Our Journey

Where The OHDSI Community Has Been
And Where We Are Going

OHDSI
Observational Health Data Sciences and Informatics
# Join The Journey

OHDSI.org

To improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care.

## Table of Contents

I. Letter To The Community ............................................ 2

II. Mission, Values, And What We Do ............................. 5
   - Research Flow .................................................. 7
   - Columbia University as Coordinating Center .............. 8

III. Collaborators .......................................................... 9
   - Map of Collaborators ......................................... 10
   - Organizations Involved with OHDSI ..................... 12
   - Testimonials ...................................................... 14
   - The Titan Awards .............................................. 16
   - OHDSI and Large Community Initiatives ............... 18

IV. Collaborative Activities ............................................. 19
   - Working Groups ............................................... 20
   - Regional Chapters ........................................... 21
   - Community Calls .............................................. 22
   - Study-A-Thons and Other Events ............................. 24
   - The Book of OHDSI ............................................ 26
   - The OHDSI Symposium ....................................... 28
   - Collaborator Showcase ....................................... 30
   - EHDEN Academy ................................................ 32

V. Data Standards ......................................................... 33
   - OMOP Common Data Model .................................. 34
   - Data Partners ..................................................... 36
   - Vocabularies ...................................................... 38

VI. Open-Source Software ................................................ 41
   - HADES ............................................................. 42
   - ATLAS ............................................................... 44

VII. Methods Research .................................................... 45
   - Empirical Calibration ......................................... 46
   - Principles of the LEGEND Project ......................... 47
   - The LEGEND Project ......................................... 48
   - Patient-Level Prediction Models ......................... 50

VIII. Publications .......................................................... 53
   - Collaborations Graph ......................................... 54
   - OHDSI Publications (2010 - July 2021) ................... 56

IX. COVID Contributions ................................................ 73
   - 88 Hours: The Story of the COVID Study-A-Thon ....... 74
   - The 2020 Global Symposium .................................. 78
   - OHDSI Impact in 2020 ENCePP Guide .................... 79
   - The Rise and Fall of Hydroxychloroquine ................. 80
   - Project CHARYBDIS ......................................... 82
   - The SCYLLA Study ............................................. 83
   - Work Around Vaccine Surveillance ....................... 84

IX. Join The Journey ........................................................ 85
   - Cheers From Our Global Community ....................... 86
   - Closing Letter .................................................. 89
   - How You Join The Journey? ................................. Inside Back Cover
OHDSI is a special group of people.

Every once in a while, a group of people working somewhat outside the system accomplishes something that the system could not accomplish and that was even thought impossible. Somewhat like the mostly apocryphal “they built it in their garage,” OHDSI was born of meetings at beaches, forests, living rooms, pubs, some musicals, and, yes, at work.

And by working together, remaining open, and being self-critical, OHDSI managed to attract thousands of researchers overseeing records on more than 10% of the world’s population, carrying out hundreds of thousands of hypothesis tests at once using systematic designs that reduce bias and multiply impact.

The clinical results have had far-reaching consequences, affecting hundreds of millions of people, including work on hypertension treatment, diabetes, and COVID-19 vaccination and treatment. I believe that OHDSI is barely understood or recognized yet, and that is due in large part on its focus on reliable research, getting it right rather than getting it advertised. Yet getting a sizable portion of the world population’s health records into a common data model and making it accessible to thousands of researchers with advanced tools and methods, and then actually following through to generate evidence that is published in the world’s top journals is a monumental achievement.

To be fair, OHDSI came out of a large initiative called Observational Medical Outcomes Partnership (OMOP), mandated by the federal government, funded centrally by the pharmaceutical industry, coordinated by a quasi-governmental office, and staffed by researchers from academics and industry. Its goal was to conduct methods research for drug safety surveillance, and it successfully delivered its remit. However, its real enduring success was innovation in a way of working, through transparency and collaboration. OMOP developed a common data model that was used not only for its own experiments but could be applied to other efforts. On its five-year completion, with its original aim delivered, OMOP researchers recognized there could still be more opportunities to impact public health—by applying what was learned about methodological best practice and collaborative innovation to the task of generating reliable evidence.

OHDSI formally began in December of 2014 as an affirmative vote in the Department of Biomedical Informatics at Columbia University to serve as its coordinating center. One of its most important initial acts was learning from similar open science-efforts like OpenMRS; OHDSI drafted a mission statement focused on community and the ultimate goal of generating evidence that promotes better health decisions and better care. Seemingly simple, it has served as the bedrock for prioritizing and decision making. It permeates not just the major decisions but also the day-to-day operations. Whether it is evolving our data standards or expanding terms in the vocabulary, conducting methodological experiments, developing new open-source software, or initiating an OHDSI network study, we want all collaborative activities aimed at advancing the mission.

As the OHDSI community grows in number, its structure evolves, including the addition of new OHDSI centers. Erasmus MC has led important efforts to build the OHDSI community across Europe. The European Health Data Evidence Network (EHDEN) started as a large IMI-funded project to build a federated data network, and it has also established the EHDEN Academy as an open educational platform for data standardization and observational research. Northeastern University has recently launched the OHDSI Center at Roux, with plans for an OHDSI laboratory, a training component, and advanced methods research. OHDSI chapters like those in the Asia-Pacific region have helped regional groups engage in OHDSI, helping to address differences in time zone and language. OHDSI strives to engage more researchers and data sources in Africa and South America.

OHDSI is perhaps still best known for its OMOP Common Data Model, as that has in effect been its biggest export. The model was created under the original initiative, and OHDSI retained its name to avoid confusion among legacy users. OHDSI has substantially evolved this open community data standard over the years and greatly expanded the vocabularies that serve as the backbone to this deep information model. The following for the OMOP Common Data Model is large and includes the All of Us Research Program, the eMERGE program, the National COVID Cohort Collaborative (N3C), the national data network in Korea, and numerous other initiatives.

Despite this data model success, OHDSI remains focused on the main mission, evidence generation. Its framework for evidence generation—characterization, estimation, and prediction—has turned out to be a valuable organizing principle. OHDSI has been a leader on several fronts. It’s focus on scale — many cases, many
variables, many hypotheses—permeates all three types of evidence, allowing OHDSI to demonstrate the operating characteristics of its analyses and to cover large areas of medicine. It practices extreme openness, with public pre-specified designs, open-source software, study diagnostics, and results. OHDSI is pushing methods research and development, advancing the state of the art in causal inference and machine learning, while also writing new statistical software because no existing tools can handle the scope of the problems we seek to answer, with hundreds of millions of patient records and tens of thousands of variables used to fit models for hundreds of thousands of hypotheses.

The emergence of COVID-19 raised the urgency of OHDSI’s mission and caused a shift in its operations and organizational structure. Current data became more important with a tight coupling between the observational researchers and the data generators. Research design and shepherding shifted from a small leadership team to a larger group engaged in the steering of research and the generation of evidence. And that, in turn, led to a multiplication of the evidence generated and expansion of influence on government policies, with examples being the recommendation against the use of hydroxychloroquine, the recommendation in favor of continuing ACE inhibitors and ARBs in the setting of COVID-19, and the reinstatement of the AstraZeneca vaccine in the setting of early clotting reports.

All of this has been achieved through the OHDSI community. It nurtured a culture of collaboration, encouragement, tolerance, generosity of time, preeminence of truth, and necessity of action. OHDSI has become a home away from home for many. OHDSI strives to improve itself, seeking to achieve equity both in its research results and among those who generate them.

Around the world, committees for funders, researchers, and industry looking for advances in evidence generation are still arguing, “imagine if we could do this,” when OHDSI has already done it. And there is much still to be done.
OHDSI Mission
To improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care.

OHDSI Vision
A world in which observational research produces a comprehensive understanding of health and disease.

OHDSI Values

Innovation: Observational research is a field which will benefit greatly from disruptive thinking. We actively seek and encourage fresh methodological approaches in our work.

Reproducibility: Accurate, reproducible, and well-calibrated evidence is necessary for health improvement.

Community: Everyone is welcome to actively participate in OHDSI, whether you are a patient, a health professional, a researcher, or someone who simply believes in our cause.

Collaboration: We work collectively to prioritize and address the real-world needs of our community’s participants.

Openness: We strive to make all our community’s proceeds open and publicly accessible, including the methods, tools and the evidence that we generate.

Beneficence: We seek to protect the rights of individuals and organizations within our community at all times.

Founded in 2014, OHDSI is a growing collaborative of more than 2,300 researchers across disciplines (including biomedical informatics, epidemiology, statistics, computer science, health policy, clinical sciences), across stakeholders (including academia, industry, government and regulatory authorities, and health providers), and across geographies (including 76 countries and six continents). OHDSI also has established an international distributed data network that applies one open community data standard and collectively contains data for more than 800 million patients around the world, and has produce a suite of open-source software packages that enables the community to translate that data into reliable evidence.

OHDSI collaborates to establish open community data standards, develop open source software, conduct methodological research, and apply best practices across the OHDSI data network to generate clinical evidence. The OHDSI distributed data network is comprised of data partners who standardize their source data through a extract-transform-load (ETL) into the OMOP Common Data Model (CDM) and apply OHDSI open-source tools securely behind their own firewall.

OHDSI network studies involve researchers collaborating to design analyses.

How OHDSI Works

Observational Health Data Sciences and Informatics (OHDSI, pronounced “Odyssey”) strives to promote better health decisions and care through globally standardized health data, continuously developing large-scale analytics and a spirit of collaboration though open science.
with pre-specified protocol and analysis code which can be executed across the OHDSI data network, allowing aggregate summary statistics (but no patient-level data) to be shared and collectively interpreted and disseminated.

OHDSI’s research has been presented across various scientific societies, such as American Medical Informatics Association (AMIA), American Statistics Association (ASA/JSM), and International Society of Pharmacoepidemiology (ISPE), and published in top medical journals, including The Lancet, JAMA, BMJ, PNAS and JAMIA.

Our growing global community is always seeking new collaborators.

Please learn more about OHDSI through this publication and Join The Journey!
The OHDSI community brings together volunteers from around the world to establish open community data standards, develop open-source software, conduct methodological research, and apply scientific best practices to both answer public health questions and generate reliable clinical evidence.

Our community is ALWAYS seeking new collaborators. Do you want to focus on data standards or methodological research? Are you passionate about open-source development or clinical applications? Do you have data that you want to be part of global network studies? Do you want to be part of a global community that truly values the benefits of open science? Add a dot to the map below and JOIN THE JOURNEY!

OHDSI By The Numbers

• 2,367 collaborators
• 74 countries
• 21 time zones
• 6 continents
• 1 community
Organizations Involved With OHDSI

OHDSI is a global community of collaborators. Many of the individuals represent organizations who benefit from their participation in the OHDSI community. OHDSI is proud to collaborate with the more than 400 organizations listed below, and looks forward to other organizations joining the journey as well.

