Reproducibility

OHDSI Community Call
March 29, 2022 • 11 am ET

April OHDSI Community Calls

Date	Topic
April 5	Name That Result
April 12	OHDSI Coordinating Center
April 19	DARWIN EU
April 26	Open-Source Community

Future OHDSI Community Calls

Date	Topic
April 5	Name That Result
April 12	OHDSI Coordinating Center
April 19	DARWIN EU
April 26	Open-Source Community

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **Kristin Kostka, Talita Duarte-Salles, Albert Prats-Uribe, Anthony G Sena, Andrea Pistillo, Sara Khalid, Lana YH Lai, Asieh Golozar, Thamir M Alshammari, Dalia M Dawoud, Fredrik Nyberg, Adam B Wilcox, Alan Andryc, Andrew Williams, Anna Ostropolets, Carlos Areia, Chi Young Jung, Christopher A Harle, Christian G Reich, Clair Blacketer, Daniel R Morales, David A Dorr, Edward Burn, Elena Roel, Eng Hooi Tan, Evan Minty, Frank DeFalco, Gabriel de Maeztu, Gigi Lipori, Hiba Alghoul, Hong Zhu, Jason A Thomas, Jiang Bian, Jimyung Park, Jordi Martínez Roldán, Jose D Posada, Juan M Banda, Juan P Horcajada, Julianna Kohler, Karishma Shah, Karthik Natarajan, Kristine E Lynch, Li Liu, Lisa M Schilling, Martina Recalde, Matthew Spotnitz, Mengchun Gong, Michael E Matheny, Neus Valveny, Nicole G Weiskopf, Nigam Shah, Osaid Alser, Paula Casajust, Rae Woong Park, Robert Schuff, Sarah Seager, Scott L DuVall, Seng Chan You, Seokyoung Song, Sergio Fernández-Bertolín, Stephen Fortin, Tanja Magoc, Thomas Falconer, Vignesh Subbian, Vojtech Huser, Waheed-Ul-Rahman Ahmed, William Carter, Yin Guan, Yankuic Galvan, Xing He, Peter R Rijnbeek, George Hripcsak, Patrick B Ryan, Marc A Suchard, and Daniel Prieto-Alhambra** the publication of **"Unraveling COVID-19: A Large-Scale Characterization of 4.5 Million COVID-19 Cases Using CHARYBDIS"** in Clinical Epidemiology.

Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

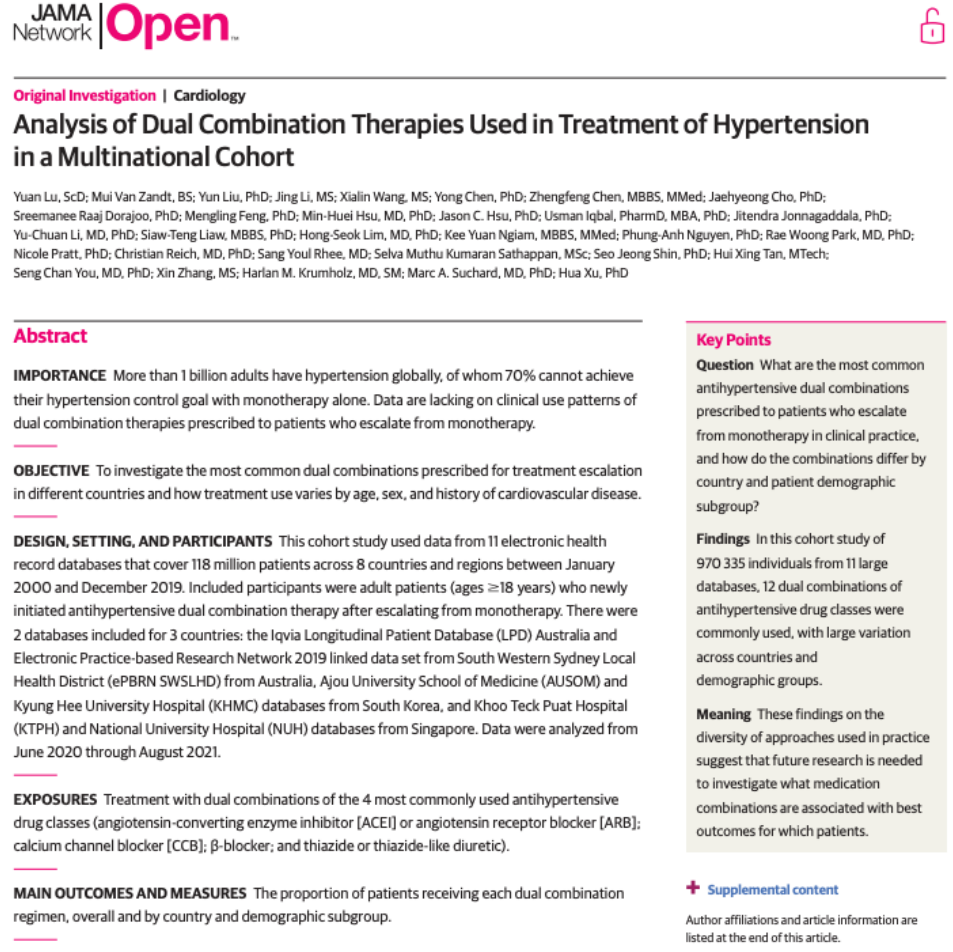
ORIGINAL RESEARCH

Unraveling COVID-19: A Large-Scale Characterization of 4.5 Million COVID-19 Cases Using CHARYBDIS

Kristin Kostka[1,2], Talita Duarte-Salles[3], Albert Prats-Uribe[4], Anthony G Sena[5,6], Andrea Pistillo[3], Sara Khalid[4], Lana YH Lai[7], Asieh Golozar[8,9], Thamir M Alshammari[10], Dalia M Dawoud[11], Fredrik Nyberg[12], Adam B Wilcox[13,14], Alan Andryc[5], Andrew Williams[15], Anna Ostropolets[16], Carlos Areia[17], Chi Young Jung[18], Christopher A Harle[19], Christian G Reich[1,2], Clair Blacketer[5,6], Daniel R Morales[20], David A Dorr[21], Edward Burn[3,4], Elena Roel[3,22], Eng Hooi Tan[4], Evan Minty[23], Frank DeFalco[5], Gabriel de Maeztu[24], Gigi Lipori[19], Hiba Alghoul[25], Hong Zhu[26], Jason A Thomas[13], Jiang Bian[19], Jimyung Park[27], Jordi Martínez Roldán[28], Jose D Posada[29], Juan M Banda[30], Juan P Horcajada[31], Julianna Kohler[32], Karishma Shah[33], Karthik Natarajan[16,34], Kristine E Lynch[35,36], Li Liu[37], Lisa M Schilling[38], Martina Recalde[3,22], Matthew Spotnitz[14], Mengchun Gong[39], Michael E Matheny[40,41], Neus Valveny[42], Nicole G Weiskopf[21], Nigam Shah[29], Osaid Alser[43], Paula Casajust[42], Rae Woong Park[27,44], Robert Schuff[21], Sarah Seager[1], Scott L DuVall[35,36], Seng Chan You[45], Seokyoung Song[46], Sergio Fernández-Bertolín[3], Stephen Fortin[5], Tanja Magoc[19], Thomas Falconer[16], Vignesh Subbian[47], Vojtech Huser[48], Waheed-Ul-Rahman Ahmed[33,49], William Carter[38], Yin Guan[50], Yankuic Galvan[19], Xing He[19], Peter R Rijnbeek[6], George Hripcsak[16,34], Patrick B Ryan[5,16], Marc A Suchard[51], Daniel Prieto-Alhambra[4]

OHDSI Shoutouts!

