

### **Networking Session**

OHDSI Community Call April 27, 2021 • 11 am ET



### **April/May OHDSI Community Calls**

Date	Topic
April 27	OHDSI Networking Session
May 3	Workgroup Updates: GIS, PEI, Registry (formerly UK Biobank), EHR
May 10	OHDSI Debates
May 17	Focus Topic: Prostate Cancer Study-A-Thon
May 24	OHDSI Fun



### **April/May OHDSI Community Calls**

Date	Topic
April 27	OHDSI Networking Session
May 3	Workgroup Updates: GIS, PEI, Registry (formerly UK Biobank), EHR
May 10	OHDSI Debates
May 17	Focus Topic: Prostate Cancer Study-A-Thon
May 24	OHDSI Fun



### **Next Week: Workgroup Presentations**



**GIS - Geographic Information System** 

**Robert Miller, Tufts** 



Registry (formerly UK Biobank)

**Maxim Moinat, The Hyve** 



PEI - Pharmacovigilance Evidence Investigation

Erica Voss, Janssen R&D



**Electronic Health Records** 

Melanie Philofsky, Odysseus Data Services



### Three Stages of The Journey

### Where Have We Been? Where Are We Now? Where Are We Going?





#### **OHDSI Shoutouts!**

> J Biomed Inform. 2021 Apr 19:103790. doi: 10.1016/j.jbi.2021.103790. Online ahead of print. **Building An OMOP Common Data Model-Compliant** Annotated Corpus for COVID-19 Clinical Trials

Yingcheng Sun 1, Alex Butler 2, Latoya A Stewart 3, Hao Liu 1, Chi Yuan 1, Christopher T Southard <sup>3</sup>, Jae Hyun Kim <sup>1</sup>, Chunhua Weng <sup>4</sup>

Affiliations + expand

PMID: 33887457 DOI: 10.1016/j.jbi.2021.103790

#### Abstract

Clinical trials are essential for generating reliable medical evidence, but often suffer from expensive and delayed patient recruitment because the unstructured eligibility criteria description prevents automatic query generation for eligibility screening. In response to the COVID-19 pandemic, many trials have been created but their information is not computable. We included 700 COVID-19 trials available at the point of study and developed a semi-automatic approach to generate an annotated corpus for COVID-19 clinical trial eligibility criteria called COVIC. A hierarchical annotation schema based on the OMOP Common Data Model was developed to accommodate four levels of annotation granularity: i.e., study cohort, eligibility criteria, named entity and standard concept. In COVIC, 39 trials with more than one study cohorts were identified and labelled with an identifier for each cohort. 1.943 criteria for non-clinical characteristics such as "informed consent", "exclusivity of participation" were annotated, 9767 criteria were represented by 18,161 entities in 8 domains. 7,743 attributes of 7 attribute types and 16,443 relationships of 11 relationship types. 17,171 entities were mapped to standard medical concepts and 1,009 attributes were normalized into computable representations. COVIC can serve as a corpus indexed by semantic tags for COVID-19 tria and analytics, and a benchmark for machine learning based criteria extraction.

Keywords: COVID-19; Clinical Trial; Eligibility Criteria; Machine Readable Dataset; Structure Corpus.

Congratulations to the team of Yingcheng Sun, Alex Butler, Latoya Stewart, Hao Liu, Chi Yuan, Christopher Southard, Jae Hyun Kim, and Chunhua Weng on this recent study in the Journal of Biomedical Informatics: Building An OMOP Common Data Model-**Compliant Annotated Corpus for COVID-19 Clinical Trials.** 



#### **OHDSI Shoutouts!**

Congratulations to the team of George Hripcsak, Martijn Schuemie, David Madigan, **Patrick Ryan and Marc Suchard** on this recent paper in the Yearbook of Medical Informatics: **Drawing Reproducible Conclusions from Observational** Clinical Data with OHDSI.