School Of Medicine • Inra University Hospital • Innovative Medical Research SA • Inova Health • Institute of Applied Biosciences • Int’l Union Of Health And Welfare • Integraal Kankercentrum Nederland • Intermountain Healthcare • IQUIA • IRIST (Italy) • Istanbul Universitesi • Istanbul University-Cerrahpasa • Janssen R&D • Janssen Scientific Affairs • Jayne Koskinas Ted Giovanis Foundation • Jiangxi Province • Johns Hopkins University • Johnson & Johnson • Juntendo University • Kagwon National University Hospital • Karolinska Institutet • Keck Medicine (USC) • Khoek Teck Puat Hospital • Ki Research Institute • King Saud University Medical City • King’s College London • Klinikko-Bolnički Centar Zvezdara • Knight Cancer Institute • Korkuk University Hospital • Koryung Hospital • Korea Advanced Inst of Sci and Tech • Korea University Anam Hospital • Korea University Ansan Hospital • Korea University Guro Hospital • Kyoto University • Kyunghee University Hospital • Kyunghee Medical Center • Kyungpook National University Hospital • Kyushu University Hospital • Leeds Teaching Hospitals NHS Trust • Ledenon MC • Lille (Luxembourg) • Loyola University (NOLA) • LTS Computing LLC • Lundbeck • Lynxcore Clinical Informatics NV • MGEMN • MameHealth • Marina Salud S.A. • General Brigham • Mayo Clinic • MDV (Japan) • Medaman BV • mellibloc • Medizir • Microsoft • MIT • MITre • Momenti AD • Montefiore/AECOM • MSU Urban Research Center • MSFP-gGmbH • MSKCC • MSU (MT) • MTPPI • MU Vienna • MUSC • HSSC • Myongji Hospital • Nanfang Hospital • National Cancer Center • National Cancer Hospital East • National Health Insurance Corporation Isan Hospital • National Institute of Public Health (Japan) • National University Hospital (SG.NUH) • NCOA • NEMUS • NIH RED • Northwesleigh • Northwell Health • Northwestern Med • Novartis • Novo Nordisk Inc • NYU Langone • Odysseus Data Services • OHDSI • Okayama University • Oklahoma Med • Omnet Patient Care Limited • OU Medical Center • Outcomes Insights • Oxford • Pareto Intelligence • Paxata • Pedianet • PEDSnet • Peking Union Medical College Hospital • Penn State • Physionet • PicnicHealth • Piramamia Hospital District • Plateforme De Données De Santé • Policlinico San Donato S.P.A • Portuguese Institute of Oncology of Porto • Premier Healthcare • PSMAR (Barcelona) • PSSJD • Pusan National University Hospital • Queen Mary University Of London • RCPG (UK) • Regeneron • Regenstrief Institute • Reliant Medical Group • Roche • Rush UMC • Rutgers • RWJ Barnabas • Sage Bionetworks • SAIL Databank • Samsung Seoul Hospital • Sanford Health • Santol • Saudi FDA • SBU (Senegal) • SanoMed • Toulouse • Seoul National University Hospital • Siemens • SERMAS & FIBBAP • Severance Hospital • Shuangshang Hospital • Siemens Health Services • SIMG (Italy) • SNOMED CT • Snowflake • Soonchunhyang University Hospital • Spectrum Health • Spok • St. Luke’s (Idaho) • Stanford University • Stichting Integraal Kankercentrum Nederland • STJIZON • Sydney LH • Taipei Medical University Affiliated Hospital • Taipei Municipal Wanfang Hospital • Takeda • Technical University Sofia • The Hyve • The Roux Institute at Northeastern • The University Court Of The University Of Edinburgh • Tokyo University • Tianjin Anding Hospital • tranSMART • TrialSpark • Tufts • Tulane • U Copenhagen • U Dundee • U Gothenburg • U Hong Kong • U IL Chicago • U Minh • U São Paulo Medical School • U South Africa • U Tartu • U Tsukuba • U Utah • U Witwatersrand • UAB • UBC • UBData • UChemical • UCOLORADO-ANCHUC Medical Camp • UCalgary • UChicago • UCinncinati • UCL (UK) • UCLAS • UCF • UFlorida Health • U Geneva • UHGS (USA) • UNO • University of Iowa • UK Biobank • UK-CRIS • UKentucky • UKER • Ulsan University Hospital • U Mass Memorial MC • UMC New Orleans • UMESSA • University of Miami • UniUE • UniUnio • UniUMichigan • University Of Michigan • University Of Minnesota • University Of Missouri • University Of Mississippi • UNC Chapel Hill • Undidade Local De Saúde De Matsinhos Epe • Université De Bordeaux • Université De Genève • University College London Hospitals NHS Foundation Trust • University of Pécs • UNMC • U Nev Lexem • UNSW Medici • U Pennsylvania • UPittsburgh • U verdict • US Department of Veterans Affairs • US Department of Defense • USP & Drug Administration • US National Cancer Institute • US National Institutes Of Health • US National Library Of Medicine • USAID (US) • UCOSA • UTEXAS Austin • UTEXAS-Houston • UTHealth • UMC • UVirginia • UWashington (Seattle) • UWisconsin-Madison • Valt De’Hbrón Hospital Campus • Vanderbilt • VCU • Veredam • Verige • Vivante Health Software • Vrije Universiteit Amsterdam • Wake Forest • Wanfang Hospital • Washington University • WashJ St Louis • Weill Cornell Medical Center • WHO Uppsala Monitoring Centre • Winship Cancer Institute of Emory University • WMichigan USOM • Wonju Severance Hospital • Wonkwang University Hospital • WU • Yale • Yong • Yoshinai Hospital • ZOL (Belgium) • Z ASSOCIATES

OHDSI.org
OHDSI Collaborators

Testimonials From The OHDSI Community

I started working for Janssen in 2015 and within my first few months of being hired I had submitted my first abstract to the OHDSI Symposium held that year. Since that time I have found incredible support in the community and I have grown in ways I never thought possible thanks to the many friends and collaborators I have met throughout my journey.

As a member of this collaborative I am constantly in awe of the quality of work that’s being produced. I am extremely proud to be a part of this community and every day I aspire to bring my best effort to the table.

Clair Blacketer
Associate Director, Observational Health Data Analytics - Janssen R&D

OHDSI is a rare place where everyone really rolls up their sleeves. It’s easy to talk, but doing takes energy and dedication. Time and again I’ve seen the community rally around supporting a need and turn it into something amazing. I think what makes OHDSI the right environment is the mission. We all want to be part of something bigger than ourselves. We all want to see healthcare change for the better. A lot of us will never get the opportunity to be at bedside treating patients. We’re removed from that piece of the equation. OHDSI provides us with a way to collaborate and share our talents to generate evidence that promotes better health decisions and better care.

It’s that commitment to doing things together, not separately, and sharing the bumps and bruises that come with the hard work that makes this the right environment for this work.

Kristin Kostka
Director of the OHDSI Center at the Roux Institute - Northeastern University

What I really like about it is the enormous energy and the true multidisciplinary focus on advancing medical research. If I’m at an OHDSI meeting, of course I’m representing The Hyve and projects we participate in, but I don’t feel like I’m put in a box, unlike other meetings where you are branded as a ‘vendor’ — there’s a genuine interest in helping out each other and what you can bring to the table. The same goes for an OHDSI study-a-thon — you can be in a call for a study team, and you don’t even notice that it’s made up of people from all sorts of backgrounds (epidemiology, medicine, data science, computer science, etc.) and types of organizations (hospitals, academics, industry, etc.). We all focus on obtaining those medical insights and evidence.

Kees van Bochove
Founder - The Hyve

Both personally and professionally, it’s great to see the number of people who care and want to help, and then actually do help and make a difference. There’s always someone else out there who knows the answer and is willing to help.

I personally have learned a lot from the community, thus I want to be able to give that knowledge back to those who haven’t had the opportunity to learn what I’ve learned. I love teaching tutorials. It allows me to help those who are new and want to be part of this community. I’m always inspired to find new ways of reaching out to more people so that they can also join our community.

Mui Van Zandt
Senior Director, OMOP Data Networks - IOVIA

2020 was the year of OHDSI for me. I’ve always been fascinated by the idea of replicating observational studies internationally, and the more I heard about the open nature of OHDSI, the more I wanted to be involved. I thoroughly enjoy the way the community deals with issues head on and tirelessly aims to drive forward change. In a year where there was so much uncertainty, I really enjoyed being part of such a dynamic and diverse group of individuals who offer their skills with the aim of improving science.

Jenny Lane
Versus Arthritis Clinical Research Fellow in Orthopaedic Surgery, NDORMS - University of Oxford

The OHDSI community is a source of inspiration for me. Take for example the OHDSI COVID-19 Study-a-thon. We had hundreds of people online, across the globe, contributing their talents and expertise to work on a problem that is impacting us all. I’ve attended a number of OHDSI events and interacted with members of the community that are doing amazing work based on the data standards and tools that are made available. OHDSI has helped me view science as a team sport — no one person can do it by themselves. I’m inspired to develop tools and contribute my talents towards OHDSI’s mission.

Anthony Sena
Associate Director - Observational Health Data Analytics - Janssen R&D
The Titan Awards

To recognize OHDSI collaborators (or collaborating institutions) for their contributions towards OHDSI’s mission, the OHDSI Titan Awards were introduced at the 2018 Symposium.

Annually, community members are invited to nominate individuals or institutions they feel have made significant contributions towards advancing OHDSI’s mission, vision and values. Once nominations are submitted, the OHDSI Titan Award Committee select the award winners, and the honorees are announced at the annual symposium.

The award categories, as well as all previous recipients, are listed here.

**Data Standards**

- 2020 - Clair Blacketer, Janssen Research and Development
  - 2019 - Oncology Workgroup (Michael Gurley, Northwestern University; Rimma Belenkaya, Memorial Sloan Kettering Cancer Center; Robert Miller, CTSI)
- 2018 - Vocabulary team (Christian Reich, IQVIA; Anna Ostropolets, Columbia University; Dmitry Dymshyts, Odysseus Data Services)

**Open-Source Development**

- 2020 - Anthony Sena, Janssen Research and Development
- 2019 - Pavil Grafkin, Odysseus Data Services
- 2018 - Christopher Knoll, Janssen Research and Development

**Methodological Research**

- 2020 - Nicholas Thurin, Université de Bordeaux
- 2019 - Jenna Reps, Janssen Research and Development
- 2018 - Martijn Schuemie, Janssen Research and Development; Marc Suchard, University of California, Los Angeles

**Clinical Applications**

- 2020 - Jenny Lane, University of Oxford
- 2019 - Oxford Study-A-Thon (Dani Prieto-Alhambra, University of Oxford, Edward Burn, University of Oxford, Jamie Weaver, Janssen Research and Development, Ross Williams, Erasmus University Medical Center)
- 2018 - Seng Chan You, Ajou University

**Community Collaboration**

- 2020 - Talita Duarte-Salles, IDIAP-JGol
- 2019 - Andrew Williams, Tufts Medical Center
- 2018 - Kristin Kostka, Deloitte; Mui Van Zandt/IQVIA

**Community Leadership**

- 2020 - Dani Prieto-Alhambra, University of Oxford
- 2019 - Peter Rijnbeek, Erasmus University Medical Center
- 2018 - Rae Woong Park, Ajou University School of Medicine

**Community Support**

- 2020 - Erasmus University Medical Center
- 2019 - James Wiggins, Amazon Web Services
- 2018 - Lee Evans, LTS Computing LLC
OHDSI is proud to collaborate with large community initiatives around the world, to support the adoption of the OMOP Common Data Model and OHDSI tools, and to advance our shared interests in generating reliable evidence.

