
# OHDSI Network Studies

## OHDSI Community Call
## Nov. 16, 2021 • 11 am ET

# Remaining 2021 OHDSI Community Calls

| Date | Topic |
|---|---|
| Nov. 16 | Open Network Studies |
| Nov. 23 | History of OHDSI |
| Nov. 30 | Collaborator Showcase Presentations |
| Dec. 7 | How Did We Do This Year? Final OKR Review |
| Dec. 14 | Holiday-Themed Final Meeting Of 2021 |

# Remaining 2021 OHDSI Community Calls

| Date | Topic |
|---|---|
| Nov. 16 | Open Network Studies |
| Nov. 23 | History of OHDSI |
| Nov. 30 | Collaborator Showcase Presentations |
| Dec. 7 | How Did We Do This Year? Final OKR Review |
| Dec. 14 | Holiday-Themed Final Meeting Of 2021 |

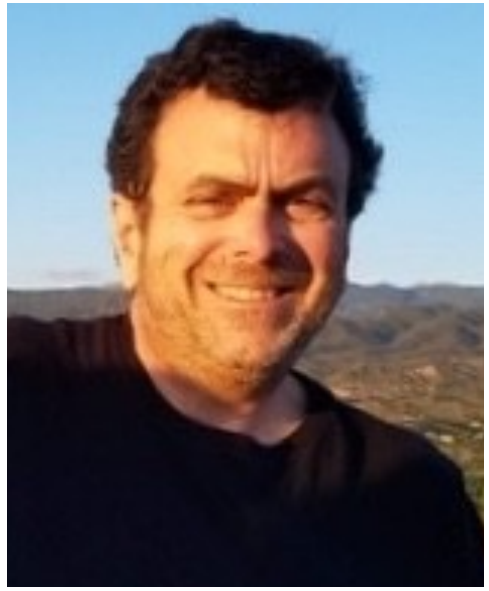
# November 23: The History of OHDSI

**Marc Overhage**

Chief Medical Informatics Officer, Anthem, Inc.

**Judy Racoosin**

FDA Deputy Director of Anesthesia, Analgesia and Addiction Products



**Paul Stang**

Vice-President: Global Epidemiology, Johnson & Johnson

# Three Stages of The Journey

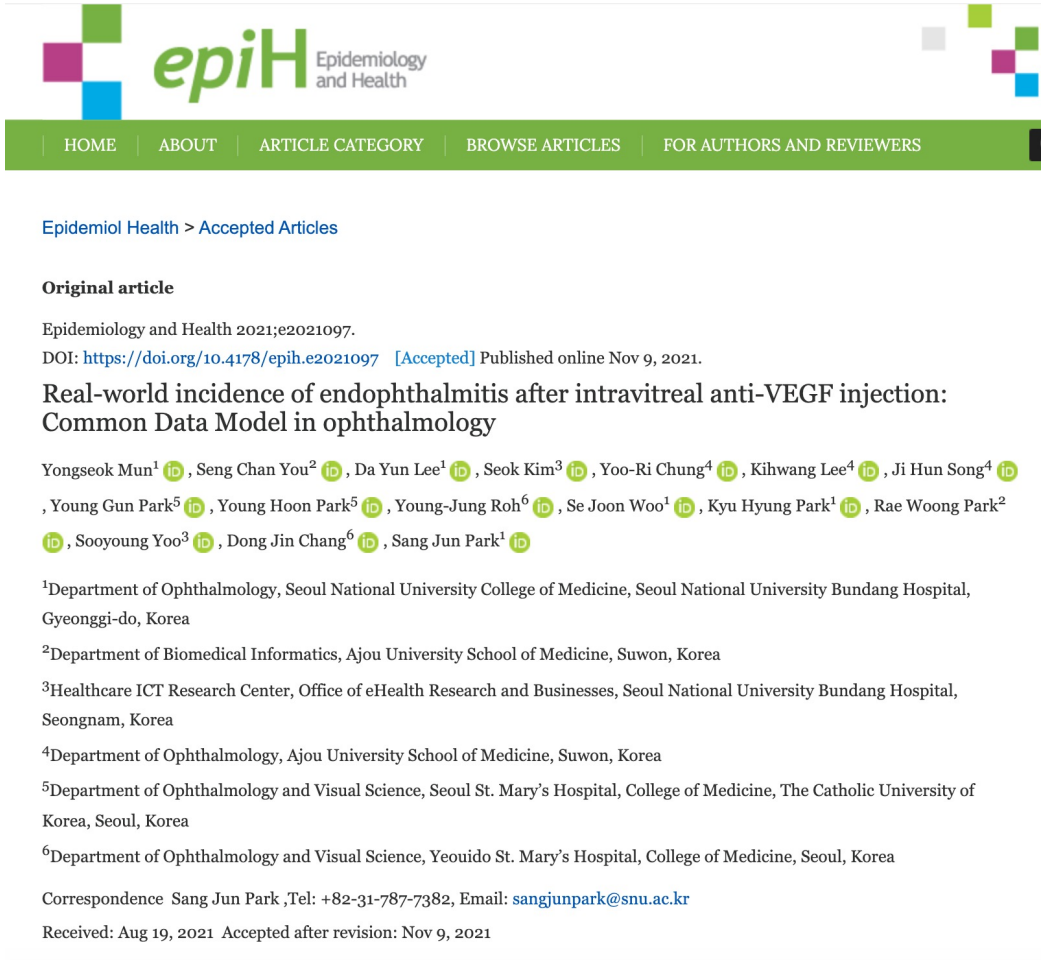
## Where Have We Been?

## Where Are We Now?

## Where Are We Going?

# OHDSI Shoutouts!

Congratulations to **Yongseok Mun Seng Chan You, Da Yun Lee, Seok Kim, Yoo-Ri Chung, Kihwang Lee, Ji Hun Song, Young Gun Park, Young Hoon Park, Young-Jung Roh, Se Joon Woo, Kyu Hyung Park, Rae Woong Park, Sooyoung Yoo, Dong Jin Chang, Sang Jun Park** for the publication of **"Real-world incidence of endophthalmitis after intravitreal anti-VEGF injection: Common Data Model in ophthalmology"** in Epidemiology and Health.



epiH Epidemiology and Health

HOME | ABOUT | ARTICLE CATEGORY | BROWSE ARTICLES | FOR AUTHORS AND REVIEWERS

Epidemiol Health > Accepted Articles

**Original article**

Epidemiology and Health 2021;e2021097.
DOI: https://doi.org/10.4178/epih.e2021097 [Accepted] Published online Nov 9, 2021.