Congratulations to the team of **Yuan Lu, Mui Van Zandt, Yun Liu, Jing Li, Xialin Wang, Yong Chen, Zhengfeng Chen, Jaehyeong Cho, Sreemanee Raaj Dorajoo, Mengling Feng, Min-Huei Hsu, Jason C. Hsu, Usman Iqbal, Jitendra Jonnagaddala, Yu-Chuan Li, Siaw-Teng Liaw, Hong-Seok Lim, Kee Yuan Ngiam, Phung-Anh Nguyen, Rae Woong Park, Nicole Pratt, Christian Reich, Sang Youl Rhee, Selva Muthu Kumaran Sathappan, Seo Jeong Shin, Hui Xing Tan, Seng Chan You, Xin Zhang, Harlan M. Krumholz, Marc A. Suchard, and Hua Xu** on the publication of **Analysis of Dual Combination Therapies Used in Treatment of Hypertension in a Multinational Cohort** in JAMA Network Open.



JAMA Network Open

Original Investigation | Cardiology

Analysis of Dual Combination Therapies Used in Treatment of Hypertension in a Multinational Cohort

Yuan Lu, ScD; Mui Van Zandt, BS; Yun Liu, PhD; Jing Li, MS; Xialin Wang, MS; Yong Chen, PhD; Zhengfeng Chen, MBBS, MMed; Jaehyeong Cho, PhD; Sreemanee Raaj Dorajoo, PhD; Mengling Feng, PhD; Min-Huei Hsu, MD, PhD; Jason C. Hsu, PhD; Usman Iqbal, PharmD, MBA, PhD; Jitendra Jonnagaddala, PhD; Yu-Chuan Li, MD, PhD; Siaw-Teng Liaw, MBBS, PhD; Hong-Seok Lim, MD, PhD; Kee Yuan Ngiam, MBBS, MMed; Phung-Anh Nguyen, PhD; Rae Woong Park, MD, PhD; Nicole Pratt, PhD; Christian Reich, MD, PhD; Sang Youl Rhee, MD; Selva Muthu Kumaran Sathappan, MSc; Seo Jeong Shin, PhD; Hui Xing Tan, MTech; Seng Chan You, MD, PhD; Xin Zhang, MS; Harlan M. Krumholz, MD, SM; Marc A. Suchard, MD, PhD; Hua Xu, PhD

Abstract

IMPORTANCE More than 1 billion adults have hypertension globally, of whom 70% cannot achieve their hypertension control goal with monotherapy alone. Data are lacking on clinical use patterns of dual combination therapies prescribed to patients who escalate from monotherapy.

OBJECTIVE To investigate the most common dual combinations prescribed for treatment escalation in different countries and how treatment use varies by age, sex, and history of cardiovascular disease.

DESIGN, SETTING, AND PARTICIPANTS This cohort study used data from 11 electronic health record databases that cover 118 million patients across 8 countries and regions between January 2000 and December 2019. Included participants were adult patients (ages ≥18 years) who newly initiated antihypertensive dual combination therapy after escalating from monotherapy. There were 2 databases included for 3 countries: the Iqvia Longitudinal Patient Database (LPD) Australia and Electronic Practice-based Research Network 2019 linked data set from South Western Sydney Local Health District (ePBRN SWSLHD) from Australia, Ajou University School of Medicine (AUSOM) and Kyung Hee University Hospital (KHMC) databases from South Korea, and Khoo Teck Puat Hospital (KTPH) and National University Hospital (NUH) databases from Singapore. Data were analyzed from June 2020 through August 2021.

EXPOSURES Treatment with dual combinations of the 4 most commonly used antihypertensive drug classes (angiotensin-converting enzyme inhibitor [ACEI] or angiotensin receptor blocker [ARB]; calcium channel blocker [CCB]; β-blocker; and thiazide or thiazide-like diuretic).

MAIN OUTCOMES AND MEASURES The proportion of patients receiving each dual combination regimen, overall and by country and demographic subgroup.

Key Points

Question What are the most common antihypertensive dual combinations prescribed to patients who escalate from monotherapy in clinical practice, and how do the combinations differ by country and patient demographic subgroup?

Findings In this cohort study of 970 335 individuals from 11 large databases, 12 dual combinations of antihypertensive drug classes were commonly used, with large variation across countries and demographic groups.

Meaning These findings on the diversity of approaches used in practice suggest that future research is needed to investigate what medication combinations are associated with best outcomes for which patients.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

OHDSI Shoutouts!