In 2020, OHDSI was awarded a $10 million contract from the U.S. Food and Drug Administration (FDA) to provide support to the Biologics Effectiveness and Safety (BEST) program, which was launched by the FDA Center for Biologics Evaluation and Research (CBER) in 2017.

The lead research team, primarily comprised of OHDSI personnel from Columbia University, UCLA, Northeastern University and Johns Hopkins University provides support to the BEST system in its mission to conduct safety and effectiveness surveillance of biologic products (vaccines, blood and blood products, tissues and advanced therapeutics).

The European Health Data & Evidence Network (EHDEN) is an IMI 2 consortium which operates in Europe within the Innovative Medicines Initiative. EHDEN was launched to address the current challenges in generating insights and evidence from real-world clinical data at scale, to support patients, clinicians, payers, regulators, governments, and the industry in understanding wellbeing, disease, treatments, outcomes and new therapeutics and devices. As of August 2021, EHDEN has created a network of 98 data partners from 23 different countries which are mapping their data to the OMOP common data model.

EHDEN has also brought together 28 small-to-medium enterprises (SMEs) to receive training and become certified to support mapping to the OMOP Common Data Model, and perform services in the ecosystem.

PIONEER is part of the Innovative Medicine Initiative’s (IMI) “Big Data for Better Outcomes” (BD4BO) umbrella program. The BD4BO mission is to improve health outcomes and healthcare systems in Europe by maximizing the potential of Big Data.

OHDSI collaborated with PIONEER in early 2021 on a five-day study-a-thon that investigated the natural history and outcomes of prostate cancer patients managed with watchful waiting.

Health Level Seven International (HL7) and OHDSI announced a collaboration to address the sharing and use of data in the healthcare and research industries by creating a single common data model on March 1, 2021. The organizations will integrate HL7 Fast Healthcare Interoperability Resources (FHIR) and OHDSI’s Observational Medical Outcomes Partnership (OMOP) common data model to achieve this goal.

The Federated E-Health Big Data for Evidence Renovation Network (FEEDER-NET) project was initiated in 2018 with a $10 million budget from the Ministry of Trade, Industry & Energy of Korea.

The main goal is to build a bio-health Big Data ecosystem, centered around an OMOP CDM-based data network. As of August 2021, the FEEDER-NET network included more than 54 million patients.

The All of Us Research Program is inviting one million people across the U.S. to help build one of the most diverse health databases in history. Researchers will use the data, which is mapped to the OMOP CDM, to learn how our biology, lifestyle, and environment affect health. This may one day help them find ways to treat and prevent disease.

The N3C is a partnership among the NCATS-supported Clinical and Translational Science Awards (CTSA) Program hubs, the National Center for Data to Health (CD2H), and NIGMS-supported Institutional Development Award Networks for Clinical and Translational Research (IDeA-CTR), with overall stewardship by NCATS. Collaborators are contributing and using COVID-19 clinical data, mapped to the OMOP CDM, to answer critical research questions to address the pandemic.
## Collaborative Activities

### OHDSI Working Groups

OHDSI’s central mission is to improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care. We work towards that goal in the areas of data standards, methodological research, open-source analytics development, and clinical applications.

Our 27 Working Groups present opportunities for all community members to find a home for their talents and passions, and make meaningful contributions. We are always looking for new collaborators.

See an area where you want to contribute? Please Join The Journey!

<table>
<thead>
<tr>
<th>Workgroup</th>
<th>Current Participants</th>
<th>Lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATLAS</td>
<td>56</td>
<td>Anthony Sena</td>
</tr>
<tr>
<td>Clinical Trials</td>
<td>111</td>
<td>Mike Hamidi, Lin Zhen</td>
</tr>
<tr>
<td>Common Data Model</td>
<td>261</td>
<td>Clair Blacketer</td>
</tr>
<tr>
<td>Data Quality Dashboard Development</td>
<td>90</td>
<td>Clair Blacketer</td>
</tr>
<tr>
<td>Early-Stage Researchers</td>
<td>44</td>
<td>Faaziha Arshad, Ross Williams</td>
</tr>
<tr>
<td>Education</td>
<td>31</td>
<td>Nigel Hughes</td>
</tr>
<tr>
<td>Electronic Health Record (EHR) ETL</td>
<td>168</td>
<td>Melanie Philofsky</td>
</tr>
<tr>
<td>Geographic Information System (GIS)</td>
<td>58</td>
<td>Robert Miller, Andrew Williams</td>
</tr>
<tr>
<td>HADES (Health Analytics Data-to-Evidence Suite)</td>
<td>120</td>
<td>Martijn Schuemie</td>
</tr>
<tr>
<td>Health Equity</td>
<td>87</td>
<td>Jake Gillberg</td>
</tr>
<tr>
<td>Latin America</td>
<td>15</td>
<td>Jose Posada</td>
</tr>
<tr>
<td>Medical Devices</td>
<td>52</td>
<td>Vojtech Huser, Asiyah Lin</td>
</tr>
<tr>
<td>Natural Language Processing</td>
<td>228</td>
<td>Hua Xu</td>
</tr>
<tr>
<td>OHDSI Asia-Pacific (APAC)</td>
<td>46</td>
<td>Mui Van Zandt</td>
</tr>
<tr>
<td>OHDSI APAC Steering Committee</td>
<td>29</td>
<td>Mui Van Zandt</td>
</tr>
<tr>
<td>Patient-Level Prediction</td>
<td>164</td>
<td>Jenna Reps, Peter Rijnbeek</td>
</tr>
<tr>
<td>Pharmacovigilance Evidence Investigation</td>
<td>48</td>
<td>Rich Boyce, Erica Voss</td>
</tr>
<tr>
<td>Phenotype Development &amp; Evaluation</td>
<td>96</td>
<td>Gowtham Rao</td>
</tr>
<tr>
<td>Patient-Generated Health Data</td>
<td>76</td>
<td>Seng Chan You</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>66</td>
<td>Shilpa Ratwani</td>
</tr>
<tr>
<td>OHDSI Steering Committee</td>
<td>129</td>
<td>Patrick Ryan</td>
</tr>
<tr>
<td>Registry (formerly UK Biobank)</td>
<td>57</td>
<td>Maxim Moinat</td>
</tr>
<tr>
<td>Pharmacovigilance Evidence Investigation</td>
<td>48</td>
<td>Rich Boyce, Erica Voss</td>
</tr>
<tr>
<td>Phenotype Development &amp; Evaluation</td>
<td>96</td>
<td>Gowtham Rao</td>
</tr>
<tr>
<td>Patient-Generated Health Data</td>
<td>76</td>
<td>Seng Chan You</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>66</td>
<td>Shilpa Ratwani</td>
</tr>
<tr>
<td>OHDSI Steering Committee</td>
<td>129</td>
<td>Patrick Ryan</td>
</tr>
<tr>
<td>Registry (formerly UK Biobank)</td>
<td>57</td>
<td>Maxim Moinat</td>
</tr>
<tr>
<td>Pharmacovigilance Evidence Investigation</td>
<td>48</td>
<td>Rich Boyce, Erica Voss</td>
</tr>
<tr>
<td>Phenotype Development &amp; Evaluation</td>
<td>96</td>
<td>Gowtham Rao</td>
</tr>
<tr>
<td>Patient-Generated Health Data</td>
<td>76</td>
<td>Seng Chan You</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>66</td>
<td>Shilpa Ratwani</td>
</tr>
<tr>
<td>OHDSI Steering Committee</td>
<td>129</td>
<td>Patrick Ryan</td>
</tr>
</tbody>
</table>

Our workgroups hold meetings, share files, chat asynchronously and more in the OHDSI Microsoft Teams environment. Collaborators can request access to any workgroup through an online form available on both OHDSI.org and our main OHDSI Microsoft Teams environment.

### OHDSI Regional Chapters

An OHDSI regional chapter represents a group of OHDSI collaborators located in a geographic area who wish to hold local networking events and meetings to address problems specific to their geographic location.

- **Africa**
  - Current Participants: 17
  - Lead: Nega Gebreyesus
- **Australia**
  - Current Participants: 36
  - Lead: Nicole Pratt
- **China**
  - Current Participants: 183
  - Lead: Hua Xu
- **Europe**
  - Current Participants: 135
  - Lead: Peter Rijnbeek
- **Japan**
  - Current Participants: 48
  - Lead: Tatsuo Hiramatsu
- **Korea**
  - Current Participants: 30
  - Lead: Mengling Feng
- **Singapore**
  - Current Participants: 26
  - Lead: Seng Chan You
- **Taiwan**
  - Current Participants: 48
  - Lead: Jason Hsu

OHDSI.org
OHDSI Community Calls

The weekly OHDSI community call is where our global network gathers to share research, discuss various topics around observational health, keep apprised on community updates, and plenty more. Our weekly calls are led by Craig Sachson, and they are both recorded and posted to both OHDSI.org and within our Teams environment.

These pages highlight just a few of the meeting topics from 2021; please check out ohdsi.org/ohdsi-community-calls to learn more about these interactive community gatherings.

COllabOraTiVe aCTiViTies

COllabOraTiVe aCTiViTies

How Can You Join Our Calls?

If you are a part of the OHDSI Teams environment, you will receive a weekly calendar invite that includes the upcoming agenda. If you don’t have access, the link is on our Community Calls page, which features all recordings and updates from past calls. Currently, our meetings are held on Tuesdays at 11 am ET. Learn more at our website!

www.ohdsi.org/ohdsi-community-calls

#JoinTheJourney

#JoinTheJourney

OHDSI.org

OHDSI.org
OHDSI Study-A-Thons & Other Events

How does OHDSI go about empowering a community to collaboratively generate the evidence that promotes better health decisions and better care? We do it by innovating on what it means to do collaborative research.

The premise of the study-a-thon is simple: bring together a diverse group of researchers aligned on a common question and focus together on collaboratively designing research protocols, executing analyses across databases, and interpreting results over an intense but fun-filled few days.