Real-world incidence of endophthalmitis after intravitreal anti-VEGF injection: Common Data Model in ophthalmology

Yongseok Mun[1], Seng Chan You[2], Da Yun Lee[1], Seok Kim[3], Yoo-Ri Chung[4], Kihwang Lee[4], Ji Hun Song[4], Young Gun Park[5], Young Hoon Park[5], Young-Jung Roh[6], Se Joon Woo[1], Kyu Hyung Park[1], Rae Woong Park[2], Sooyoung Yoo[3], Dong Jin Chang[6], Sang Jun Park[1]

[1]Department of Ophthalmology, Seoul National University College of Medicine, Seoul National University Bundang Hospital, Gyeonggi-do, Korea

[2]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Korea

[3]Healthcare ICT Research Center, Office of eHealth Research and Businesses, Seoul National University Bundang Hospital, Seongnam, Korea

[4]Department of Ophthalmology, Ajou University School of Medicine, Suwon, Korea

[5]Department of Ophthalmology and Visual Science, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea

[6]Department of Ophthalmology and Visual Science, Yeouido St. Mary's Hospital, College of Medicine, Seoul, Korea

Correspondence Sang Jun Park ,Tel: +82-31-787-7382, Email: sangjunpark@snu.ac.kr

Received: Aug 19, 2021 Accepted after revision: Nov 9, 2021

# OHDSI Shoutouts!

Both **Jamie Weaver** and **Kristin Kostka** were quoted in a Sunday USA Today feature that highlighted the OHDSI community titled **“Electronic medical records have been around decades. Their power to help other patients is starting to be unleashed.”**

HEALTH

## Electronic medical records have been around decades. Their power to help other patients is starting to be unleashed.

Karen Weintraub USA TODAY

Published 6:01 a.m. ET Nov. 14, 2021

For decades, electronic medical records have offered the potential for making one patient's experience useful to the next.

Such real-world experience could offer insights into which typ on which medication, how best to sequence different drugs or unusual symptoms into a diagnosis.

But unless it was part of a formal research trial, or passed dow the next, that type of information hasn't been available.

"It's really frustrating to imagine we're this far along in our tec still asking really basic questions like what works for whom an said Kristin Kostka, a computational epidemiologist at Northe

James Weaver, shown here with his wife Clare and kids Lucy and Raines, is helping develop computer algorithms that can mine patient health records to help others. *Courtesy James Weaver*

Weaver wanted to figure out whether it would be OK to take just the second drug. Clinical trials focus on people earlier in the course of their disease, not post-surgery and radiation, so they weren't much help.

Instead, he dug into OHDSI databases.

# Three Stages of The Journey

## Where Have We Been?

## Where Are We Now?

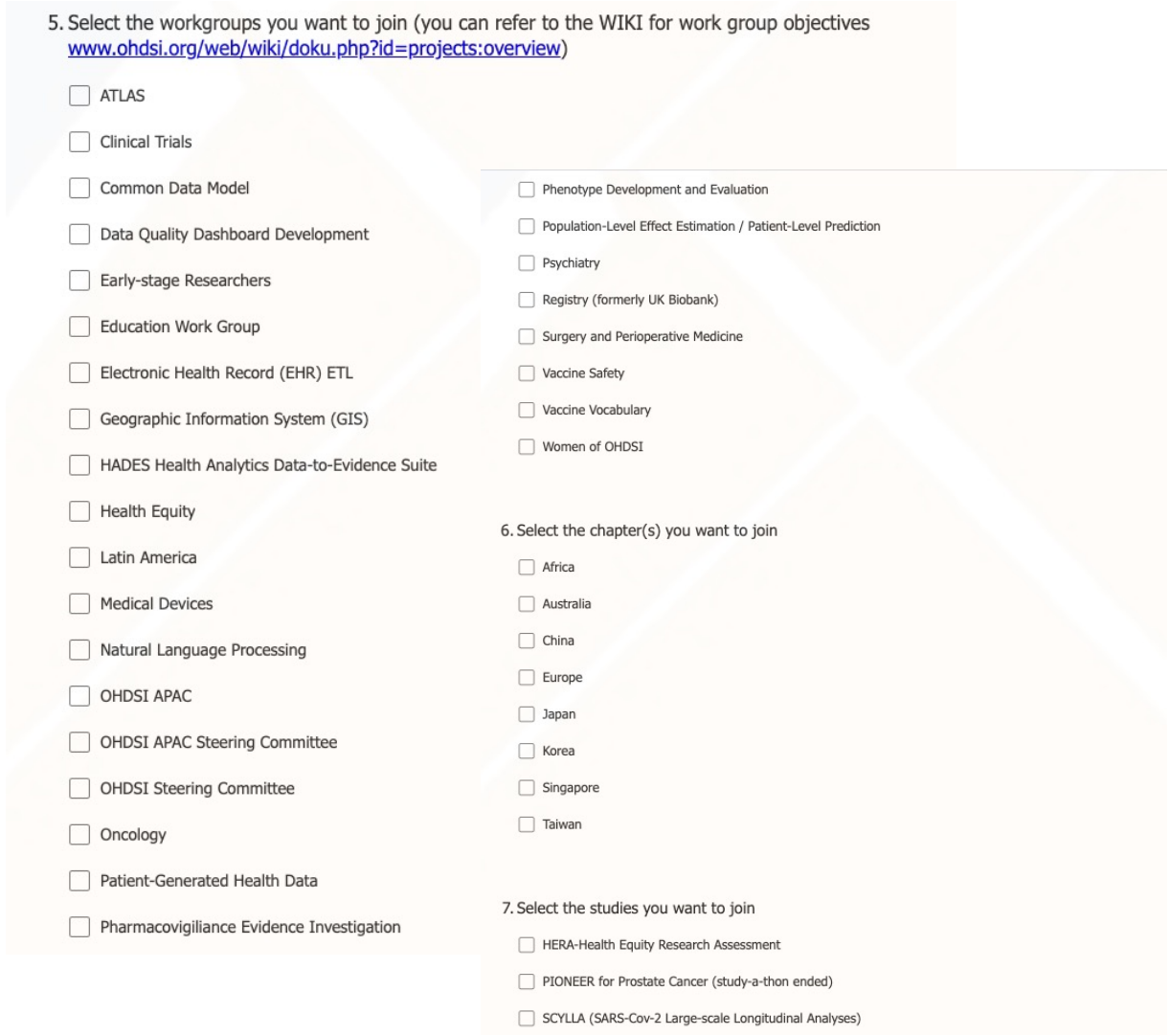
## Where Are We Going?