Congratulations to both **Aki Nishimura and Marc Suchard** on the publication of **Prior-preconditioned conjugate gradient method for accelerated Gibbs sampling in 'large n & large p' Bayesian sparse regression** in the Journal of the American Statistical Association.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2022, VOL. JUST-ACCEPTED, NO. JA, 1-40
https://doi.org/10.1080/01621459.2022.2057859

Taylor & Francis
Taylor & Francis Group

Open access

Prior-preconditioned conjugate gradient method for accelerated Gibbs sampling in 'large *n* & large *p*' Bayesian sparse regression

Akihiko Nishimura[a] and Marc A. Suchard[b]

[a] Department of Biostatistics, Johns Hopkins University; [b] Department of Biomathematics, Biostatistics, and Human Genetics, University of California - Los Angeles

ABSTRACT

In a modern observational study based on healthcare databases, the number of observations and of predictors typically range in the order of $10^6 \sim 10^9$ and of $10^4 \sim 10^5$. Despite the large sample size, data rarely provide sufficient information to reliably estimate such a large number of parameters. Sparse regression techniques provide potential solutions, one notable approach being the Bayesian method based on shrinkage priors. In the "large n & large p" setting, however, the required posterior computation encounters a bottleneck at repeated sampling from a high-dimensional Gaussian distribution, whose precision matrix $\mathbf{\Phi}$ is expensive to compute and factorize. In this article, we present a novel algorithm to speed up this bottleneck based on the following observation: we can cheaply generate a random vector b such that the solution to the linear system $\mathbf{\Phi}\beta = b$ has the desired Gaussian distribution. We can then solve the linear system by the conjugate gradient (CG) algorithm through matrix-vector multiplications by $\mathbf{\Phi}$; this involves no explicit factorization or calculation of $\mathbf{\Phi}$ itself. Rapid convergence of CG in this context is guaranteed by the theory of prior-preconditioning we develop. We apply our algorithm to a clinically relevant large-scale observational study with $n = 72,489$ patients and $p = 22,175$ clinical covariates, designed to assess the relative risk of adverse events from two alternative blood anti-coagulants. Our algorithm demonstrates an order of magnitude speed-up in posterior inference, in our case cutting the computation time from two weeks to less than a day.

ARTICLE HISTORY
Received 11 February 2020
Revised 2 March 2022
Accepted 18 March 2022

KEYWORDS
Big Data, Conjugate gradient, Markov chain Monte Carlo, numerical linear algebra, sparse matrix, variable selection

OHDSI Shoutouts!

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

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

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

Upcoming Workgroup Calls

Date	Time (ET)	Meeting
Wednesday	10 am	FHIR and OMOP Digital Quality Measurements Subgroup (ZOOM)
Wednesday	11 am	GIS – Geographic Information Systems Development
Thursday	12 pm	FHIR and OMOP Oncology Subgroup
Thursday	1 pm	OMOP CDM Oncology Vocabulary Subgroup
Thursday	6 pm	FHIR and OMOP Terminologies Subgroup (ZOOM)
Friday	10:30 am	Clinical Trials
Tuesday	10 am	Common Data Model

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

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
- [ ] FHIR and OMOP
- [ ] Geographic Information System (GIS)
- [ ] HADES Health Analytics Data-to-Evidence Suite
- [ ] Healthcare Systems Interest Group (formerly EHR)
- [ ] Health Equity
- [ ] Latin America
- [ ] Medical Devices
- [ ] Medical Imaging
- [ ] Natural Language Processing
- [ ] OHDSI APAC
- [ ] OHDSI APAC Steering Committee
- [ ] OHDSI Steering Committee
- [ ] Oncology
- [ ] Open-source Community
- [ ] Phenotype Development and Evaluation
- [ ] Population-Level Effect Estimation / Patient-Level Prediction
- [ ] Psychiatry
- [ ] Registry (formerly UK Biobank)
- [ ] Surgery and Perioperative Medicine
- [ ] Vaccine Evidence
- [ ] Vaccine Vocabulary