#### **Drawing Reproducible Conclusions from Observational Clinical** Data with OHDSI

George Hripcsak, Martijn J. Schuemie, David Madigan, Patrick B. Ryan, Marc A. Suchard

- > Author Affiliations
- > Further Information

Abstract References

> Permissions and Reprints

#### Summary

Objective: The current observational research literature shows extensive publication bias and contradiction. The Observational Health Data Sciences and Informatics (OHDSI) initiative seeks to improve research reproducibility through open science.

Methods: OHDSI has created an international federated data source of electronic health records and administrative claims that covers nearly 10% of the world's population. Using a common data model with a practical schema and extensive vocabulary mappings, data from around the world follow the identical format. OHDSI's research methods emphasize reproducibility, with a large-scale approach to addressing confounding using propensity score adjustment with extensive diagnostics; negative and positive control hypotheses to test for residual systematic error; a variety of data sources to assess consistency and generalizability; a completely open approach including protocol, software, models, parameters, and raw results so that studies can be externally verified; and the study of many hypotheses in parallel so that the operating characteristics of the nethods car

Results: OHDSI has already produced findings in areas like hypertension treatment that are being into practice, and it has produced rigorous studies of COVID-19 that have aided government agen treatment decisions, that have characterized the disease extensively, that have estimated the coneffects of treatments, and that the predict likelihood of advancing to serious complications.

Conclusions: OHDSI practices open science and incorporates a series of methods to address reproduct has produced important results in several areas, including hypertension therapy and COVID-19 research



### **HADES** — And How You Can Help

### The Quiet Driving Force for Observational Research, HADES Empowers a Global Community — Beyond Just OHDSI — To Generate Real-World Evidence

Certain factors for the success of an open-science community like OHDSI are more obvious than others. When hundreds of people come together to research a common cause, or studies are run against millions of patient records in a global database, it becomes clear that something impactful is happening.

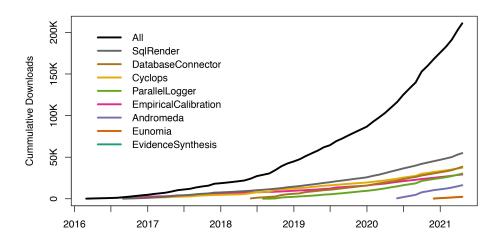
One critical factor in OHDSI's ability to perform rigorous, ground-breaking analyses lies under the surface, but it holds an equally important role in the overall community mission. A core foundation for OHDSI is open-source software development, and a small group of community collaborators, led by Martijn Schuemie, has generated a collection of analytics tools that enable research both in and out of the OHDSI community.

HADES — the Health Analytics Data-to-Evidence Suite — is a set of 20 open-source R packages for large scale analytics, including population characterization,

Contractational
 Commission Controlled Case Series analysis static commission of series and controlled controlled Case Series analysis static commission of series and controlled Case Series analysis static commission of series and controlled Case Series analysis static commission of series and controlled Case Series analysis static segment of series and controlled Case Series analysis static segment of series and controlled Case Series analysis static segment of series and controlled controlled Case Series analysis static segment of segment of series series and controlled controlled Case Series analysis static segment of segment of series series and controlled Case Series analysis state of series segment segment of series segment segment of series segment segment

HADES is a set of 20 open-source R packages that have been developed within the OHDSI community to aid observational research.

population-level causal effect estimation, and patient-level prediction, as well as supporting packages that are critical throughout the journey of observational research. The packages offer a robust set of functions that together can be used to perform all the steps required to conduct a network study, from connecting to a database, translating queries into the appropriate SQL dialect, generating cohorts and extracting features, fitting large-scale statistical models, compiling results for meta-analysis and empirical calibration, and enabling exploration through interactive visualization dashboards.





We are very proud of the impact that HADES continues to make on real-world evidence generation. Our team develops, tests and continuously monitors a set of tools that empowers global research using best practices developed within our community.