# Upcoming Workgroup Calls

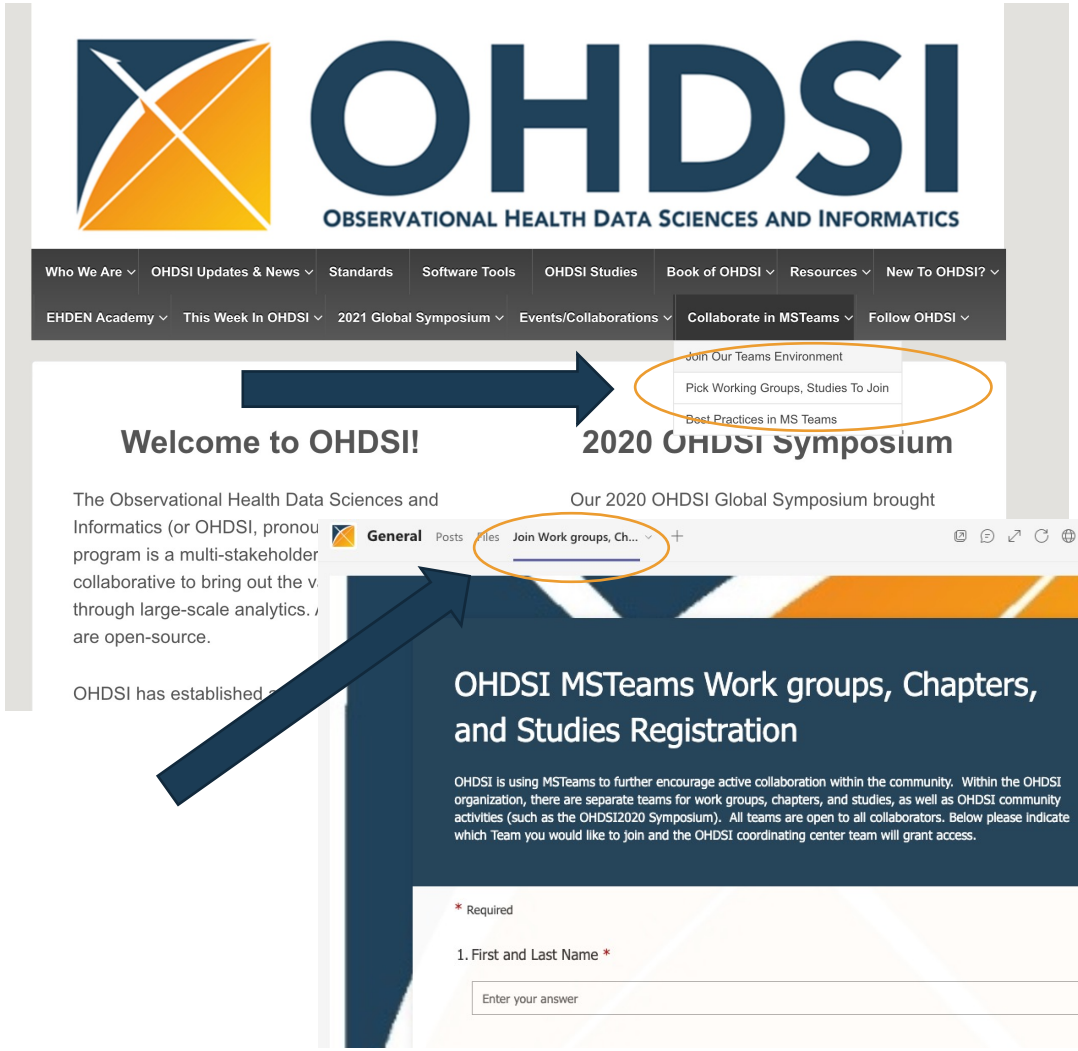
| Date | Time (ET) | Meeting |
|---|---|---|
| Tuesday | 2 pm | Health Equity |
| Wednesday | 10 am | OMOP CDM Oncology – Development Subgroup |
| Wednesday | 1 pm | Data Quality Dashboard Development |
| Wednesday | 7 pm | Medical Imaging |
| Thursday | 12 pm | HADES |
| Thursday | 1 pm | OMOP CDM Oncology – CDM/Vocabulary Subgroup |
| Friday | 9 am | Vaccine Vocabulary |
| Friday | 10 am | Registry |
| Friday | 4 pm | Health Equity Journal Club |
| Monday | 10 am | GIS-Geographic Information System |
| Monday | 2 pm | FHIR and OMOP Terminologies Subgroup (ZOOM) |
| Tuesday | 9 am | OMOP CDM Oncology – Genomic Subgroup |

www.ohdsi.org/upcoming-working-group-calls

# Get Access To Different Teams/WGs/Chapters

# Get Access To Different Teams/WGs/Chapters




5. Select the workgroups you want to join (you can refer to the WIKI for work group objectives www.ohdsi.org/web/wiki/doku.php?id=projects:overview)