OHDSI collaborators have held multiple study-a-thons on a wide array of topics, including orthopedic surgery, rheumatoid arthritis, colorectal cancer, cardiovascular prediction, prostate cancer, and COVID-19. Each event has demonstrated our collective ability to accomplish in a short time what may be unimaginable alone, and it has provided further reinforcement of the power of community and the value of multi-disciplinary collaboration.
The Book of OHDSI

Thank You To Our Book of OHDSI Contributors

| Hamed Abedtash                      | Brian Christian                        |
| Mustafa Ascha                      | Sergio Eslava                           |
| Mark Beno                          | Sunah Song                              |
| Frank DeFalco                     | Mark Khayter                            |
| Thomas Fältner                    | Hamed Abedtash                          |
| Kristin Koettker                   | Brian Christian                        |
| Sindhooosa Malay                   | Sergio Eslava                           |
| Jose Posada                       | Sunah Song                              |
| Peter Rijnbeek                    | Mark Beno                               |
| Martijn Schuemie                   | Frank DeFalco                           |
| Marc Schuemie                      | Thomas Fältner                          |
| Hal Itani                         | Kristin Koettker                        |
| Miles Safford                     | Sindhooosa Malay                        |
| Lisa Cori                         | Jose Posada                             |
| Matt Piotrowski                    | Peter Rijnbeek                          |
| Andrew Williams                   | Martijn Schuemie                        |
| Rachel Seager                     | Miles Safford                           |
| David Waddell                     | Lisa Cori                               |
| Sarah Seager                      | Matt Piotrowski                         |
| David Waddell                     | Rachel Seager                           |
| Bob Wenning                      | David Waddell                           |
| Sanjiv Srinivas                   | Bob Wenning                             |
| Michael Stavros                   | Sanjiv Srinivas                         |
| Sarah Seager                      | Michael Stavros                         |
| Sarah Seager                      | Michael Stavros                         |
| Sarah Seager                      | Michael Stavros                         |
| Sarah Seager                      | Michael Stavros                         |
| Sarah Seager                      | Michael Stavros                         |
| Sarah Seager                      | Michael Stavros                         |
| Sarah Seager                      | Michael Stavros                         |
| Sarah Seager                      | Michael Stavros                         |
| Sarah Seager                      | Michael Stavros                         |

The Book of OHDSI is available for free online in English, Korean, and Chinese, and can also be purchased through Amazon (all links on OHDSI.org).

What Will You Find In The Book of OHDSI?

<table>
<thead>
<tr>
<th>Preface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1: The OHDSI Community</td>
</tr>
<tr>
<td>1.1 The Journey from Data to Evidence</td>
</tr>
<tr>
<td>1.2 Observational Medical Outcomes Partnership</td>
</tr>
<tr>
<td>1.3 OHDSI as an Open-Science Collaborative</td>
</tr>
<tr>
<td>1.4 OHDSI’s Progress</td>
</tr>
<tr>
<td>1.5 Collaborating in OHDSI</td>
</tr>
<tr>
<td>1.6 Summary</td>
</tr>
<tr>
<td>Chapter 2: Where To Begin</td>
</tr>
<tr>
<td>2.1 Why You Should Join OHDSI</td>
</tr>
<tr>
<td>2.2 Where You Fit In</td>
</tr>
<tr>
<td>2.3 Summary</td>
</tr>
<tr>
<td>Chapter 3: Open Science</td>
</tr>
<tr>
<td>3.1 Open Science</td>
</tr>
<tr>
<td>3.2 Open Science In Action: the Study—a Thorn</td>
</tr>
<tr>
<td>3.3 Open Standards</td>
</tr>
<tr>
<td>3.4 Open Source</td>
</tr>
<tr>
<td>3.5 Open Data</td>
</tr>
<tr>
<td>3.6 Open Disclosure</td>
</tr>
<tr>
<td>3.7 OHDSI and the FAIR Guiding Principles</td>
</tr>
<tr>
<td>3.8 Summary</td>
</tr>
<tr>
<td>Chapter 4: The Common Data Model</td>
</tr>
<tr>
<td>4.1 Design Principles</td>
</tr>
<tr>
<td>4.2 Data Model Conventions</td>
</tr>
<tr>
<td>4.3 CDM Standardized Tables</td>
</tr>
<tr>
<td>4.4 Additional Information</td>
</tr>
<tr>
<td>4.5 Summary</td>
</tr>
<tr>
<td>4.6 Design Exercises</td>
</tr>
<tr>
<td>Chapter 5: Standardized Vocabularies</td>
</tr>
<tr>
<td>5.1 Why Vocabularies, and Why Standardizing</td>
</tr>
<tr>
<td>5.2 Concepts</td>
</tr>
<tr>
<td>5.3 Interoperability</td>
</tr>
<tr>
<td>5.4 Hierarchy</td>
</tr>
<tr>
<td>5.5 Internal Reference Tables</td>
</tr>
<tr>
<td>5.6 Special Situations</td>
</tr>
<tr>
<td>5.7 Summary</td>
</tr>
<tr>
<td>5.8 Exercises</td>
</tr>
<tr>
<td>Chapter 6: Extract Transform Load</td>
</tr>
<tr>
<td>6.1 Introduction</td>
</tr>
<tr>
<td>6.2 Design the ETL</td>
</tr>
<tr>
<td>6.3 Step 2: Create the Load Maps</td>
</tr>
<tr>
<td>6.4 Step 3: Implement the ETL</td>
</tr>
<tr>
<td>6.5 Step 4: Quality Control</td>
</tr>
<tr>
<td>6.6 ETL Conventions and THEMIS</td>
</tr>
<tr>
<td>6.7 ETL Maintenance</td>
</tr>
<tr>
<td>6.8 Final Thoughts on ETL</td>
</tr>
<tr>
<td>6.9 Summary</td>
</tr>
<tr>
<td>6.10 Exercises</td>
</tr>
<tr>
<td>Chapter 7: Data Analytics Use Cases</td>
</tr>
<tr>
<td>7.1 Characterization</td>
</tr>
<tr>
<td>7.2 Population-Level Estimation</td>
</tr>
<tr>
<td>7.3 Patient-Level Prediction</td>
</tr>
<tr>
<td>Chapter 8: OHDSI Analytics Tools</td>
</tr>
<tr>
<td>8.1 Analysis Implementation</td>
</tr>
<tr>
<td>8.2 Analysis Strategies</td>
</tr>
<tr>
<td>8.3 ATLAS</td>
</tr>
<tr>
<td>8.4 Methods Library</td>
</tr>
<tr>
<td>8.5 Development Stagedata</td>
</tr>
<tr>
<td>8.6 Summary</td>
</tr>
<tr>
<td>Chapter 9: SQL and R</td>
</tr>
<tr>
<td>9.1 Software</td>
</tr>
<tr>
<td>9.2 DatabaseConnector</td>
</tr>
<tr>
<td>9.3 Quering the CDM</td>
</tr>
<tr>
<td>9.4 Using the Vocabulary When Querying</td>
</tr>
<tr>
<td>9.5 QueryLibrary</td>
</tr>
<tr>
<td>9.6 Designing a Simple Study</td>
</tr>
<tr>
<td>9.7 Implementing the Study using SQL and R</td>
</tr>
<tr>
<td>9.8 Summary</td>
</tr>
<tr>
<td>9.9 Exercises</td>
</tr>
<tr>
<td>Chapter 10: Defining Cohorts</td>
</tr>
<tr>
<td>10.1 What is a Cohort?</td>
</tr>
<tr>
<td>10.2 Rule-Based Cohort Definitions</td>
</tr>
<tr>
<td>10.3 Concept Sets</td>
</tr>
<tr>
<td>10.4 Periodic Cohort Definitions</td>
</tr>
<tr>
<td>10.5 Cohort Definition Valdity</td>
</tr>
<tr>
<td>10.6 Defining a Cohort for Hypertension</td>
</tr>
<tr>
<td>10.7 Implementing a Cohort using ATLAS</td>
</tr>
<tr>
<td>10.8 Implementing the Cohort using SQL</td>
</tr>
<tr>
<td>10.9 Summary</td>
</tr>
<tr>
<td>10.10 Exercises</td>
</tr>
<tr>
<td>Chapter 11: Characterization</td>
</tr>
<tr>
<td>11.1 Database Level Characterization</td>
</tr>
<tr>
<td>11.2 Cohort Characterization</td>
</tr>
<tr>
<td>11.3 Treatment Pathways</td>
</tr>
<tr>
<td>11.4 Incidence</td>
</tr>
<tr>
<td>11.5 Characterizing Hypertensive Persons</td>
</tr>
<tr>
<td>11.6 Database Characterization in ATLAS</td>
</tr>
<tr>
<td>11.7 Cohort Characterization in ATLAS</td>
</tr>
<tr>
<td>11.8 Cohort Characterization in R</td>
</tr>
<tr>
<td>11.9 Cohort Pathways in ATLAS</td>
</tr>
<tr>
<td>11.10 Incidence Analysis in ATLAS</td>
</tr>
<tr>
<td>11.11 Summary</td>
</tr>
<tr>
<td>11.12 Exercises</td>
</tr>
<tr>
<td>Chapter 12: Population-Level Estimation</td>
</tr>
<tr>
<td>12.1 The Cohort Method Design</td>
</tr>
<tr>
<td>12.2 The Self-Contained Cohort Design</td>
</tr>
<tr>
<td>12.3 The Case-Control Design</td>
</tr>
<tr>
<td>12.4 The Case-Crossover Design</td>
</tr>
<tr>
<td>12.5 The Self-Contained Case Series Design</td>
</tr>
<tr>
<td>12.6 Designing a Hypertension Study</td>
</tr>
<tr>
<td>12.7 Implementing the Study using ATLAS</td>
</tr>
<tr>
<td>12.8 Implementing the Study using R</td>
</tr>
<tr>
<td>12.9 Cohort Pathways in ATLAS</td>
</tr>
<tr>
<td>12.10 Summary</td>
</tr>
<tr>
<td>12.11 Exercises</td>
</tr>
<tr>
<td>Chapter 13: Patient-Level Prediction</td>
</tr>
<tr>
<td>13.1 The Prediction Problem</td>
</tr>
<tr>
<td>13.2 Data Extraction</td>
</tr>
<tr>
<td>13.3 Fitting the Model</td>
</tr>
<tr>
<td>13.4 Evaluating Prediction Models</td>
</tr>
<tr>
<td>13.5 Designing a Patient-Level Prediction Study</td>
</tr>
<tr>
<td>13.6 Implementing the Study in ATLAS</td>
</tr>
</tbody>
</table>

The Book of OHDSI aims to be a central knowledge repository for OHDSI, and it focuses on describing the OHDSI community, OHDSI data standards, and OHDSI tools.

It is intended for both OHDSI newcomers and veterans alike, and aims to be practical, providing the necessary theory and subsequent instructions on how to design and implement your research yourself.

You will learn about the OMOP common data model and standard vocabularies, and how they can be used to standardize an observational healthcare database. You will learn about three analytic use cases for these data: characterization, population-level estimation, and patient-level prediction. You will read about OHDSI’s open-source tools and how they can be applied to your data and how you can design and implement your own analyses following OHDSI’s best practices.

Chapters on data quality, clinical validity, software validity, and method quality will explain how to establish the quality of the generated evidence. Lastly, you will learn how to use the OHDSI tools to execute these studies in a distributed research network.

The Book of OHDSI is purchased through Amazon (all links on OHDSI.org).
The OHDSI Symposium

There is nothing quite like an OHDSI symposium. Whether it is held in the U.S., Europe or Asia, or even virtually, our community has turned the symposium into a can’t-miss event each year. While we are proud of the scientific contributions we share, there is far more to the symposium that makes it such a special event. Take a look at some images from past symposia, and we hope to all return together in 2022 and celebrate the incredible work we have done together.

Oct. 20, 2015 • Washington, D.C.

Sept. 23-24, 2016 • Washington, D.C.

Oct. 18-20, 2017 • Bethesda, Md.

Mar. 23-24, 2018 • Rotterdam, Neth.

Oct. 20, 2019 • Guangzhou, China

Sept. 15-17, 2019 • Bethesda, Md.