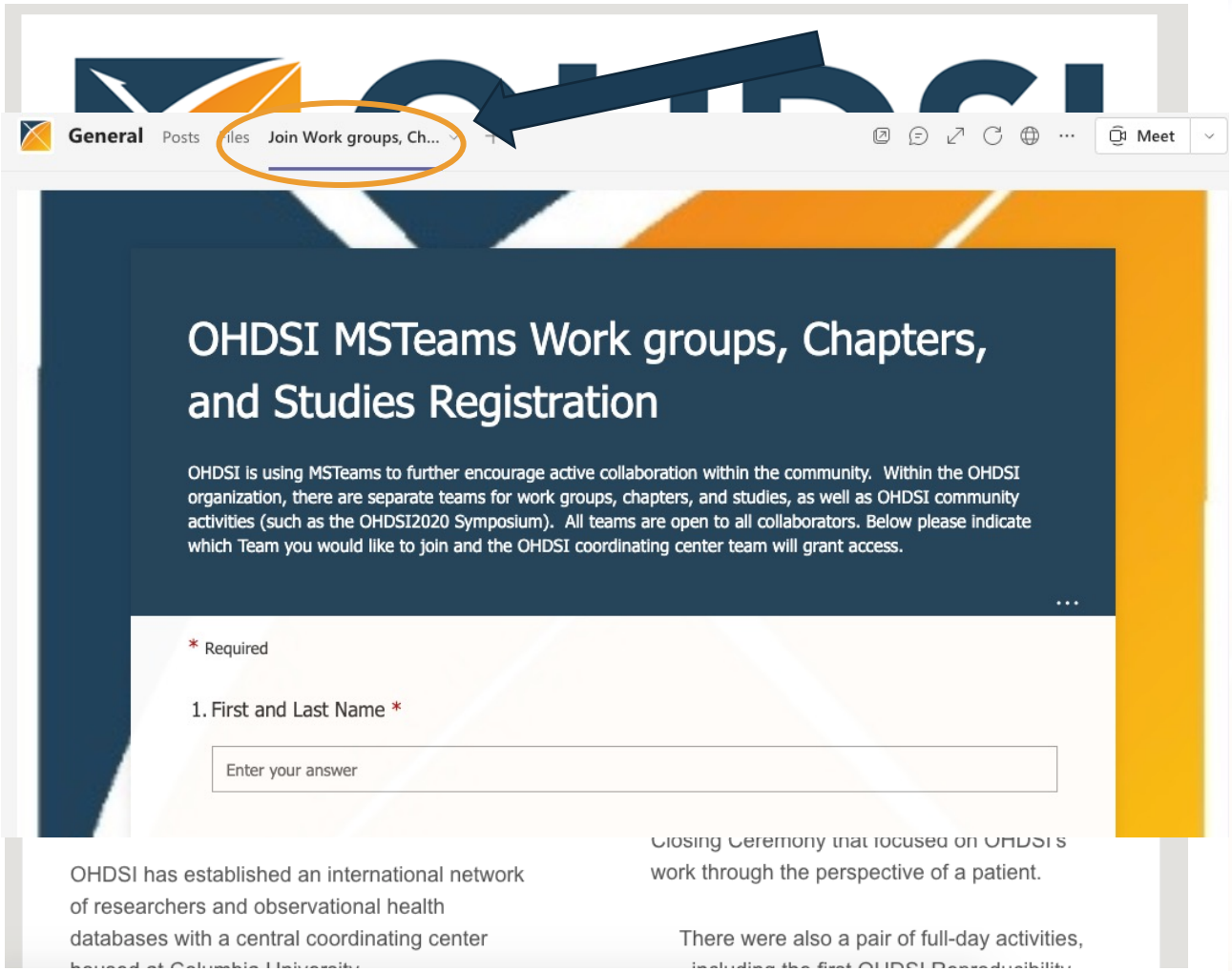
6. Select the chapter(s) you want to join

- [ ] 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)