- ☐ ATLAS
- ☐ Clinical Trials
- ☐ Common Data Model
- ☐ Data Quality Dashboard Development
- ☐ Early-stage Researchers
- ☐ Education Work Group
- ☐ Electronic Health Record (EHR) ETL
- ☐ Geographic Information System (GIS)
- ☐ HADES Health Analytics Data-to-Evidence Suite
- ☐ Health Equity
- ☐ Latin America
- ☐ Medical Devices
- ☐ Natural Language Processing
- ☐ OHDSI APAC
- ☐ OHDSI APAC Steering Committee
- ☐ OHDSI Steering Committee
- ☐ Oncology
- ☐ Patient-Generated Health Data
- ☐ Pharmacovigiliance Evidence Investigation
- ☐ Phenotype Development and Evaluation
- ☐ Population-Level Effect Estimation / Patient-Level Prediction
- ☐ Psychiatry
- ☐ Registry (formerly UK Biobank)
- ☐ Surgery and Perioperative Medicine
- ☐ Vaccine Safety
- ☐ Vaccine Vocabulary
- ☐ Women of OHDSI

6. Select the chapter(s) you want to join

- ☐ Africa
- ☐ Australia
- ☐ China
- ☐ Europe
- ☐ Japan
- ☐ Korea
- ☐ Singapore
- ☐ Taiwan

7. Select the studies you want to join

- ☐ HERA-Health Equity Research Assessment
- ☐ PIONEER for Prostate Cancer (study-a-thon ended)
- ☐ SCYLLA (SARS-Cov-2 Large-scale Longitudinal Analyses)

# 2021 APAC Symposium • Nov. 18

| Nov. 18 (APAC Time Zone) | Time (Korea time) | Contents | Speaker(s) |
|---|---|---|---|
| Morning | 9:00 – 9:25 am | OHDSI State of the Community | George Hripcsak/Patrick Ryan |
| | 9:25 – 9:50 am | OHDSI APAC State of the Community | Mui Van Zandt |
| | 9:50 – 10:00 am | Energy Break | |
| | 10:00 – 10:25 am | EHDEN | Peter Rijnbeek |
| | 10:25 – 10:50 am | FHIR and OHDSI Collaboration | Christian Reich |
| | 10:50 – 11:00 am | Energy Break | |
| | 11:00 - 12:30 pm | APAC Chapter Visions for 2022 | Chapter Leads |
| Lunch Break | 12:30 – 13:00 pm | | |
| Afternoon (in GatherTown) | 13:00 – 14:00 pm | Workgroup Sessions (Medical Image, FHIR, CDM Tables) | |
| | 14:00 – 15:00 pm | Collaboration Showcase | |
| | 15:00 – 16:00 pm | APAC Study Sessions | |

www.ohdsi.org/apac

# 2021 APAC Symposium • Nov. 18

## WG – Medical Imaging

**Seng Chan You**
Assist. Professor
Yonsei University Health System

## WG – FHIR Collaboration

**Christian Reich**
VP IQVIA, OHDSI founder

**Adam Chee**
Chief of Smart Health Leadership Centre, NUS

## WG – CDM v5.4

**Clair Blacketer**
Assoc. Director Janssen

## APAC Study Session

**Marc Suchard**
Professor, UCLA

**www.ohdsi.org/apac**

# Open-Source Governance Workshop • Nov 29, 9 am

The Open Source Governance Workshop is open to all members of the OHDSI community. The workshop is graciously being organized by the Johns Hopkins Open Source Program Office (OSPO) and Stephen Walli, an open source community advocate and expert at Microsoft.

The goal of the workshop is to learn from very large open source communities on how they onboard, train, and organize contributors to fill a myriad of technical and non-technical roles within an open source eco-system.

# Next CBER Best Seminar Series

BEST

Webinar Registration

| | |
|---|---|
| Topic | CBER BEST Initiative Seminar Series - Vaccine safety evaluation using the self-controlled case series method |
| Description | Background: The CBER BEST Initiative Seminar Series is designed to share and discuss recent research of relevance to ongoing and future surveillance activities of CBER regulated products, namely biologics. The series focuses on safety and effectiveness of biologics including vaccines, blood components, blood-derived products, tissues and advanced therapies. The seminars will provide information on characteristics of biologics, required infrastructure, study designs, and analytic methods utilized for pharmacovigilance and pharmacoepidemiologic studies of biologics. They will also cover information regarding potential data sources, informatics challenges and requirements, utilization of real-world data and evidence, and risk-benefit analysis for biologic products. The length of each session may vary, and the presenters will be invited from outside FDA. Please see the details below for our upcoming seminar. Anyone can register and join for free. Stay tuned for more details and additional webinars during the year.<br><br>Topic: Vaccine safety evaluation using the self-controlled case series method<br><br>Description: The self-controlled case series (SCCS) method will be introduced in the context of vaccine safety evaluation, with examples. We will also consider some of the methodological extensions of the SCCS method that have been used in vaccine safety studies, particularly those relevant to studies of COVID-19 vaccine safety. |
| Time | Dec 1, 2021 11:00 AM in Eastern Time (US and Canada) |

# #OHDSISocialShowcase This Week

## Development of an ETL Process for Bulk and Incremental Load of German Patient Data into OMOP CDM Using FHIR

Elisa Henke[1], Yuan Peng[1], Ines Reinecke[1], Michéle Zoch[1], Martin Sedlmayr[1]
[1] Institute of Medical Informatics and Biometry, Carl Gustav Carus Faculty of Medicine, Technische Universität Dresden, Germany

*This work is part of the project MIRACUM, funded by the German Ministry of Education and Research (FKZ 01ZZ1801A/L).*

### Background

**Motivation:**

- The Use Case 'Alerting in Care – IT Support for Patient Recruitment' [1] in MIRACUM (Medical Informatics in Research and Care in University Medicine) aims to develop a Clinical Trials Recruitment Support System (CTRSS).
- This system suggests patients for clinical trials based on data in the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM).

**Objective:**

- To provide data for the CTRSS we need to design and develop an ETL (Extract-Transform-Load) process for filling OMOP CDM using Fast Healthcare Interoperability Resources (FHIR) profiles from MI-I and MIRACUM as data source.
- The ETL process has to support an initial (bulk) load as well as near real time or at least once a day updates (incremental load) of the data in OMOP CDM, to enable quick recruitment.