Dec. 12-14, 2019 • Gwangju, Korea

The 2020 Global Symposium (Oct. 18-21), and the first ever Asia-Pacific (APAC) Symposium (Dec. 5-6) were both held virtually due to the pandemic. While we missed being in person, we still shared ideas, learned from each other, and had plenty of fun. A few memories are below!
Collaborator Showcase

A highlight of our annual symposium is the Collaborator Showcase, when members of the community come together to share research and learn from each other. We received a record number of submissions for the 2021 showcase, and that followed a 2020 Symposium that produced more than 100 accepted posters, talks or software demonstrations.

Collaborator showcase research is shared beyond the symposium. OHDSI posts each presentation on both Twitter and LinkedIn as part of the #OHDSISocialShowcase series. Each submission since 2019 is also posted on OHDSI.org.

2020 Showcase Awards

The community votes on top awards within OHDSI’s four major categories of research each year. Below are the 2020 honorees.

**Observational Data Standards and Management**

Clinical trial data conventions for the OMOP Common Data Model
(Chris Roeder, Katy Sadowski, Maxim Moinat, Philip Solovyev, Sonia Araujo)

**Methodological Research**

Noisy-Or Risk Allocation: A Probabilistic Model for Attributable Risk Estimation
(Amelia Averitt, Adler Perotte)

**Open-Source Analytics Development**

Large-scale evaluation of treatment effect heterogeneity in hypertension
(Alexandros Rekkas, David Van Klaveren, Peter Rijnbeek)

**Clinical Applications**

OHDSI Alexa Skill for a Personalized COVID-19 Outcomes Risk Calculator
(Lisa Evans)
The EHDEN Academy (academy.ehden.eu) serves as a free, publicly available online educational resource for anyone working in the domain of real-world data and real-world evidence.

Originating in the European Health Data & Evidence Network (EHDEN) IMI2 project, its goal is to build upon the foundations of that project and its collaboration with the OHDSI community.

The EHDEN Academy aims to be a resource for all those who generate and utilize data, work technically with it (e.g. ETL and mapping), and are involved in methodological development and the use of standardized analytical tools.

Current Courses in the EHDEN Academy

- Getting Started
- EHDEN Foundation
- ETL Learning Pathway: Data Partner & SME
- Real World Use Cases
- OHDSI-in-a-Box
- OMOP CDM and Standardised Vocabularies
- ATLAS
- Extract, Transform and Load
- Infrastructure
- R for Patient-Level Prediction
- Population-Level Effect Estimation
- Phenotype Definition, Characterisation and Evaluation
- Characterisation
- Citizen and Patient Group Training
- HTA & RWD
- EHDEN Platform Training
- USAGI
- Drug Utilisation Studies
- HADES
- Data Quality Assessment & Reporting

Courses In Development

- Characterisation
- Citizen and Patient Group Training
- Estimation Library I HADES
- Data Quality Assessment & Reporting

The European Health Data & Evidence Network (EHDEN) aspires to be the trusted observational research ecosystem to enable better health decisions, outcomes and care.

Its mission is to provide a new paradigm for the discovery and analysis of health data in Europe, by building a large-scale, federated network of data sources standardized to the OMOP common data model.

As of the summer of 2021, EHDEN has built a federated network of 98 data partners from across 23 European nations, and has trained 28 small-to-medium enterprises to support mapping of this data to OMOP.
OMOP Common Data Model

The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) is an open community data standard, designed to standardize the structure and content of observational data and to enable efficient analyses that can produce reliable evidence.

“The OMOP Common Data Model serves as the foundation of all our work in the OHDSI community, and I’m proud that our open community data standard has been so widely adopted and so extensively used to generate reliable evidence.”

- Clair Blacketer
2020 Titan Award for Data Standards recipient

OMOP CDM By The Numbers

37 tables
- 17 to standardize clinical data
- 10 to standardize vocabularies

394 fields
- 193 with _id to standardize identification
- 101 with _concept_id to standardize content
- 43 with _source_value to preserve original data

1 Open Community Data Standard
What does it take to be an OHDSI data partner? Anyone with access to observational data, representing more than 810 million unique patient records, approximately 11% of the population, can be an OHDSI data partner. Together, these databases represent 331 databases, including 284 electronic health records and 28 administrative claims.

### OHDSI Data Partners

An OHDSI data partner is anyone with access to observational data representing more than 810 million unique patient records, approximately 11% of the population. Together, these databases represent 331 databases, including 284 electronic health records and 28 administrative claims.

### Data Standards

- **SCHEMA**
  - **National**
  - **Regional**
  - **International**
- **ODS**
  - **National**
  - **Regional**
- **IHE**
- **HAPI EE**
- **EDS**
- **DPC**
- **ODIN**
- **CDM**
- **CDR**
- **EHR**
- **Claims**
- **CDW**
- **Survey**
- **Other**

### Join the Journey

- **#JoinTheJourney**
- **OHDSI.org**
- **OHDSI.org**

### DaTa sTanDarDs

**331 databases**, including **284 electronic health records** and **28 administrative claims**

### What does it take to be an OHDSI data partner? Anyone with access to observational data, representing more than 810 million unique patient records, approximately 11% of the population, can be an OHDSI data partner. Together, these databases represent 331 databases, including 284 electronic health records and 28 administrative claims.
The OHDSI vocabularies allow organization and standardization of medical terms to be used across the various clinical domains of the OMOP common data model, and enables standardized analytics that leverage the knowledge base when constructing exposure and outcome phenotypes and other features within characterization, population-level effect estimation, and patient-level prediction studies.

This treemap shows all concepts in the OHDSI vocabularies, organized by domain (color) and vocabularies (boxes sized by the number of concepts).

- 9,833,611 concepts
- 3,392,214 standard concepts
- 701,277 classification concepts
- 133 vocabularies
- 40 domains

1 Shared Resource to Enable Data Standards

Want to learn more about the OHDSI vocabularies?
Read: book.ohdsi.org
Download: athena.ohdsi.org
Learn: academy.ehden.edu

“If we really want to achieve global collaboration, we need more than just standardizing data format. We have to establish a shared understanding of data meaning and speak the same language when expressing clinical ideas. The OHDSI vocabularies is a community resource that makes it possible to work to reach this common goal.”

- Christian Reich
2018 Titan Award for Data Standards recipient
The open-source tools that empower OHDSI research are not only available to the community, but they are DEVELOPED by the community. Leaders within our global network, including 2018 Titan Award recipient Martijn Schuemie (pictured), have developed the foundation for OHDSI collaborators to engage in robust, reliable and reproducible observational health research.
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, a study is 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, 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.

The packages interact directly with any observational data in the OMOP Common Data Model, and are designed to support network research across large datasets with millions of patients and billions of observations, as well as smaller populations. HADES scales to enable larger numbers of analyses so that researchers can systematically explore populations and hypotheses across a range of outcomes.

These packages, available on the HADES home page (ohdsi.github.io/Hades), have empowered at least 34 network studies. These include the OHDSI LEGEND study on hypertension, COVID AESI characterization, and many more. All packages are matured to be additionally released on CRAN (The Comprehensive R Archive Network, a public repository for all R users). The HADES community and the breadth of work has expanded as collaboration efforts have matured.

But for success to follow these positive developments, the HADES foundation and team continues to need greater support.

A small portion of the community maintains the set of packages, and one 2021 HADES objective is to diversify the leadership within the ecosystem. There are several ways that OHDSI collaborators can support this critical piece of the puzzle. Developers can contribute by helping develop and test code. Users of the tools can help with testing, user documentation and other training resources. Those with the means can provide financial support to help pay for developers specifically focused on open-source development. Anybody can contribute ideas as part of the HADES workgroup, which meets every second Thursday of the month at noon ET.

Just as every piece of the HADES toolset has added the value of OHDSI, every small contribution from the community can aid the advancement of HADES.

“Open-source development tools within the HADES ecosystem has been critical to our growth and success as a community,” said George Hripcsak, Chair and Vivian Beaumont Allen Professor of Biomedical Informatics at Columbia, the coordinating center for OHDSI. “Martijn and the HADES team have done extraordinary work to put us in position to run observational health studies that make a difference to patients around the world, but we can’t overlook the burden on this small core of our community who have enabled this growth. I believe we have people who are generous with both their time and talents to help take HADES to a sustainable level as we continue to mature as a community.”

We need your support to continue developing, maintaining and testing our open-source software.

How can you help?

- developers can contribute by helping develop and test code
- users of the tools can help with testing, user documentation and other training resources
- those with the means can provide financial support to help pay for developers specifically focused on open-source development
- anybody can contribute ideas as part of the HADES workgroup, which meets every second Thursday of the month at noon ET.

Open-Source Software

HADES — Health Analytics Data-to-Evidence Suite

Our community, and the breadth of work has expanded as collaboration efforts have matured.
ATLAS

ATLAS is a free, publicly available, web-based tool developed by the OHDSI community that facilitates the design and execution of analyses on standardized, patient-level, observational data in the OMOP CDM format.

Enabling A Journey From Data To Evidence

“ATLAS makes it possible for everyone in the OHDSI community to collaboratively design high-quality observational studies and produce reproducible code that can be shared and executed on OMOP CDM databases around the world.”

- Christopher Knoll

2018 Titan Award for Open-Source Development recipient

Want to learn more about ATLAS?
Experience: atlas-demo.ohdsi.org/
Download: github.com/ohdsi/atlas
Read: book.ohdsi.org/
Train: academy.ehden.eu

VII.

Methods Research

This graphic is taken from the Suchard et al study, published in The Lancet, that is featured on page 49.
Empirical Calibration

Methodological research is a foundational aspect of OHDSI work. We seek to evaluate the performance of analytics methods so we understand when they can be appropriately applied and how confident we can be in the reliability of the evidence we generate. This research has provided the empirical evidence to allow OHDSI to establish best practices for the design and implementation of population-level effect estimation, as applied for safety surveillance and comparative effectiveness research.

Negative controls – exposure-outcome pairs with no causal relationship – offer a powerful diagnostic to evaluate the reliability of a population-level effect estimation study. By applying the same method on the same data to a large collection of negative controls, one can determine if there is systematic error in the analysis, whether due to selection bias, confounding, or measurement error. Empirical calibration is a statistical procedure developed by OHDSI collaborators to use the error distribution estimated from negative controls and correct the original study statistics – point estimates, confidence intervals, and p-values – to restore their nominal operating characteristics and allow for a more honest interpretation of what really has been learned from observational data.

LEGEND (Large-scale Evidence Generation and Evaluation across a Network of Databases) applies high-level analytics to perform observational research on hundreds of millions of patient records within OHDSI’s international database network. LEGEND is based on 10 guiding principles that were published in JAMIA (August, 2020) and are listed below.

1. **LEGEND will generate evidence at a large scale.** Instead of answering a single question at a time (eg, the effect of 1 treatment on 1 outcome), LEGEND answers large sets of related questions at once (eg, the effects of many treatments for a disease on many outcomes). **Aim:** Avoids publication bias. Achieves comprehensiveness of results, and allows for an evaluation of the overall coherence and consistency of the generated evidence.

2. **Dissemination of the evidence will not depend on the estimated effects.** All generated evidence is disseminated at once. **Aim:** Avoids publication bias and enhances transparency.

3. **LEGEND will generate evidence using a prespecified analysis design.** At analyses, including the research questions that will be answered, will be decided prior to analysis execution. **Aim:** Avoids P hacking.