- Martijn Schuemie



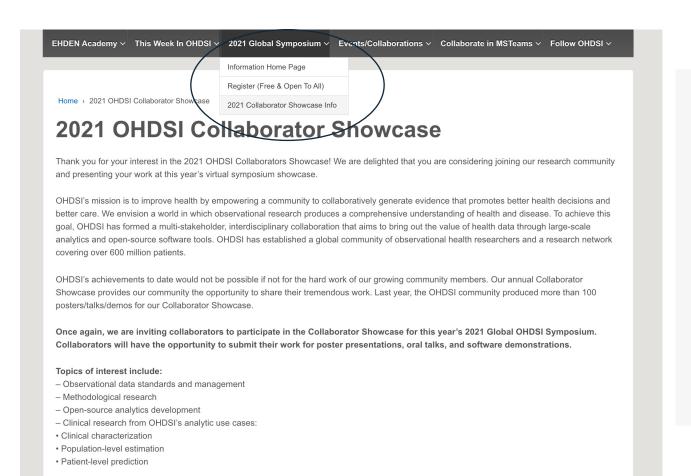
ohdsi.org/hades-empowers-global-community/





#### **#OHDSI2021 Call For Participation**





### **OHDSI**

#### 2021 OHDSI Collaborators Showcase Brief Report Submission Form-Posters, Oral Talks and Software Demonstrations

Thank you for your interest in the 2021 OHDSI Collaborators Showcase! We are delighted that you are interested in showcasing your work at this year's virtual symposium showcase. For planning purposes, please take a few minutes to fill out this submission form to help the Scientific Review Committee better understand your solution.

By filling out this form you may choose if you would like to present your work as a poster, an oral talk or a software demonstration (or all three). If a poster or software demo, you will present it during the Collaborator Showcase at the symposium. If an Oral talk, you will present an estimated 7-minute talk at the symposium.



#### **OHDSI Shoutouts!**



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

Have a study published? Please send to <a href="mailto:sachson@ohdsi.org">sachson@ohdsi.org</a> 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	8 am	Vaccine Vocabulary
Wednesday	10 am	OMOP CDM Oncology – Development Subgroup
Wednesday	1 pm	Data Quality Dashboard (rescheduled from last week)
Thursday	8 am	Psychology
Thursday	1 pm	OMOP CDM Oncology – CDM/Vocabulary Subgroup
Friday	10 am	Electronic Health Records
Tuesday	9 am	OMOP CDM Oncology – Genomic Subgroup

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



#### **Get Access To Different Teams/WGs/Chapters**



4. Select the workgroups you want to join (you can r www.ohdsi.org/web/wiki/doku.php?id=projects:ov	
_ ATLAS	
Clinical Trials	
Common Data Model	Phenotype Development and Evaluation
Data Quality Dashboard Development	Population-Level Effect Estimation / Patient-Level Prediction
Early-stage Researchers	Psychiatry
Education Work Group	Registry (formerly UK Biobank)
Electronic Health Record (EHR) ETL	Surgery and Perioperative Medicine  Vaccine Vocabulary
Geographic Information System (GIS)	☐ Women of OHDSI
HADES Health Analytics Data-to-Evidence Suite	
Latin America	5. Select the chapter(s) you want to join
Medical Devices	Australia
Natural Language Processing	China
OHDSI APAC	☐ Japan
OHDSI APAC Steering Committee	☐ Korea
OHDSI Steering Committee	Singapore
Oncology	Taiwan
Patient-Generated Health Data	
Pharmacovigiliance Evidence Investigation	6. 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)



#### **APAC Update**



#### **OHDSI APAC - Our Asia-Pacific Community**

OHDSI is a global, multi-stakeholder, interdisciplinary and openscience network that collaborates to bring out the value of health data through large-scale analytics. Our Asia-Pacific (APAC) community comprises six regional chapters (Australia, China, Japan, Singapore, South Korea, Taiwan) and has led important OHDSI initiatives around the world.