### Methods

**Semantic Mapping**

**FHIR Profiles**
- Patient
- Encounter
- Condition

[2]

- Medication
- MedicationAdministration
- Observation
- Procedure

[3]

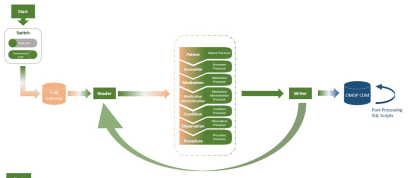
**OMOP CDM tables** [4]
- PERSON
- LOCATION
- DEATH
- VISIT_OCCURRENCE
- VISIT_DETAIL
- OBSERVATION_PERIOD
- CONDITION_OCCURRENCE
- FACT_RELATIONSHIP
- PROCEDURE_OCCURRENCE
- OBSERVATION
- MEASUREMENT
- DRUG_EXPOSURE

**SpringBatch Framework** [5]


### Results

We have designed and implemented an ETL process which transforms MI-I and MIRACUM FHIR resources to OMOP CDM. This ETL process uses a switch to select whether the ETL process is executed as bulk load or as incremental load.

bulk load
with predefined order
of processing
FHIR resources



incremental load
with random order
of processing
FHIR resources

### Conclusions

- The developed ETL process can transform and load data from FHIR into OMOP CDM as bulk load or incremental load.
- Thus, patient data can be updated to enable rapid recruitment with the CTRSS based on OMOP CDM.
- In the future, it is our aim to:
  - use meta data from FHIR and OMOP CDM to automate the ETL process
  - update the ETL process to new versions of the FHIR profiles from MI-I

References:
[1] Reinecke I, Gulden C, Kümmel M, Nassirian A, Blasini R, Sedlmayr M. Design for a Modular Clinical Trial Recruitment Support System Based on FHIR and OMOP. Stud Health Technol Inform. 2020 16;270:158-162.
[2] Medical Informatics Initiative Germany. Basismodule des Kerndatensatzes der MII. Available from: https://www.medizininformatik-initiative.de/de/basismodule-deskerndatensatzes-der-mii.
[3] MIRACUM. MIRACUM Core Implementation Guide – Table of Contents. 2020. Available from: https://fhir.miracum.org/core/toc.html.
[4] Observational Health Data Sciences and Informatics. OMOP CDM v5.3.1. Available from: https://ohdsi.github.io/CommonDataModel/cdm531.html.
[5] Ward L, Syer D, Risberg T, Kasanicky R, Garrette D, Lund W, Minella M, Schaefer C, Hillert G, Renfro G, Bryant J, Hassine M B. Spring Batch – Reference Documentation. 2021. Available from: https://docs.spring.io/spring-batch/docs/current/reference/html/index.html.

Contact: Elisa.Henke@ukdd.de

## MONDAY

## The secret (outpatient) garden: Using the OMOP CDM to better differentiate between out-patient and inpatient diagnoses in the VA EHR data

**Authors: Elisa Henke, Yuan Peng, Ines Reinecke, Michele Zoch, Martin Sedlmayr**

# #OHDSISocialShowcase This Week

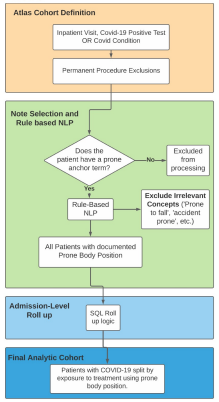
Lightning Talk!

**Detection of prone positioning in hospitalized COVID patients using NLP**

PRESENTER:
**Patrick R. Alba**

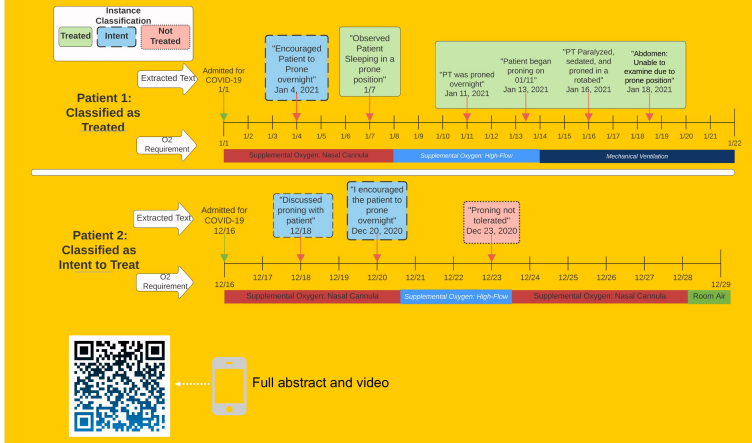
**BACKGROUND:**
**A**wake prone positioning has been widely adopted in many centers during the COVID-19 pandemic, and was conditionally endorsed by organizations such as the World Health Organization and the Intensive Care Society. However, estimating the use of proning as a treatment is not possible with current structured data. We developed a natural language processing(NLP) tool to identify documentation of proning.

**METHODS**




**RESULTS**
- 5,212 patients identified at the Department of Veterans Affairs (VA) from March 2020 – May 2021
- A preliminary external validation was completed using data from the Stanford Health system data warehouse. (STARR)

Awake prone positioning of COVID-19 patients is **novel and experimental, yet broadly adopted.** We used NLP to estimate its use in real world pandemic data.



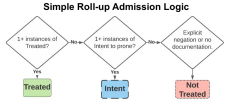

Full abstract and video

**Concept Definition:**

Clinical notes for COVID-19 patients were often found to contain recommendations regarding prone positioning in awake patients (e.g. notes indicating that proning was encouraged, or that education was provided). This reflects an 'intention to prone', regardless of the patient's subsequent exposure to prone positioning. To permit this flexibility in subsequent analysis, we decided to roll up all mentions of proning into three primary categories.

**Instance Classification**
- **Treated:** positive mention indicating the patient was exposed to prone positioning.
- **Intent to treat:** proning appears to have been encouraged, with no clear negation or indication as to its tolerability.
- **Not treated:** typically reflects an inability to prone, or intention to prone that is negated.