4. **LEGEND will generate evidence by consistently applying a systematic process across all research questions.** This principle precludes modification of analyses to obtain a desired answer to any specific question. This does not imply a simple one-size-fits-all process, rather that the logic for modifying an analysis for specific research questions should be explicated and applied systematically. **Aim:** Avoids P hacking and allows for the evaluation of the operating characteristics of this process (Principle 6).

5. **LEGEND will generate evidence using best practices.** LEGEND answers each question using current best practices, including advanced methods to address confounding, such as propensity scores. Specifically, we will not employ suboptimal methods (in terms of bias) to achieve better computational efficiency. **Aim:** Minimizes bias.

6. **LEGEND will include empirical evaluation through the use of control questions.** Every LEGEND study includes control questions. Control questions are questions where the answer is known. These allow for measuring the operating characteristics of our systematic process, including residual bias. We subsequently account for this observed residual bias in our P values, effect estimates, and confidence intervals using empirical calibration. [7,8] **Aim:** Enhances transparency on the uncertainty due to residual bias.

7. **LEGEND will generate evidence using open-source software that is freely available to all.** The analysis software is open to review and evaluation, and is available for replicating analyses down to the smallest detail. **Aim:** Enhances transparency and allows replication.

8. **LEGEND will not be used to evaluate new methods.** Even though the same infrastructure used in LEGEND may also be used to evaluate new causal inference methods, generating clinical evidence should not be performed at the same time as method evaluation. This is a corollary of Principle 5, since a new method that still requires evaluation cannot already be best practice. Also, generating evidence with unproven methods can hamper the interpretability of the clinical results. Note that LEGEND does evaluate how well the methods it uses perform in the specific context of the questions and data used in a LEGEND study (Principle 6). **Aim:** Avoids bias and improves interpretability.

9. **LEGEND will generate evidence across a network of multiple databases.** Multiple heterogeneous databases (different data capture processes, health-care systems, and populations) will be used to generate the evidence to allow an assessment of the replicability of findings across sites. **Aim:** Enhances generalizability and uncovers potential between-site heterogeneity.

10. **LEGEND will maintain data confidentiality; patient-level data will not be shared between sites in the network.** Not sharing data will ensure patient privacy, and comply with local data governance rules. **Aim:** Privacy.
METHODS RESEARCH

LEGEND in Action

LEGEND (Large-scale Evidence Generation and Evaluation Across a Network of Databases) principles have been applied to studying the effects of treatments for depression, hypertension, and COVID-19, and are being applied to Type 2 diabetes. The clinical impact of LEGEND has already been observed, with important evidence that promotes better health decisions published in Lancet, JAMA Internal Medicine, and Hypertension.


Research and Applications

Large-scale evidence generation and evaluation across a network of databases (LEGEND): assessing validity using hypertension as a case study

Martijn J. Schumez, 1, 2 Patrick B. Ryan, 3 Nicole Pratt, 4 Ruixin Chen, 5, 6 Song Chen You, 2 Harlan M. Kravetski, 2, 7 David Madigan, 4 George Hripcsak, 5, 8 and Marc A. Suchard 9, 10

LEGEND Hypertension project used state-of-the-art causal methods to address both observational biases and potential confounders.

The clinical impact of LEGEND has already been observed, with important evidence that promotes better health decisions published in Lancet, JAMA Internal Medicine, and Hypertension.

The paper also reported that non-dihydropyridine calcium channel blockers proved inferior to thiazides from any of five drug classes, including both thiazides and ACE inhibitors. Within the LEGEND project, ACE inhibitors produced both worse cardiovascular outcomes and worse side effects than thiazides.

First-line thiazide new-users experienced three major medical outcomes (heart attack, hospitalization for heart failure, and stroke) at an approximate 15% lower event rate than those who began treatment with an ACE inhibitor. Furthermore, among potential side effects associated with first-line hypertensive drugs, ACE inhibitor new-users experienced a higher rate of 19 potential side effects — and a lower rate of 2 — than thiazide diuretic new-users.

In spite of these differences, the majority of patients from this study who initiated treatment were prescribed ACE inhibitors (48%) over thiazides (17%). The results, however, indicate that over 3,100 major cardiovascular events could potentially have been avoided had those approximately 2.4 million ACE inhibitor new-users chosen a thiazide diuretic instead.

Filling the Evidence Gaps

"The LEGEND project attempts to fill the evidence gaps in treatment choices that randomized controlled trials (RCTs) leave unanswered," said lead author Marc A. Suchard, MD, PhD (University of California, Los Angeles). "We were able to compare all antihypertensive drug classes against each other at a massive scale and in a transparent and reproducible manner to study what patients worry about. Heart attack. Stroke. Heart failure. Drug safety. LEGEND synthesizes real-world evidence to determine how different drug classes impact the people who have to choose between them."

"We did not execute our study to prove one particular drug class was most effective," Suchard added. "Instead, we used the high-level analytics and best practices developed within OHDSI to study all of these drug classes against each other and openly report on all possible comparisons. Researchers can then interpret specific results in the context of their own research questions."

OHDSI researchers believe LEGEND will continue to significantly enhance how real-world evidence is used to study important healthcare questions that impact millions of patients worldwide.

First-Line Thiazide Diuretic Users Experience 15% Fewer Adverse Cardiovascular Outcomes Than ACE Inhibitor Users

The 2017 ACC/AHA guidelines on antihypertensives recommend initiating hypertension (high blood pressure) treatment with prescription medications from any of five drug classes, including both thiazides and ACE inhibitors. Within the LEGEND project, ACE inhibitors produced both worse cardiovascular outcomes and worse side effects than thiazides.

First-line thiazide new-users experienced three major medical outcomes (heart attack, hospitalization for heart failure, and stroke) at an approximate 15% lower event rate than those who began treatment with an ACE inhibitor. Furthermore, among potential side effects associated with first-line hypertensive drugs, ACE inhibitor new-users experienced a higher rate of 19 potential side effects — and a lower rate of 2 — than thiazide diuretic new-users.

In spite of these differences, the majority of patients from this study who initiated treatment were prescribed ACE inhibitors (48%) over thiazides (17%). The results, however, indicate that over 3,100 major cardiovascular events could potentially have been avoided had those approximately 2.4 million ACE inhibitor new-users chosen a thiazide diuretic instead.

Filling the Evidence Gaps

"The LEGEND project attempts to fill the evidence gaps in treatment choices that randomized controlled trials (RCTs) leave unanswered," said lead author Marc A. Suchard, MD, PhD (University of California, Los Angeles). "We were able to compare all antihypertensive drug classes against each other at a massive scale and in a transparent and reproducible manner to study what patients worry about. Heart attack. Stroke. Heart failure. Drug safety. LEGEND synthesizes real-world evidence to determine how different drug classes impact the people who have to choose between them."

"We did not execute our study to prove one particular drug class was most effective," Suchard added. "Instead, we used the high-level analytics and best practices developed within OHDSI to study all of these drug classes against each other and openly report on all possible comparisons. Researchers can then interpret specific results in the context of their own research questions."

OHDSI researchers believe LEGEND will continue to significantly enhance how real-world evidence is used to study important healthcare questions that impact millions of patients worldwide.

First-Line Thiazide Diuretic Users Experience 15% Fewer Adverse Cardiovascular Outcomes Than ACE Inhibitor Users

The 2017 ACC/AHA guidelines on antihypertensives recommend initiating hypertension (high blood pressure) treatment with prescription medications from any of five drug classes, including both thiazides and ACE inhibitors. Within the LEGEND project, ACE inhibitors produced both worse cardiovascular outcomes and worse side effects than thiazides.

First-line thiazide new-users experienced three major medical outcomes (heart attack, hospitalization for heart failure, and stroke) at an approximate 15% lower event rate than those who began treatment with an ACE inhibitor. Furthermore, among potential side effects associated with first-line hypertensive drugs, ACE inhibitor new-users experienced a higher rate of 19 potential side effects — and a lower rate of 2 — than thiazide diuretic new-users.

In spite of these differences, the majority of patients from this study who initiated treatment were prescribed ACE inhibitors (48%) over thiazides (17%). The results, however, indicate that over 3,100 major cardiovascular events could potentially have been avoided had those approximately 2.4 million ACE inhibitor new-users chosen a thiazide diuretic instead.

Filling the Evidence Gaps

"The LEGEND project attempts to fill the evidence gaps in treatment choices that randomized controlled trials (RCTs) leave unanswered," said lead author Marc A. Suchard, MD, PhD (University of California, Los Angeles). "We were able to compare all antihypertensive drug classes against each other at a massive scale and in a transparent and reproducible manner to study what patients worry about. Heart attack. Stroke. Heart failure. Drug safety. LEGEND synthesizes real-world evidence to determine how different drug classes impact the people who have to choose between them."

"We did not execute our study to prove one particular drug class was most effective," Suchard added. "Instead, we used the high-level analytics and best practices developed within OHDSI to study all of these drug classes against each other and openly report on all possible comparisons. Researchers can then interpret specific results in the context of their own research questions."

OHDSI researchers believe LEGEND will continue to significantly enhance how real-world evidence is used to study important healthcare questions that impact millions of patients worldwide.

First-Line Thiazide Diuretic Users Experience 15% Fewer Adverse Cardiovascular Outcomes Than ACE Inhibitor Users

The 2017 ACC/AHA guidelines on antihypertensives recommend initiating hypertension (high blood pressure) treatment with prescription medications from any of five drug classes, including both thiazides and ACE inhibitors. Within the LEGEND project, ACE inhibitors produced both worse cardiovascular outcomes and worse side effects than thiazides.

First-line thiazide new-users experienced three major medical outcomes (heart attack, hospitalization for heart failure, and stroke) at an approximate 15% lower event rate than those who began treatment with an ACE inhibitor. Furthermore, among potential side effects associated with first-line hypertensive drugs, ACE inhibitor new-users experienced a higher rate of 19 potential side effects — and a lower rate of 2 — than thiazide diuretic new-users.

In spite of these differences, the majority of patients from this study who initiated treatment were prescribed ACE inhibitors (48%) over thiazides (17%). The results, however, indicate that over 3,100 major cardiovascular events could potentially have been avoided had those approximately 2.4 million ACE inhibitor new-users chosen a thiazide diuretic instead.

Filling the Evidence Gaps

"The LEGEND project attempts to fill the evidence gaps in treatment choices that randomized controlled trials (RCTs) leave unanswered," said lead author Marc A. Suchard, MD, PhD (University of California, Los Angeles). "We were able to compare all antihypertensive drug classes against each other at a massive scale and in a transparent and reproducible manner to study what patients worry about. Heart attack. Stroke. Heart failure. Drug safety. LEGEND synthesizes real-world evidence to determine how different drug classes impact the people who have to choose between them."

"We did not execute our study to prove one particular drug class was most effective," Suchard added. "Instead, we used the high-level analytics and best practices developed within OHDSI to study all of these drug classes against each other and openly report on all possible comparisons. Researchers can then interpret specific results in the context of their own research questions."

OHDSI researchers believe LEGEND will continue to significantly enhance how real-world evidence is used to study important healthcare questions that impact millions of patients worldwide.