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
- ☐ FHIR and OMOP
- ☐ Geographic Information System (GIS)
- ☐ HADES Health Analytics Data-to-Evidence Suite
- ☐ Healthcare Systems Interest Group (formerly EHR)
- ☐ Health Equity
- ☐ Latin America
- ☐ Medical Devices
- ☐ Medical Imaging
- ☐ Natural Language Processing
- ☐ OHDSI APAC
- ☐ OHDSI APAC Steering Committee
- ☐ OHDSI Steering Committee
- ☐ Oncology
- ☐ Open-source Community
- ☐ Phenotype Development and Evaluation
- ☐ Population-Level Effect Estimation / Patient-Level Prediction
- ☐ Psychiatry
- ☐ Registry (formerly UK Biobank)
- ☐ Surgery and Perioperative Medicine
- ☐ Vaccine Evidence
- ☐ Vaccine Vocabulary

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)

2022 OHDSI U.S. Symposium

Registration is OPEN for #OHDSI2022!

The 2022 OHDSI Symposium will be held Oct. 14-16 at the Bethesda North Marriott Hotel & Conference Center.





2022 OHDSI Symposium (Oct. 14-16)

Oct. 14: Main Conference

The main conference at the 2022 OHDSI Symposium returns Friday, Oct. 14, at the Bethesda North Marriott Hotel & Conference Center. The conference will include the collaborator showcase; the submission deadline for that is Friday, June 24.

Oct. 15: Full-Day Tutorial

An Introductory Journey From Data To Evidence

This tutorial is meant to introduce newer members of the community to steps along the journey from data to evidence using the OMOP Common Data Model, OHDSI tools and scientific best practices.

2022 OHDSI U.S. Symposium

Saturday: Full-Day Tutorial

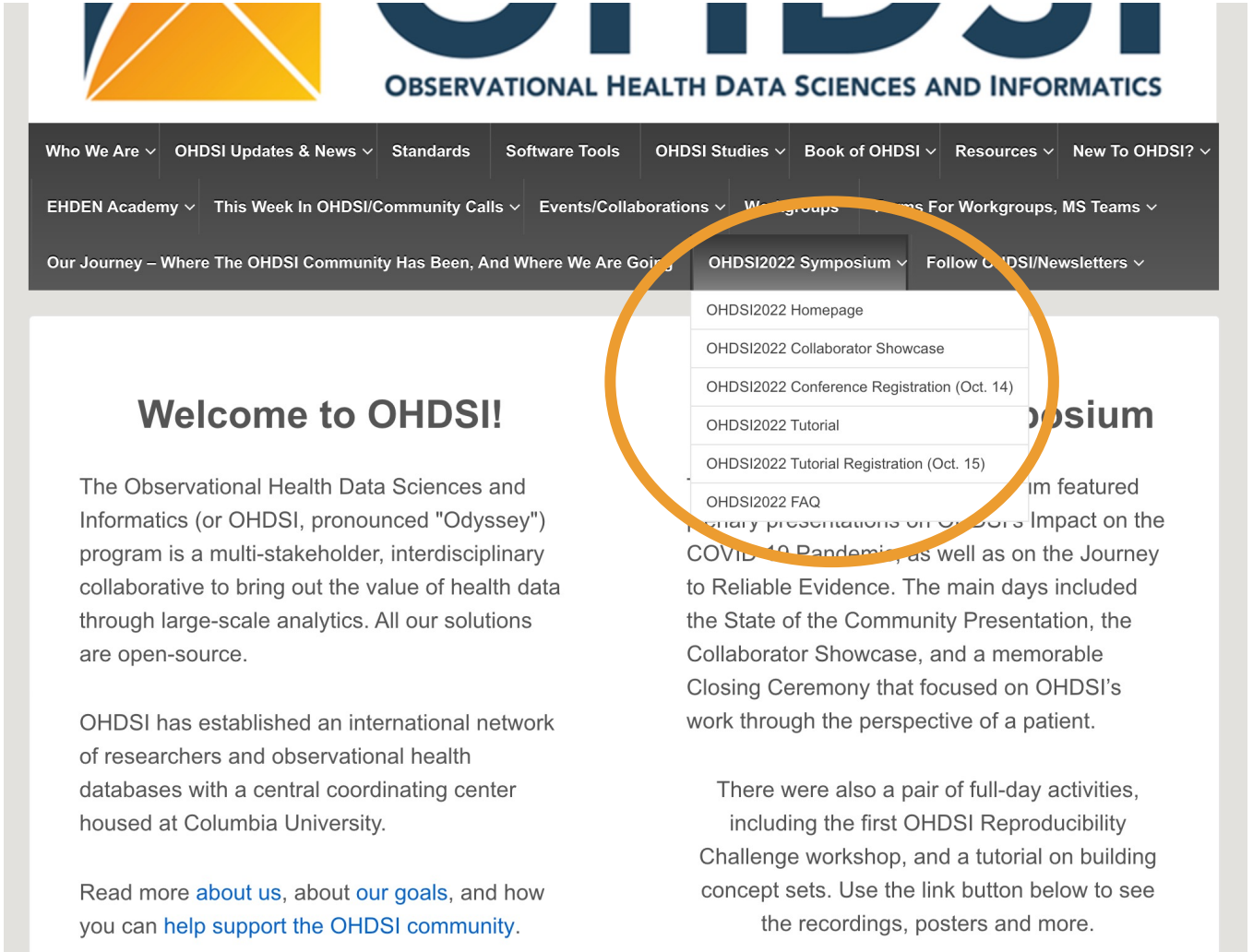
An Introductory Journey From Data To Evidence