The first OHDSI APAC Symposium was held virtually Dec. 5-6, 2020, and it provided an opportunity for our Asia-Pacific collaborators (along with others around the world) to share research, provide community updates, collaborate in regional breakouts or network discussions, and have some OHDSI fun.

Date	Topic
March 25	Network Research
April 8	OMOP Projects
April 22	Regional Updates
May 6	10-Minute Tutorials
May 20	Collaborator Showcase Presentations
June 3	Network Research
June 17	Welcome To OHDSI
July 1	OMOP Projects

Upcoming APAC Community Calls

The APAC community has its own group in the OHDSI MS Teams environment to promote greater collaboration on our collaborative efforts. It will also be the home of our bi-weekly APAC Community Calls, where we will provide updates, share research presentations, collaborate on topics of shared interest, and plenty more. How can you get involved? First, request access to our MS Teams Environment, then request access to our OHDSI APAC workgroup. You can also join the meetings by using this direct link.

Our calls are scheduled on a bi-weekly basis, held every other Thursday/Wednesday depending on where you live. All calls are recorded, so if you miss one or want to share with a collaborator, check out the call recordings below.

- Apr. 22, 2021 - Regional Updates



# APAC Community Call Regional Updates

April 22<sup>nd</sup> 2021



#### Korea Chapter Update



 Nationwide Korean administrative claim data (HIRA) has been mapped to OMOP-CDM ranging 2010 to 2019 for entire Korean population



 More Korean hospitals including three hospitals in Yonsei University Health System (Severance; Gangnam Severance; Yongin Severance hospital) joined OHDSI with available OMOP-CDM v5.3.1 databases



### China Chapter Update



- Organization change
  - Three new co-Chairs:
    - Dr. Yi Zhou, Zhongshan University
    - Dr. Lei Liu, Fudan University
    - Dr. Hui Lv, Jiaotong University
  - Three directions
    - Training and engagement
    - Methods and tools
    - Clinical studies

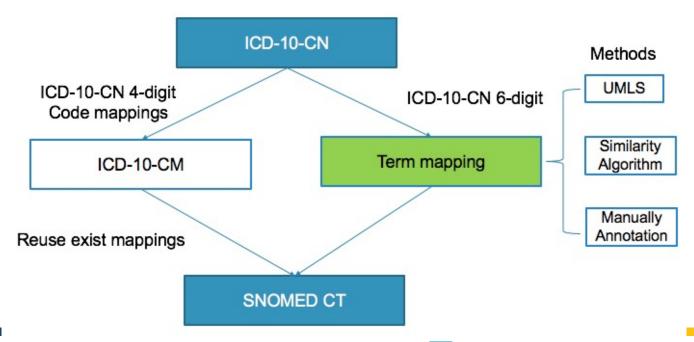


### Chinese Vocabulary Work



- Five volunteers for Chinese Vocabulary Development
- Task 1: Local concepts to ICD-10-CN and SNOMED
- Task 2: ICD-10-CN 6-digit to SNOMED

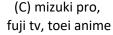
N	Number of Concepts
1	15707
2	231537
3	6245 (2013)
	78658 (2014)
	83295 (2015)
4	1000 outpatient, 1000 inpatient
5	1910
6	1166





### Japan Chapter Update









- A person who wants to work on drug mapping has came along and created a plan.
- Going to start a mapping project for drug.

#### Taiwan Chapter Update



 Goal 1: OHDSI Taiwan website currently under development









- Goal 3: Build TMU OHDSI OMOP-CDM
  - OHDSI-OMOP CDM kick-off meeting on September 21, 2020, to start the data bridging platform between TMUCRD and OHDSI-CDM.
  - OHDSI global members provided technical guidance online meeting on September 23, 2020.