First-Line Thiazide Diuretic Users Experience 15% Fewer Adverse Cardiovascular Outcomes Than ACE Inhibitor Users

The 2017 ACC/AHA guidelines on antihypertensives recommend initiating hypertension (high blood pressure) treatment with prescription medications from any of five drug classes, including both thiazides and ACE inhibitors. Within the LEGEND project, ACE inhibitors produced both worse cardiovascular outcomes and worse side effects than thiazides.

First-line thiazide new-users experienced three major medical outcomes (heart attack, hospitalization for heart failure, and stroke) at an approximate 15% lower event rate than those who began treatment with an ACE inhibitor. Furthermore, among potential side effects associated with first-line hypertensive drugs, ACE inhibitor new-users experienced a higher rate of 19 potential side effects — and a lower rate of 2 — than thiazide diuretic new-users.

In spite of these differences, the majority of patients from this study who initiated treatment were prescribed ACE inhibitors (48%) over thiazides (17%). The results, however, indicate that over 3,100 major cardiovascular events could potentially have been avoided had those approximately 2.4 million ACE inhibitor new-users chosen a thiazide diuretic instead.
"Patient-level prediction can make a huge impact on the way we deliver medicine, but a lot more work is needed to ensure quality models are developed. OHDSI is leading research to establish best practices, answering important questions that will ensure future predictive models generate reliable evidence."

- Jenna Reps
2019 Titan Award for Methodological Research recipient
Members of the OHDSI community have published many papers together. Often, these studies are first showcased at our annual OHDSI Symposia, like the 2019 event pictured here. These events also provide opportunities for networking, which leads to new collaborations, and new collaborations lead to new evidence generation that impacts patients around the world.
Collaborations Within

In this chapter, you will see both the depth and wide range of peer-reviewed publications that our community has produced over the last decade. How has OHDSI accomplished so much in so little time?

We work together.

This graphic highlights just how much our community collaborates to produce high-quality observational research.

Since our community writes many, MANY papers together, this graphic can’t have everybody in the perfect spot. But it clearly shows how the culture of ‘we’ over ‘me’ has powered OHDSI to incredible heights.

Our OHDSI Community

• Each dot is an OHDSI collaborator with at least 2 OHDSI papers, which include studies involving OMOP
• Size of the dot indicates the number of OHDSI/OMOP papers
• The color indicates the first year someone wrote an OHDSI paper (see legend below)
• A line means two authors were on the same paper. The darker the color of the line, the more papers they co-authored
• The layout is based on co-authorships, so people who collaborated more end up close together in the graph


suspected adverse drug reactions, and observational data from large health-care databases.

How Confident Are We about Observational Findings in Large-scale Evidence Generation and Evaluation across a Network of Databases (LEGEND).

Validation of models predicting stroke in female patients newly diagnosed with atrial fibrillation at scale in the OHDSI network.


Deep phenotyping of 34,128 adult patients hospitalised with COVID-19 in an international network

The Counterfactual χ-GAN: Finding comparable cohorts in observational health data.

Establishment and evaluation of a multicenter collaborative prediction model construction framework supporting OHDSI publications

Deep Learning Approach to Parse Eligibility Criteria in Dietary Supplements Clinical Trials

Large-scale Evidence Generation and Evaluation across a Network of Databases (LEGEND).

The European medical information framework: A novel ecosystem for sharing healthcare data across Europe.

A novel ecosystem for sharing healthcare data across Europe. Learn Health Syst.

Building an international data network: Validation of models predicting stroke in female patients newly diagnosed with atrial fibrillation at scale in the OHDSI network.

Phenotyping of major symptomatic hemorrhagic transformation in acute ischemic stroke at scale in the OHDSI network.

This does not alter our adherence to PLOS ONE policies on sharing data and materials. Prediction Study

De-identification and filtering immune-related adverse events signal based on deep learning.

Comparison of approaches for de-identification of health data: A systematic review.

Deep Learning Approach to Parse Eligibility Criteria in Dietary Supplements Clinical Trials

Design for a Modular Clinical Trial Recruitment Support System Based on FHIR and OMOP.

One such comparator is thegratis cmk project (CMK-GP). The CMK-GP project is funded by the Swiss National Science Foundation (SNSF) under Grant No. 330030_172864/1. This work was supported by the Swiss National Science Foundation (SNSF) under Grant No. 330030_172864/1. The views expressed in this article reflect those of the authors and do not necessarily represent those of the Swiss National Science Foundation or any of the other funding bodies.

The Longitudinal Epidemiologic

How Confident Are We about Observational Findings in Large-scale Evidence Generation and Evaluation across a Network of Databases (LEGEND).
IX. COVID-19 Contributions

This timeline represents only some of OHDSI's global efforts in response of the global pandemic between March 2020 and March 2021.

OHDSI Covid-19 Study-A-Thon begins

CHARYBDIS package released; first CHARYBDIS results available

EMA references two OHDSI studies in ENCePP guide on best practices

OHDSI obtains COVID-19 Therapeutics Accelerator grant towards global research on COVID-19 treatments

Project CHARYBDIS, SCYLLA, more COVID research presented at OHDSI Symposium

AEMAEUS methods study on vaccine safety surveillance begins

AESI study towards monitoring vaccine surveillance opens

Use of repurposed and adjacent drugs in hospital patients with covid-19: a multinational cohort study

Characterising the background incidence rates of adverse events of special interest for covid-19 vaccines in eight countries: a multinational cohort study

OHDSI Covid-19

OHDSI obtains COVID-19 Therapeutics Accelerator grant towards global research on COVID-19 treatments

Project CHARYBDIS, SCYLLA, more COVID research presented at OHDSI Symposium

AEMAEUS methods study on vaccine safety surveillance begins

AEMAEUS study towards monitoring vaccine surveillance opens

This timeline represents only some of OHDSI's global efforts in response of the global pandemic between March 2020 and March 2021.

Use of repurposed and adjacent drugs in hospital patients with COVID-19: A multinational cohort study

Characterising the background incidence rates of adverse events of special interest for COVID-19 vaccines in eight countries: a multinational cohort study
The time was originally meant for highlighting OHDSI capabilities, not testing them. The hours were meant for sharing global research, not sharing in global research. The Observational Health Data Sciences and Informatics (OHDSI) community held a COVID-19 global, virtual study-a-thon March 26-29, 2020, believing that a network of people who valued both collaboration and open science could make a meaningful impact on the current global pandemic.

How? Nobody was quite sure in the moment, but they were confident they would figure it out. “We chose an ambitious path and relied on our community and infrastructure to lead the way,” said Patrick Ryan, Vice President of Observational Health Data Analytics at Janssen Research and Development. “In simple terms, efforts within our community over the past 88 months set the foundation for OHDSI’s most important and impactful 88 hours.”

The Observational Health Data Sciences and Informatics (OHDSI) community, by definition, is a multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics. In plainer terms, it’s a community of people who volunteer their time and talents for the shared goal of improving healthcare through observational research. A global network of OHDSI colleagues planned to celebrate recent research initiatives and discuss future efforts during the annual European Symposium at Oxford University in January. But that session was canceled due to the rapidly spreading COVID-19 virus; in its place, the organizing committee planned a study-a-thon, which OHDSI has experienced significant success with several times over. The twists?

The COVID-19 data was limited (a significant issue for observational health data science). He hosted the 2018 OHDSI European Symposium, and is leading the recently created EHDEN consortium, which is building a large-scale, federated network of European data sources for observational health data science. He presented on it during the 2019 U.S. Symposium, led another one in Barcelona to focus on rheumatoid arthritis and volunteered to host the global community for the 2020 European Symposium.

“We were thrilled to bring the OHDSI community to Oxford, and we were excited about some new aspects, including new tutorials,” Prieto-Alhambra said. “It was crushing to cancel it in the moment, but we quickly looked ahead and saw an opportunity to make the most of our time and talents. From that moment, we never looked back.”

88 hours.

That was the time between the global kickoff and closing calls, both of which have combined for more than 2,300 views on YouTube (the entire set of calls plus the final keynote). More than 330 people from at least 30 nations registered to collaborate in the event, offering their services in areas like literature review, protocol development, study execution, etc. Peter Rijnbeek, Associate Professor Health Data Science at the Erasmus University Medical Center in the Netherlands, has a history of bringing together leaders in observational health data science. He led the 2018 OHDSI European Symposium, and is leading the recently created EHDEN consortium, which is building a large-scale, federated network of European data sources for the discovery and generation of real-world evidence. He took a leadership role once again; his Erasmus team set up the Microsoft Teams virtual platform and created 17 different rooms.

Community involvement was sought in suggesting such questions, but a group that truly believes in collaborative open science knew this was a time to reach outside the circle. Stakeholders around the world reached out to open science knew this was a time to reach outside the circle. Stakeholders around the world reached out to open science knew this was a time to reach outside the circle. Stakeholders around the world reached out to OHDSI to run packages against a more robust set of COVID data than anywhere in the United States. A handful of American institutions, including Columbia and Stanford, signed on to provide deidentified COVID data as well. “The data owners chose to donate their data for use in these critical studies simply because they want to help,” kostka said. “They share our belief in the power of the OHDSI community, and because of that trust, we are able to generate the world’s largest observational studies to help inform decision-making in this major public health issue. I think that’s the coolest thing imaginable, and I am so proud to be part of this effort.”

Laying the groundwork was the necessary warmup for the sprint that was to come — and the marathon that would follow. 88 hours.

It began Thursday, March 26, at 7 am in Oxford, as Prieto-Alhambra took a leadership role once again; his Erasmus team set up the Microsoft Teams virtual platform and created 17 different rooms, each of which worked with OHDSI to run packages against a more robust set of COVID data than anywhere in the United States. A handful of American institutions, including Columbia and Stanford, signed on to provide deidentified COVID data as well. “The data owners chose to donate their data for use in these critical studies simply because they want to help,” kostka said. “They share our belief in the power of the OHDSI community, and because of that trust, we are able to generate the world’s largest observational studies to help inform decision-making in this major public health issue. I think that’s the coolest thing imaginable, and I am so proud to be part of this effort.”

88 hours.

It began Thursday, March 26, at 7 am in Oxford, as Prieto-Alhambra took a leadership role once again; his Erasmus team set up the Microsoft Teams virtual platform and created 17 different rooms, each of which worked with OHDSI to run packages against a more robust set of COVID data than anywhere in the United States. A handful of American institutions, including Columbia and Stanford, signed on to provide deidentified COVID data as well. “The data owners chose to donate their data for use in these critical studies simply because they want to help,” kostka said. “They share our belief in the power of the OHDSI community, and because of that trust, we are able to generate the world’s largest observational studies to help inform decision-making in this major public health issue. I think that’s the coolest thing imaginable, and I am so proud to be part of this effort.”

Laying the groundwork was the necessary warmup for the sprint that was to come — and the marathon that would follow.

88 hours.

It began Thursday, March 26, at 7 am in Oxford, as Prieto-Alhambra took a leadership role once again; his Erasmus team set up the Microsoft Teams virtual platform and created 17 different rooms, each of which worked with OHDSI to run packages against a more robust set of COVID data than anywhere in the United States. A handful of American institutions, including Columbia and Stanford, signed on to provide deidentified COVID data as well. “The data owners chose to donate their data for use in these critical studies simply because they want to help,” kostka said. “They share our belief in the power of the OHDSI community, and because of that trust, we are able to generate the world’s largest observational studies to help inform decision-making in this major public health issue. I think that’s the coolest thing imaginable, and I am so proud to be part of this effort.”