This tutorial is meant to introduce newer members of the community to steps along the journey from data to evidence using the OMOP Common Data Model, OHDSI tools and scientific best practices.

Agenda

8:30 am • Overview of the OHDSI Journey: where are we going? — **Patrick Ryan**
9 am • OMOP Common Data Model and vocabulary — **Clair Blacketer**
9:50 am • Energy break
10 am • ETL a source database into OMOP CDM – **TBA**
10:50 am • Energy break
11 am • Creating cohort definitions – **Asieh Golozar**
11:50 am • Lunch break
12:30 pm • Phenotype evaluation – **Gowtham Rao**
1:20 pm • Energy break
1:30 pm • Characterization – **Kristin Kostka**
2:20 pm • Energy break
2:30 pm • Estimation – **Martijn Schuemie**
3:20 pm • Energy break
3:30 pm • Prediction – **Jenna Reps**
4:20 pm • Recap of the OHDSI Journey, where do we go from here? – **George Hripcsak**

2022 OHDSI U.S. Symposium



symposium@ohdsi.org

OHDSI Vocabulary Journey

The March 22 OHDSI Community Call provided an in-depth look at the OHDSI vocabulary, from how it is developed, to how it can be utilized, and where it should grow from here. Three leaders from the vocabulary workgroup joined to present a trio of topics for this session:

"A peek into the OHDSI vocabulary engine room"
– Michael Kallfelz, Physician Executive • Odysseus Data Services

"Fun things you can learn with the OHDSI standardized vocabularies"
– Patrick Ryan, Vice President, Observational Health Data Analytics • Janssen Research & Development; Adjunct Assistant Professor • Columbia University

"Time for reflection • Where are we? Where should we be?"
– Christian Reich, Vice President, RWE Systems • IQVIA

Watch The Presentation Here

Michael Kallfelz, Patrick Ryan and Christian Reich led the above presentation during the OHDSI community call.

The slides from this presentation are available here.

Both Reich and Anna Ostropolets led the Standardized Vocabularies chapter of the Book of OHDSI, which was first released in 2019 and has been updated online since. The introductory paragraph for that chapter is below. You can read the whole chapter here.

The OMOP Standardized Vocabularies, often referred to simply as "the Vocabulary", are a foundational part of the OHDSI research network, and an integral part of the Common Data Model (CDM). They allow standardization of methods, definitions and results by defining the content of the data, paving the way for true remote (behind the firewall) network research and analytics. Usually, finding and interpreting the content of observational healthcare data, whether it is structured data using coding schemes or laid down in free text, is passed all the way through to the researcher, who is faced with a myriad of different ways to describe clinical events. OHDSI requires harmonization not only to a standardized format, but also to a rigorous standard content.

ohdsi.org/ohdsi-vocabulary-journey

OHDSI Dev Con

April 22, 2022 (8 am – 12 pm)

The Open-Source Community is hosting the first **Dev Con** as a way of accepting and mentoring new contributors to our environment. We are planning multiple workshops, talks and a panel discussion to both welcome and engage both current and future developers within OHDSI.