#### Goal 2: Interaction between Universities in Taiwan

















#### **OMOP** mapping in Australia



#### **PRIMARY CARE EMR**

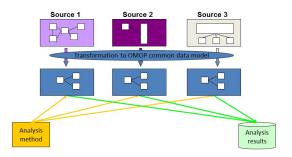
- UNSW ePBRN
  - 18 general practices and hospitals in an integrated health neighbourhood in South Western Sydney
- Victoria
  - 3 GP systems (**120 practices**, several million patients)
    - Medical Director, Best Practice and Zedmed
  - Outcome Health (1,000 GP surgeries)
- NPS MedicineWise MedicineInsight database
  - 700+ general practice information systems from all over Australia

#### **SECONDARY CARE EMR**

- Hospital EMR warehouses
  - CERNER2OMOP: Maps CERNER EMR data into OMOP (customized to SLHD and SESLHD. Contact Blanca Gallego Luxan
  - CaVa: Maps oncology data into OMOP oncology extension (MOSAIQ and ARIA) (customized to SWSLHD) Contact Timothy Churches
  - Victorian Comprehensive Cancer Centre Data Hub (Extensive Victorian hospital and GP coverage)
  - Proposed CERNER based EMR
    - Queensland Health all public hospitals
    - Austin Health
    - Western Health
    - Provisional: Peter MacCallum Cancer Centre

#### **ADMINISTRATIVE CLAIMS**

- Phamaceutical Benefits Scheme
   10% Sample dispensed medicines
- Department of Veterans' Affairs
   Claims Data
  - National Department of Veterans' Affairs administrative health claims





### OHDSI translation at UniSA



- Successfully converted PBS10 to the OHDSI/OMOP CDM
  - Version 5.3.1
  - Clinical OMOP CDM tables:
    - Person, Obs\_Period, Drug\_Exposure, Drug\_Era, Death
  - Atlas set up in the locked down computing environment

#### • Stats:

- Data period: March 2005 July 2020 inclusive
- ~ 3 million people (~10% of Aus population)
- ~ 300 million records
- 95%++ conversion of drug records to standardised drug concepts



### **OHDSI Network Collaboration**



- Ongoing work from OHDSI COVID-19 study-a-thon
- UNSW contributed flu data 2012-2019 for two COVID-19 studies related to predictive modelling.
- One study is published <a href="https://medinform.jmir.org/2021/4/e21547">https://medinform.jmir.org/2021/4/e21547</a> and other under review with BMC Research Methodology.
- OHDSI APAC research study on hypertension
  - A Comprehensive Comparative Effectiveness and Safety Study of the Second Antihypertensive Agent after Monotherapy at scale using the OHDSI AP Network
- EUMAEUS
  - large-scale evaluations of methods for vaccine safety to help us understand how these methods will perform when applied to COVID-19 vaccines
- Others?



### **Data Quality**



- Extend OHDSI CDM-based data quality assessments
  - "White bandicoot" tool
- Mapping Australia's metadata and ontologies for national health & welfare metadata standards
  - Sustainable conformance to data quality and interoperability standards



#### Please join us for the start of the 2021 ODHSI Australia Webinar Series

- Presented by Dr Tim Churches, Georgie Kennedy and Nasreen Kadaan:
- -"CaVa: Populating the OMOP oncology extension directly from clinical data"
- A description of the implementation of the OMOP CDM from a cancer clinical information system.
- -Thursday 29th April at 2pm AEST



# Collaboration with Al Singapore on Federated Learning





#### **Four Key Pillars**



#### **Al Research**

Invest in deep capabilities to catch the next wave of scientific innovations and breakthroughs.



#### **Al Technology**

Address major challenges
that affect our economy
and society thereby
promoting bold ideas and
the application of innovative
Al technologies



#### **Al Innovation**

Broaden the use and adoption of AI in Singapore and groom local AI talents to support industry growth



#### **Al Makerspace**

Provide a platform and suite of AI tools, APIs and prebuilt solutions (BRICKS), training, consulting and engineering services for Startups and SMEs to jumpstart their AI journey