**Simple Roll-up Admission Logic**




**NLP Methods:**

The NLP system was built using the Apache Unstructured Information Management Architecture Asynchronous Scaleout(UIMA-AS) and the libraries and tools contained in the VA Leo framework.

**Validation**

At an instance-level using VA data precision and inter-annotator agreement(IAA) were assessed showing a **precision of 94.6%** and an IAA **Kappa of 0.73**. After applying the admission-level rollup logic, 100 complete patient timelines were reviewed using VA data, and a preliminary validation of 40 patients was completed using Stanford STARR data.

| | Sensitivity | Specificity | PPV |
|---|---|---|---|
| VA | 84.09 | 94.64 | 91.28 |
| STARR* | 100.00* | 50.00 | 85.71 |

* While the results were promising, this initial sample and validation was completed primarily to assess the feasibility and application of the algorithm on Stanford Data.

**Patients:**

5,212 patients identified at the VA as having received treatment from March 2020 – May 2021.

**Future work:**

The creation of this algorithm is an important first step towards understanding the incidence and utility of awake proning as a treatment of hypoxemic respiratory failure. Future work will focus on completion of multi-institutional proning incidence study through large scale deployment within the Stanford STARR cohort.


**Patrick R. Alba, Jose Posada, Jason Weatherald, Kristine E. Lynch, Annie Bowles, Nigam Shah, Olga V. Patterson, Scott L. DuVall*, Evan Minty***

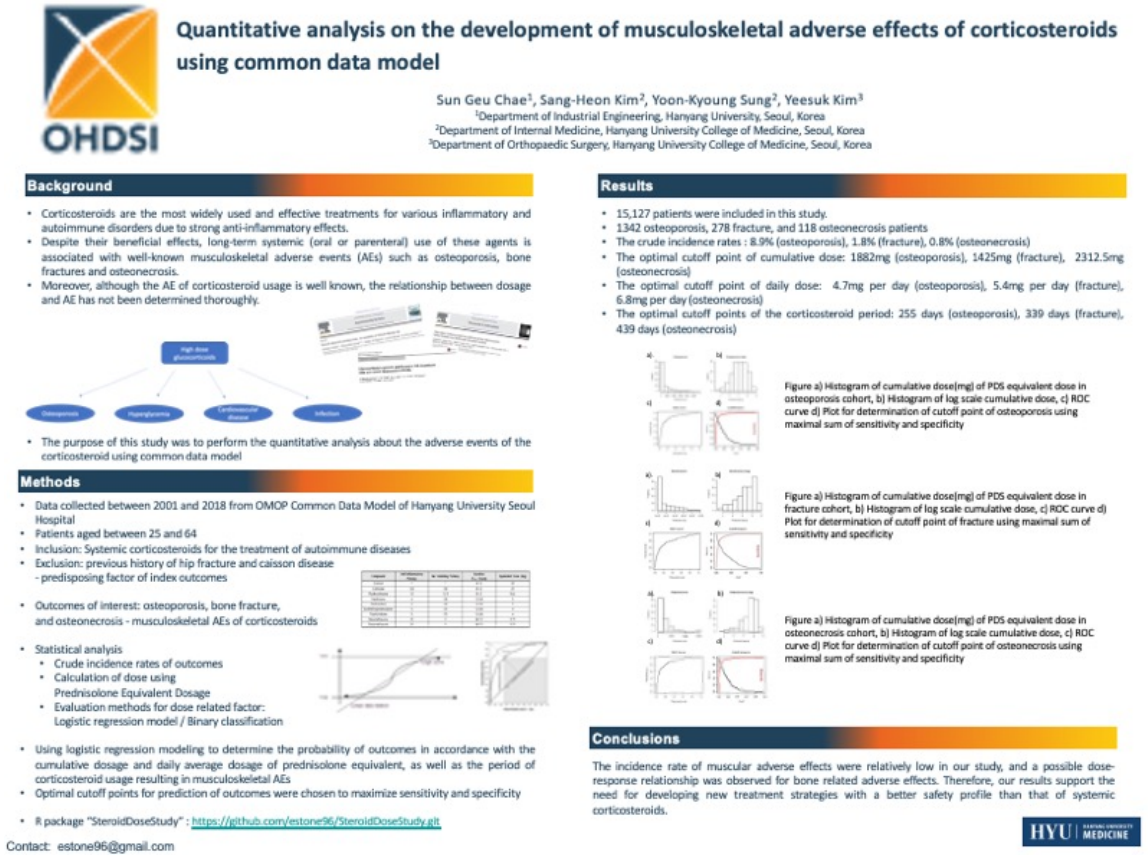
OHDSI


**TUESDAY**

## Detection of prone positioning in hospitalized COVID patients using NLP


Authors: Patrick R. Alba (presenter), Jose Posada, Jason Weatherald, Kristine Lynch, Annie Bowles, Nigam Shah, Olga V. Patterson, Scott L. DuVall, Evan Minty

# #OHDSISocialShowcase This Week



Quantitative analysis on the development of musculoskeletal adverse effects of corticosteroids using common data model

Sun Geu Chae[1], Sang-Heon Kim[2], Yoon-Kyoung Sung[2], Yeesuk Kim[3]
[1]Department of Industrial Engineering, Hanyang University, Seoul, Korea
[2]Department of Internal Medicine, Hanyang University College of Medicine, Seoul, Korea
[3]Department of Orthopaedic Surgery, Hanyang University College of Medicine, Seoul, Korea



**WEDNESDAY**

## Quantitative analysis on the development of musculoskeletal adverse effects of corticosteroids using common data model