Don't miss this opportunity! Use the link at the bottom to register!

Time	Topic
8 am	Open-Source Workshops
10 am	State of the OHDSI Community (Paul Nagy, Adam Black)
10:20 am	Keynote – Grand Vision for OHDSI Software Ecosystem (Martijn Schuemie)
11 am	Industry Panel Discussion (How Do/Should We Connect It All Together?)

bit.ly/OHDSIDev22

Are You Interested In …

- participating with an OHDSI project team?
- seeing 'under the hood' of the OHDSI engine?
- being mentored by professional developers?

Use This Link To Register Today!

DevCon Agenda

Time (ET)	Track 1	Track 2
8 am	ATLAS **(Anthony Sena)**	HADES Introduction **(Adam Black)**
8:30 am	WebAPI **(Anthony Sena)**	CohortDiagnostics **(James Gilbert)**
9 am	White Rabbit/Rabbit In A Hat **(Maxim Moinat)**	Patient-Level Prediction **(Jenna Reps)**
9:30 am	Data Quality Dashboard **(Clair Blacketer)**	Cyclops **(Marc Suchard)**
10 am	State of OHDSI Development **(Adam Black and Paul Nagy)**	
10:20 am	Keynote **(Martijn Schuemie)**	
11 am	Panel Discussion (Putting The Pieces Together) **Lee Evans** - Broadsea (OHDSI) **Cory Stevenson** - OHDSI on Azure (Microsoft) **Paul Sexson** - OHDSI In A Box (AWS) OHDSI on GCP	

Next CBER Best Seminar

Topic CBER BEST Seminar Series - Addressing Selection and Confounding Bias in Test-Negative Study Designs for Flu and COVID-19 Monitoring

Description: The test-negative design (TND) has become a standard approach to evaluate vaccine effectiveness against the risk of acquiring infectious diseases such as Influenza, Rotavirus, Dengue fever and more recently COVID-19 in real world settings. Despite the TND's potential to reduce unobserved differences in healthcare seeking behavior (HSB) between vaccinated and unvaccinated subjects, substantial variability in unobserved HSB may remain among study participants. As latent HSB is likely also a strong predictor of selection into the TND sample, confounding bias of the vaccine's causal effect by latent HSB may be induced by collider stratification bias resulting from the TND.

Speakers

Dr. Eric Tchetgen Tchetgen

Luddy Family President's Distinguished Professor @Wharton School of the University of Pennsylvania

Eric J. Tchetgen Tchetgen is the Luddy Family President's Distinguished Professor at the Wharton School of the University of Pennsylvania. Professor Tchetgen Tchetgen comes to the University of Pennsylvania from Harvard University, where he has served since 2008 as Professor of Biostatistics and Epidemiologic Methods with joint appointments in the departments of Biostatistics and Epidemiology at the T.H. Chan School of Public Health. He researches infectious diseases, including HIV/AIDS, and the role of genetic and social factors in the patterns, causes, and effects of public health. Professor Tchetgen Tchetgen has received grants from the National Institutes of Health and the Centers for Disease Control. He completed his Ph.D. in Biostatistics at Harvard University in 2006 under the supervision of Professor James M. Robins. He received his B.S. in Electrical Engineering from Yale University in 1999.

Wed., April 27, 11 am ET

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?

March 29: Reproducibility

Anna Ostropolets

PhD Student • Columbia University

Topic: The OHDSI2021 Reproducibility Challenge

Martijn Schuemie

Research Fellow, Epidemiology Analytics • Janssen Research and Development

Topic: Developing Reproducible Studies

Asieh Golozar

Vice President, Global Head of Data Science • Odysseus Data Services, Inc.

Topic: The Reproducibility Service, via the OHDSI Center at the Roux Institute

Polina Talapova, MD