#### Collaboration with AI Singapore on Federated Learning

#JoinTheJourney





**Dr Jianshu Weng Project Lead, AI SG** 



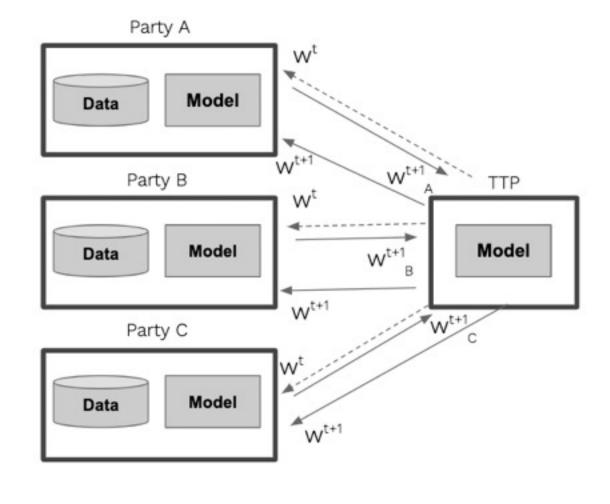




### **AISINGAPORE**



Smart Nation NGAPORE





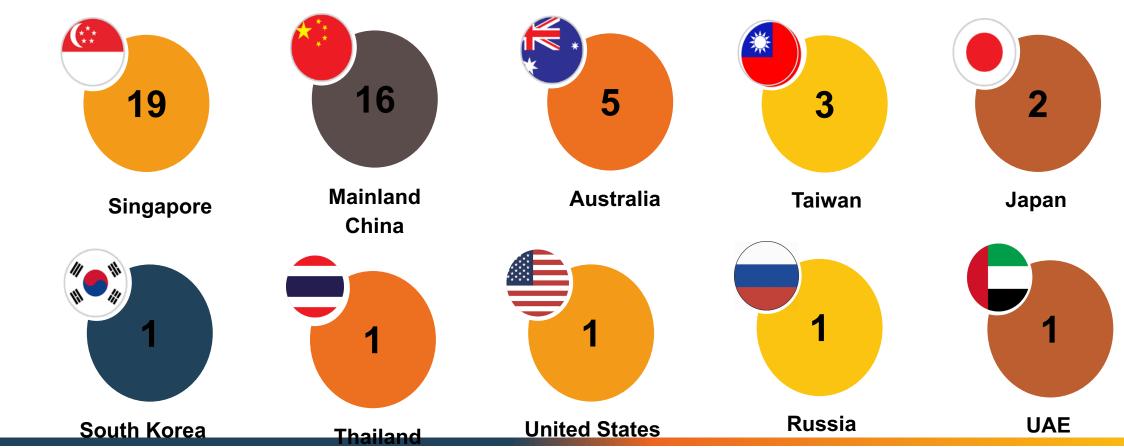
# SG Healthcare AI Datathon and EXPO 10-18 July 2021



#### www.nus-datathon.com



Attracted 1000+ physicians and data science around the world







#### **Job Opening**



#### Vacancy – OMOP expert / health database researcher

23rd April 2021

One of our EHDEN data partners, the UK Clinical Practice Research Datalink (CPRD), has a vacancy for an OMOP expert / health database researcher. CPRD has converted the CPRD Aurum primary care electronic health record (EHR) database into an OMOP common data model (CDM) version and are seeking an individual with expertise in using the OMOP CDM and strong technical skills to support ongoing validation work. The successful candidate will also train users within CPRD in the use of the OHDSI analytical infrastructure. The position will be hosted within the Observational Research team, but will require close collaboration with the Health Data Science team who manage the technical implementation.

For more information and a detailed vacancy specification please contact:

Darren Lunn, Head of Health Data Science, CPRD (darren.lunn@mhra.gov.uk)



#### **CBER Best Initiative Seminar Series**

The next edition of the CBER Best Initiative Seminar Series will be held Wednesday, May 5, at 11 am ET. Dr. Ben Goldstein will discuss "Understanding Informed Presence in Electronic Health Records Data" during that session.

