permits for increased reproducibility and transparency of research, we posit the SNOMED adaptation as a substantial improvement to the current implementation of the CCI used by OHDSI.

THURSDAY Adaptation and Validation of the Charlson Comorbidity Index in Administrative Claims Data Using the SNOMED CT Standardized Vocabulary (Stephen Fortin, Jenna Reps, Patrick Ryan)

#OHDSISocialShowcase This Week

FeederNet (Federated E-Health Big Data for Evidence Renovation Network) platform in Korea

PRESENTER: **Seongwon Lee**

INTRO:
- The FeederNet platform was successfully developed, and research using OMOP-CDM and FeederNet platform has been being actively conducted in Korea.
- In this study, we share the status of the FeederNet platform and the future R&D plans on it in Korea.

CDM DATA NETWORK in KOREA
- 57 hospitals established OMOP-CDM
- 72% of all tertiary hospitals in Korea developed OMOP-CDM
- 73,888,154 patients' data (including duplicates) were converted into CDM

57 Hospitals | 73M Patients

FEEDERNET COORDINATING PLATFORM
- Currently, 48 hospitals' CDM DBs are integrated with the platform
- *FeederNet Central*
  - A portal to harmonize distribute joint research
  - Located in cloud
- *FeederNet Node*
  - A client module to execute the analytic engine and communicate with FeederNet central
  - Installed in each hospital's server

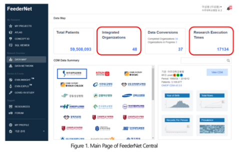


Figure 1. Main Page of FeederNet Central

This work was supported by the Bio Industrial Strategic Technology Development Program (20003883) funded By the Ministry of Trade, Industry & Energy (MOTIE, Korea), a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health & Welfare (grant number: HI16C0992) and Government-wide R&D Fund project for infectious disease research (GFID), Republic of Korea (grant number: HG22C0024)

CDM data-network of 57 hospitals and 73M patients' medical data was established in South Korea.

Currently, joint research using it have been being actively conducted; **97 papers since 2017 and 55 papers in only 2021** were published by Korean first authors.

Scan QR to **download** the abstract or poster.

RESEARCH FREE ZONE (RFZ)
- *Clauses of RFZ*
  1) Grant same authorities on CDM analysis to all researchers in RFZ hospitals
  2) Approve single IRB among RFZ hospitals

18 hospitals joined RFZ

Figure 2. Ranking of CDM DBs most used for analysis on FeederNet in 2021

RESEARCH STATUS in KOREA
- *Analyses on FeederNet*
  - Since May 2019, 17,134 analyses have been executed on FeederNet (see Figure1)
  - Number of analyses on FeederNet has been increasing

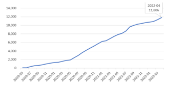
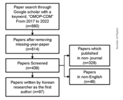


Figure 3. Cumulative number of analysis using FeederNet

- *Paper Publications*
  - 97 papers of Korean first author have been published since 2017 (as of June 2022)
  - 55 papers were published in only 2021



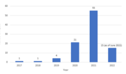


Figure 4. OMOP-CDM research status by year by Korean researcher as the first author

FUTURE R&D
- *(Near) real-time CDM ETL*
  - Periodic full ETL + daily/weekly/monthly incremental ETL
- *Specialized small CDM*
  - Building small but specific CDMs for disease or in-depth medical data
- *Conversion of unstructured medical data to CDM*

Seongwon Lee, Chungsoo Kim, Junhyuk Chang, Rae Woong Park

OHDSI

FRIDAY

FeederNet (Federated E-Health Big Data for Evidence Renovation Network) platform in Korea (Seongwon Lee, Chungsoo Kim, Junhyuk Chang, 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?

OHDSI Speed Dating

CNN US Crime + Justice Energy + Environment Extreme Weather Space + Science

A single winning ticket for Friday's $1.35 billion Mega Millions jackpot drawing was sold in Maine

By Tina Burnside and Aya Elamroussi, CNN
Updated 5:50 AM EST, Sat January 14, 2023

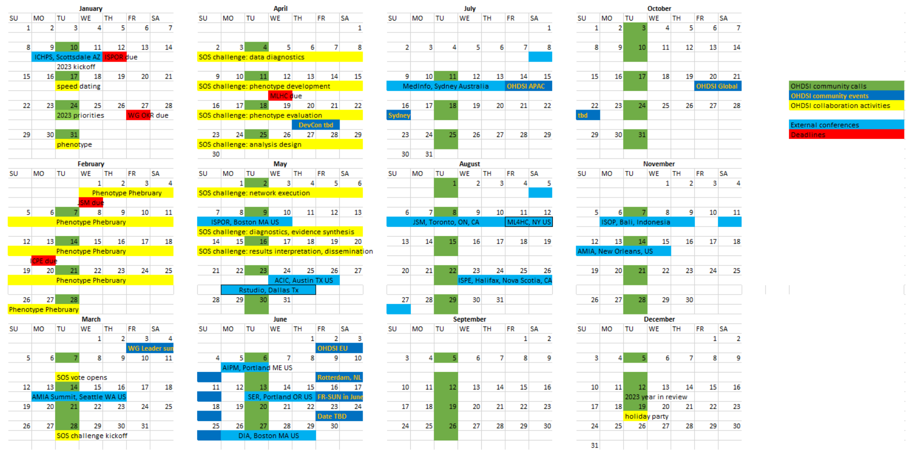
Speed dating questions – Session 1

- What's your name?
- Where do you live?
- What organization(s) are you affiliated with?
- Where are you on the OHDSI journey?
- If Andrew did win the Mega Millions, how much should he invest in OHDSI and what should the money fund?

Speed dating questions – Session 2

- What’s your name?
- Where do you live?
- What organization(s) are you affiliated with?
- Where are you on the OHDSI journey?
- If YOU won the Mega Millions, what would be your first big splurge purchase for yourself?

Speed dating questions – Session 3



- What’s your name?
- Where do you live?
- What organization(s) are you affiliated with?
- Where are you on the OHDSI journey?
- What type of collaboration are you hoping to engage in with fellow OHDSI colleagues this year?

Speed dating questions – Session 4

- What's your name?
- Where do you live?
- What organization(s) are you affiliated with?
- Where are you on the OHDSI journey?
- What was your 2023 New Year's Resolution and (be honest!) have you kept it through 17 days?