**Authors: Sun Geu Chae, Sang-Heon Kim, Yoon-Kyoung Sung, Yeesuk Kim**

# #OHDSISocialShowcase This Week

***Title: PheValuator 2.0: Changes to Improve the Performance of the Phenotype Algorithm Evaluation Tool***

PRESENTER: **Joel** Swerdel

**INTRO:**

- **Phenotype algorithms are used in nearly all observational research**
- **The performance of these algorithms is usually unknown**
- **Traditional validation through chart review is expensive, time consuming, and provides incomplete results**
- **PheValuator is an open-source tool in the OHDSI toolkit for estimating all the performance characteristics of algorithms**
- **Objective: Evaluate performance of PheValuator 2.0 v. 1.0 as compared to gold standard phenotype benchmark**

**METHODS**

**Changes to PheValuator 2.0 include:**

1. **New xSpec and xSens**

Example Psoriasis xSpec:

≥2 Diagnosis codes for Psoriasis -30 to -1 days prior to any visit
**AND** 1 Diagnosis code for Psoriasis -30 to -1 days prior to any visit for the first time in a subject's history

Example Psoriasis xSens:

≥ 1 Diagnosis code for Psoriasis All time to -1 days prior to any visit

2. **Use of up to 3 time windows for predictive modeling features**

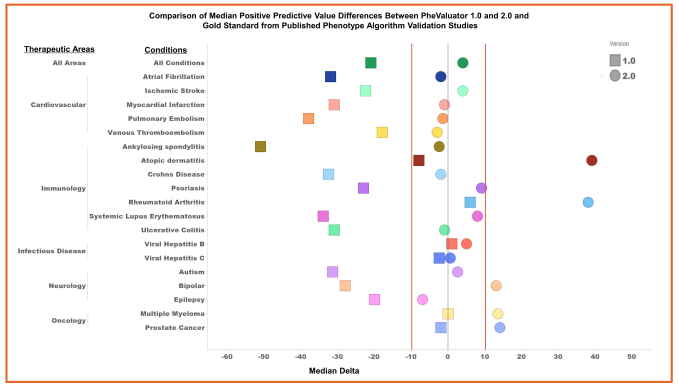
Example for chronic disease:

Window 1: 0-30 days after index
Window 2: 31-365 days after index
Window 3: 366-9999 days after index

**Analysis:** Calculated differences in Positive Predictive Value, Sensitivity, and Specificity from PheValuator 1.0 + 2.0 and the phenotype benchmark set of traditionally validated algorithms

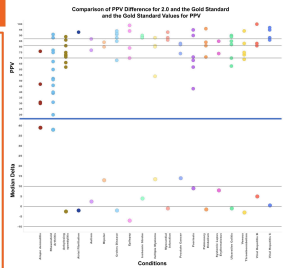
- 19 phenotypes
- 90 algorithms
- Across 5 datasets

**PheValuator 2.0 produces estimates of Positive Predictive Value for phenotype algorithms comparable to chart review.**




| Therapeutic Area | Condition | Positive Predictive Value Difference Version 1 Median (IQR) | Version 2 Median (IQR) |
|---|---|---|---|
| Overall | Overall | -2 (-34, -3) | 4 (-3, 15) |
| Cardiovascular | Overall | -25 (-34, -18) | 0 (-7, 4) |
| | Atrial Fibrillation | -32 (-39, -28) | -2 (-3, 1) |
| | Pulmonary Embolism | -38 (-46, -24.25) | -1.5 (-8.25, 3.25) |
| | Venous Thromboembolism | -18 (-25, -5) | -3 (-10, 1) |
| | Ischemic Stroke | -22.5 (-28, -19.75) | 4 (2, 5) |
| | Myocardial Infarction | -31 (-35, -20) | -1 (-6.5, 0.5) |
| Immunology | Overall | -27 (-39, -4) | 7 (-2, 30.25) |
| | Ankylosing spondylitis | -51 (-59, -41.5) | -2.5 (-10, 5) |
| | Atopic dermatitis | -8 (-16.5, -2) | 39 (26.5, 42.5) |
| | Ulcerative Colitis | -31 (-37.5, -21.25) | -1 (-5.75, 4) |
| | Crohns Disease | -32.5 (-39, -26.25) | -2 (-6, 2.75) |
| | Rheumatoid Arthritis | 6 (-11, 16) | 38 (23, 48) |
| | Psoriasis | -23 (-34, -6.5) | 9 (-3, 26) |
| | Systemic Lupus Erythematosus | -34 (-42.5, -30) | 8 (0, 9) |
| Infectious Disease | Overall | -1 (-12.5, 3) | 2 (-3, 6) |
| | Viral Hepatitis C | -2.5 (-11.5, 3) | 0.5 (-3, 3.75) |
| | Viral Hepatitis B | 1 (-14.5, 1) | 5 (-5.5, 6) |
| Neurology | Overall | -24 (-31.75, -17) | 7 (-8, 12.75) |
| | Autism | -31.5 (-36, -26.5) | 2.5 (-5.25, 7.5) |
| | Bipolar | -28 (-31.5, -26.25) | 13 (11.5, 14.75) |
| | Epilepsy | -20 (-24, -10.75) | -7 (-18.75, 9.5) |
| Oncology | Overall | -1 (-4, 4.5) | 14 (10, 18) |
| | Multiple Myeloma | 0 (-3.25, 6.5) | 13.5 (10, 22) |
| | Prostate Cancer | -2 (-4, 1.25) | 14 (9.5, 16.5) |

| Therapeutic Area | Condition | Sensitivity Difference Version 1.0 Median (IQR) | Version 2.0 Median (IQR) |
|---|---|---|---|
| Overall | Overall | -39 (-51, -20) | -16 (-37, -6) |
| Cardiovascular | Pulmonary Embolism | -23 (-31.75, -12.75) | -20 (-28.75, -15.25) |
| | Venous Thromboembolism | -55.5 (-56, -47) | -42.5 (-44, -40) |
| Immunology | Rheumatoid Arthritis | -45.5 (-63.25, -34) | -23.5 (-39, -12.25) |
| | Psoriasis | -20 (-28.5, -16) | 4 (-1, 6) |
| Neurology | Epilepsy | -41 (-41, -36) | -12 (-18, -11) |
| Oncology | Multiple Myeloma | -37 (-48, -21.75) | -14.5 (-21, -3.25) |
| | Prostate Cancer | -85 (-87, -85) | -58 (-62, -44) |