Registration is free, and everybody is invited. You can register at:

https://northeastern.zoom.us/webinar/register/WN\_UkPMnpakQ4 qRMJS7O459pA

#### **CBER BEST Initiative Seminar Series**



Date:

May 5, 2021

Time

11:00 -12:00 PM EDT

Topi

Understanding Informed Presence in Electronic Health Records Data

#### 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 course of the year.

**Description:** Electronic Health Records (EHR) data have become a key data source for clinical research. Their relative availability and abundance make them very appealing for analytic tasks ranging from comparative effectiveness research, disease surveillance, population health, and predictive modeling. However, as has been widely documented, as a real-world data source, there are a number of analytic challenges with EHR data. One of the key challenges of EHR data relates to the observability of the underlying data. While fundamentally a missing data problem, we have termed this process *informed presence* to highlight that what we observe is informative. In this talk I will provide some illustrations for how informed presence can bias insights and inference with EHR data as well as discuss some approaches – both analytical and design based – that can mitigate these biases. Ultimately EHR data are an extremely useful data source, though like any complicated data source need to be used thoughtfully.

#### Presenter:

Dr. Ben Goldstein, PhD, MPH



Benjamin Goldstein is Associate Professor of Biostatistics and Bioinformatics at Duke University. He is a member of the Duke Clinical Research Institute and serves as the Science Lead for the Children's Health Discovery Initiative. Dr. Goldstein's research focuses on the meaningful use of electronic health records data. His work sits at the intersection of biostatistics, biomedical informatics, machine learning and epidemiology. He works closely with the Duke University Health System developing, implementing and evaluating risk prediction and clinical decision support tools. He also studies how patients' informative visit process can impact inference in EHR based studies. Dr. Goldstein received his PhD in Biostatistics and MPH in Biostatistics and Epidemiology from UC Berkeley.

Registration:

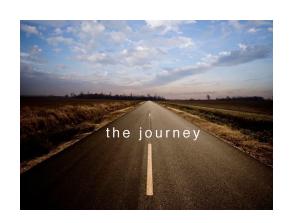
https://northeastern.zoom.us/webinar/register/WN\_UkPMnpakQ4qRMJS7O459pA





### 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?













Top Left: Network Studies, COVID, Vaccines, Oncology, T2DM, Psychiatry

Bottom Left: Characterization, New To OHDSI, Estimation, Methods Research, Prediction, ATLAS, Phenotyping, HADES, OMOP CDM, Mentoring

**Top Right:** Open Tables

Also In Classroom: 3 Whiteboards





#### **URL:**

https://gather.town/app/QGqIYUyhyvYIlaKE/OHDSINetworkSessionAPR27

Password: OHDSI2021!

Links and passwords will be posted in chat. If you can't access the chat, they are posted on our Teams General page. Create your avatar, enter and network with our community!



#### Apr 27 Community Call Focus: OHDSI Networking Session 🖋

General



#### CraigSachson



Please join us Tuesday, April 27, at 11 am ET for our next OHDSI Community Call, which will be our first Network Session-themed call.

Following our community updates, we will use the website GatherTown 11 to facilitate this session and allow all attendees to connect with small or large groups to discuss various topics, or simply to meet a friend and catch up. No download is required beforehand, and everybody is invited!

Many of you will have the calendar link in your OHDSI Teams environment. If not, you will need the specific meeting link to access the community call; it will also be posted in the main OHDSI Teams environment. As always, if you can't make it, the recordings will be posted to both our Community Calls page and the General OHDSI Teams recordings folder.

Link for Gathertown session: https://gather.town/app/QGqIYUyhyvYllaKE/OHDSINetworkSessionAPR27.





#### **URL:**

https://gather.town/app/QGqIYUyhyvYllaKE/OHDSINetworkSessionAPR27

Password: OHDSI2021!

#JoinTheJourney

Links and passwords will be posted in chat. If you can't access the chat, they are posted on our Teams General page. Create your avatar, enter and network with our community!