Joel N. Swerdel, PhD MS MPH[1,2];
Martijn Schuemie, PhD[2,3];
Patrick B. Ryan, PhD[1,2,3]
[1] Janssen Research and Development, Titusville, NJ, USA
[2] Columbia University, New York, NY, USA
[3] Observational Health Data Sciences and Informatics (OHDSI), New York, NY


## THURSDAY

## PheValuator 2.0: Changes to Improve the Performance of the Phenotype Algorithm Evaluation Tool


**Authors: Joel Swerdel, Martijn Schuemie, Patrick B. Ryan**

# #OHDSISocialShowcase This Week

***Ontology of Cancer Diagnosis in the OMOP Vocabulary***

PRESENTER: **Dmitry Dymshyts**

**INTRODUCTION:**

- Observational research in cancer requires substantially more detail to represent conditions than most other therapeutic areas.
- Cancer attributes are covered in many terminology systems and data collection standards including ICD-O, SNOMED CT, LOINC, AJCC, NCIt, NAACCR, CAP
- For the first version of the OMOP CDM Oncology Module a new vocabulary "Cancer Modifiers" was developed based on the content of LOINC, NCIt, NAACCR, and CAP.
- In the new version of the OMOP Oncology Module, we addressed the problem of missing ontological relationships and mappings between the source and the Cancer Modifier vocabulary.

**Methods**

Each cancer modifier concept is a result of pre-coordination of two or more attributes. For example, *Prostate Cancer by AJCC 7th edition Stage 4* is pre-coordinated from the dimensions of *Stage*, *Staging System*, and diagnostic *Schema* (Table 1). Concepts in the Cancer Modifier vocabulary are grouped into concept classes each covering certain diagnostic aspects (e.g. *Staging/Grading*, *Metastasis*). Each class of cancer modifiers has a set of defined dimensions. Cancer modifiers are recorded in the measurement table and can be explicitly connected with the base cancer diagnosis through modifier_of_event_id and modifier_of_field_concept_id to represent a comprehensive patient's cancer diagnosis. †

We extended the Cancer Modifier vocabulary with relationships to their respective dimensions. These dimensions serve as nodes in the concept hierarchy. A complete set of dimensions for each cancer modifier class is depicted in Table 1.

Adding relationships between base diagnosis and Staging/Grading modifiers via diagnostic schemas

Adding modifier attributes enable aggregation and hierarchical queries.

Building mappings from NAACCR and CAP vocabularies to Cancer Modifier vocabulary

Note, this work is in progress, please join the Oncology Working group if you want to contribute

Take a picture to go to the repository or click this message

Table 1. Cancer modifier concepts and their attributes

| Staging and Grading | Stage | Stage IV | Prostate Cancer by AJCC 7th edition Stage 4 |
|---|---|---|---|
| | Staging system | AJCC v7 | |
| | Schema | Prostate Cancer | |
| Metastasis | Fact of metastasis | Metastasis to | Metastasis to the Body of Stomach |
| | Metastatic site | Body of Stomach | |
| Nodes | Type of involvement | Macrometastasis | Paraaortic Lymph Nodes with Macrometastasis |
| | Group of lymph nodes involved | Paraaortic Lymph Nodes | |
| Dimension | Measurement type | Greatest Dimension | Greatest Dimension of viable tumor |
| | Measured entity | Viable tumor | |
| Margin | Margin type | Distal Margin | Distal Margin Involved by High-Grade squamous dysplasia |
| | Tumor growth | High-Grade squamous dysplasia | |
| Extension/Invasion | Fact of invasion of extension | Microscopic invasion | Microscopic Lymphovascular Invasion (LVI) into the Lymphatics |
| | Target structure of invasion | Lymphatics | |
| Histopattern | Histological finding | Nodular Growth Pattern | Nodular Growth Pattern |
| Topography | Tumor location | 10 o'clock | 10 o'clock |

Results

- Added relationships between base diagnosis and Staging/Grading modifiers via diagnostic schemas enable post-coordination of base diagnosis with stage and grade when these links are not available in the source data. This enhancement is critical for precise definition of cancer diagnosis and supports high specificity in cohort and phenotype building.

- Added modifier attributes enable aggregation and hierarchical queries.

- Mappings from NAACCR and CAP vocabularies to Cancer Modifier vocabulary serve two purposes. One as a crosswalk from these representations to OMOP. Another as a deduplication of highly redundant concepts in NAACCR and CAP. These mappings enable vocabulary-driven data conversion to OMOP from the two major sources of US cancer data, tumor registries and synoptic pathology reports.

Dmitry Dymshyts, Vlad Korsik, Denys Kaduk, Mike Nerovnya, Alex Davydov, Christian Reich, Michael Gurley, Asieh Golozar, Shilpa Ratwani, S. Joseph Sirintrapun, W. Scott Campbell, Rimma Belenkaya - OHDSI Oncology Workgroup

REGENERON UNMC ODYSSEUS DATA SERVICES INC Memorial Sloan Kettering Cancer Center IQVIA OHDSI

**FRIDAY**

## Ontology of Cancer Diagnosis in the OMOP Vocabulary

**Authors: Dmitry Dymshyts, Vlad Korsik, Denys Kaduk, Mike Nerovnya, Alex Davydov, Christian Reich, Michael Gurley, Asieh Golozar, Shilpa Ratwani, Joseph Sirintrapun , W. Scott Campbell, Rimma Belenkaya**

# 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?

# November 16 Community Call: OHDSI Studies

**Prognostic Significance of Liver Metastasis in Non-Small Cell Lung Cancer**

Asieh Golozar

**Assessing Health Equity in Mental Healthcare Delivery Using a Federated Network Research Model**

Jacob Zelko

**Redefining Polypharmacy: A Longitudinal Study in Routinely Collected Data**

Leena Elhussein

**Long COVID phenotyping and vaccine effectiveness methods**

Annika Jodicke and Kristin Kostka

**Health Equity Research Assessment (HERA) Characterization**

Noémie Elhadad

**Adverse Events of Special Interest within COVID-19 Subjects**

Erica Voss