Laying the groundwork was the necessary warmup for the sprint that was to come — and the marathon that would follow. 88 hours.

It began Thursday, March 26, at 7 am in Oxford, as Prieto-Alhambra took a leadership role once again; his Erasmus team set up the Microsoft Teams virtual platform and created 17 different rooms, each of which worked with OHDSI to run packages against a more robust set of COVID data than anywhere in the United States. A handful of American institutions, including Columbia and Stanford, signed on to provide deidentified COVID data as well. “The data owners chose to donate their data for use in these critical studies simply because they want to help,” kostka said. “They share our belief in the power of the OHDSI community, and because of that trust, we are able to generate the world’s largest observational studies to help inform decision-making in this major public health issue. I think that’s the coolest thing imaginable, and I am so proud to be part of this effort.”

Laying the groundwork was the necessary warmup for the sprint that was to come — and the marathon that would follow.

88 hours.

It began Thursday, March 26, at 7 am in Oxford, as Prieto-Alhambra took a leadership role once again; his Erasmus team set up the Microsoft Teams virtual platform and created 17 different rooms, each of which worked with OHDSI to run packages against a more robust set of COVID data than anywhere in the United States. A handful of American institutions, including Columbia and Stanford, signed on to provide deidentified COVID data as well. “The data owners chose to donate their data for use in these critical studies simply because they want to help,” kostka said. “They share our belief in the power of the OHDSI community, and because of that trust, we are able to generate the world’s largest observational studies to help inform decision-making in this major public health issue. I think that’s the coolest thing imaginable, and I am so proud to be part of this effort.”

Laying the groundwork was the necessary warmup for the sprint that was to come — and the marathon that would follow. 88 hours.

It began Thursday, March 26, at 7 am in Oxford, as Prieto-Alhambra took a leadership role once again; his Erasmus team set up the Microsoft Teams virtual platform and created 17 different rooms, each of which worked with OHDSI to run packages against a more robust set of COVID data than anywhere in the United States. A handful of American institutions, including Columbia and Stanford, signed on to provide deidentified COVID data as well. “The data owners chose to donate their data for use in these critical studies simply because they want to help,” kostka said. “They share our belief in the power of the OHDSI community, and because of that trust, we are able to generate the world’s largest observational studies to help inform decision-making in this major public health issue. I think that’s the coolest thing imaginable, and I am so proud to be part of this effort.”

Laying the groundwork was the necessary warmup for the sprint that was to come — and the marathon that would follow. 88 hours.

It began Thursday, March 26, at 7 am in Oxford, as Prieto-Alhambra took a leadership role once again; his Erasmus team set up the Microsoft Teams virtual platform and created 17 different rooms, each of which worked with OHDSI to run packages against a more robust set of COVID data than anywhere in the United States. A handful of American institutions, including Columbia and Stanford, signed on to provide deidentified COVID data as well. “The data owners chose to donate their data for use in these critical studies simply because they want to help,” kostka said. “They share our belief in the power of the OHDSI community, and because of that trust, we are able to generate the world’s largest observational studies to help inform decision-making in this major public health issue. I think that’s the coolest thing imaginable, and I am so proud to be part of this effort.”

Laying the groundwork was the necessary warmup for the sprint that was to come — and the marathon that would follow.
Alhambra welcomed an international community of people to this unique and critical initiative. A panel including Ryan Rijnbeek and George Hripcsak — chair of the Department of Biomedical Informatics at Columbia University, the coordinating center for OHDSI — discussed the long journey from the formation of OHDSI to this moment, and what they believed could be accomplished over four days.

Subgroup calls immediately followed to set the course for their respective work plans. Teams within characterizing, estimation and prediction studies discussed study questions, varied responsibilities, and timetables over the four days; those timetables were dependent on the phenotype group, which had to develop standard cohorts that could be used within all studies.

It was the ultimate team environment. And the clock was now ticking.

88 hours.

Leadership from institutions including Oxford, Erasmus, Columbia, UCLA, Ajou University, Janssen Research and Development, and IQVIA helped put this event in motion, but OHDSI empowers collaborators at different stages of their own journey to make important contributions.

Jennifer Lane, an orthopedic surgeon pursuing her PhD at Oxford, led the literature review efforts and co-authored the manuscript for the largest safety profile on hydroxychloroquine ever executed. Ed Burn, a recent PhD graduate from Oxford, led the characterization team; he had also served as lead author for the Lancet Rheumatology paper on knee replacement.

Ross Williams, Cynthia Yang and Aniek Markus are each PhD students at Erasmus, and they worked on co-authoring a prediction study that could help critical hospitalization and triage decisions healthcare workers are making daily. Anna Ostropolets, a PhD student at Columbia, shared in the leadership of the phenotype team and presented on the 114 validated & reviewed cohorts developed and distributed by the team during the closing call.

Many others within academia contributed to the initiative, while global stakeholders from both industry and healthcare agencies provided critical efforts, ranging from protocol design to data support. “The OHDSI community has an open approach to everything,” said Lane, co-lead author of the hydroxychloroquine study, which had its preprint recently posted on MedRxiv. “It is based upon clear communication, that all contributions are valuable. Everyone is playing to their strengths, which means that the combined effort is precise in many areas that would be incredibly difficult or impossible within one research group or institution. I have met people who will shape the way I work in the future, both through their leadership and their willingness to help me learn novel research approaches.”

Many registrants were newcomers to the OHDSI process who found the idea of a COVID-19 study-a-thon either inspirational and interesting. Their contributions may have been more limited than others over the 88 hours. Some from that group quickly found their footing in the community afterwards and joined studies either developed or brainstormed over the four days.

The 88th hour.

The global closing call was broadcast live to a global audience and provided a series of presentations about how OHDSI arrived at this moment. It was an opportunity to celebrate shared efforts, announce study designs and preliminary findings, and plan for the future.

When Prieto-Alhambra signed off for the final time, COVID-19 did not go away. OHDSI won’t either. The efforts continued immediately. As protocols continue to be designed or improved, data partners work to run studies and generate evidence. The first manuscript was submitted by the team during the closing call.

The real-world evidence we are generating to inform decision-making in this pandemic is the most important thing to come from these four days,” Ryan said. “Reflecting on what a community of volunteers achieved in this collaborative setting is humbling. We had a shared goal that mattered to everybody, but OHDSI has a way of attracting people that enjoy being around. I don’t take that for granted. The people that make up our community are our greatest strength.”

It’s easy to have that positive feeling on Day 1, or as you reach the close, but to have it in the middle of a four-day marathon is a testament to the energy created organically. Friday night chat messages and Saturday morning team calls mattered — in that short a time, it all matters — and maintaining focus and enthusiasm powered the process from start to finish.


- Argentina
- Australia
- Belarus
- Belgium
- Brazil
- Canada
- China
- Colombia
- Croatia
- Denmark
- England
- France
- Germany
- Hungary
- India
- Israel
- Italy
- Netherlands
- New Zealand
- Peru
- Saudi Arabia
- Singapore
- South Korea
- Spain
- Sweden
- Switzerland
- Taiwan
- Ukraine
- UAE
- United States
Lesson Learned Converting Patient-Level Data to the OMOP Common Data Model to Support the COVID-19 Crisis
(Presenter: Erica Voss)

Towards Clinical Data-Driven Eligibility Criteria Optimization for Interventional COVID-19 Clinical Trials
(Presenter: Jaehyun Kim)

Renin-Angiotensin System Blockers and Susceptibility to COVID-19: a Multinational Open Science Cohort Study
(Presenter: Daniel Morales)

OHDSI Alexa Skill for a Personalized COVID-19 Outcomes Risk Calculator
(Presenter: Lisa Evans)

Collaborator Showcase

The Collaborator Showcase is a favorite part of the annual symposia, and our 2020 showcase featured more than a dozen presentations (poster, demo or talk) that focused on studying natural disease history of COVID-19, and it resulted in several published studies.

Project CHARYBDIS
Presented by Talita Duarte-Salles
Project CHARYBDIS (Characterizing Health Associated Risks, and Your Baseline Disease In SARS-COV-2) focused on studying natural disease history of COVID-19, and it resulted in several published studies.

Project SCYLLA
Presented by Daniel Prieto-Alhambra
Project SCYLLA (SARS-Cov-2 Large-scale Longitudinal Analyses) focused on assessing comparative effectiveness and safety among treatments administered during hospitalization and prior to intensive services.

OHDSI Practices Cited In Revision of EMA’s ENCepPP Guide on Methodological Standards

When the European Medicines Agency (EMA) published both the eighth (July 2020) and ninth (July 2021) revisions of The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology, a pair of OHDSI studies were referenced for having informed and supported the EMA's recommendations.

ENCePP noted that “combining data across different databases affords insight into the generalisability of the results and may improve precision if outcomes or exposure of interest are rare or when there is interest in subgroup effects.” The network study led by Jennifer Lane that evaluated the safety profile of hydroxychloroquine, alone and in combination with azithromycin, was specifically cited in this section.

ENCePP also highlighted the critical value of transparency in relation to observational science. The EMA cited the renin-angiotensin system blockers and susceptibility to COVID-19 study, authored by Daniel Morales, for supporting “the reproducibility of their study by publishing the study protocol in the EU PAS Register ahead of time, providing a start-to-finish executable code, facilitating the sharing and exploration of the complete result set with an interactive web application and asking clinicians and epidemiologists to perform a blinded evaluation of propensity score diagnostics for the treatment comparisons.”

EMA/95098/2010 Rev.8

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP)
Guide on Methodological Standards in Pharmacoepidemiology
(Revision 8)

Hydroxychloroquine received significant attention as a potential COVID-19 treatment early in the pandemic. The OHDSI community recognized an insufficient amount of real-world evidence on the safety profile of hydroxychloroquine, so it became an immediate focus during the COVID-19 study-a-thon in late March, 2020.

Lane J CE, et al., observed a significant cardiovascular risk related to the combination of hydroxychloroquine and azithromycin. Shortly after that study was released via preprint, the EMA put out a press release warning of the risks associated with it.

That study, now published in The Lancet Rheumatology (press release below), generated real-world evidence that impacted clinical care, as shown on the next page by a later OHDSI drug utilization study led by Albert Prats-Urice.

Hydroxychloroquine, both alone and in combination with azithromycin, gained early attention during the pandemic as a potential COVID-19 treatment. The short-term (>30 days) safety profile did not identify excess risk in any of the 16 adverse event comparisons as compared to a similar RA drug, sulfasalazine (SSZ). The long-term HCQ therapy was associated with a 66% increase in cardiovascular mortality as compared to SSZ.

The combination of hydroxychloroquine (HCQ) and azithromycin (AZM) has been linked to significant cardiovascular risks, including mortality, in the largest safety study ever performed on both HCQ and HCQ+AZM. This OHDSI network study was led by a large-scale study of its overall safety profile, with a 65% increase in cardiovascular-related mortality, compared to sulfasalazine.

HCQ + AZM had a cardiovascular mortality risk that was more than twice (2.14) as high as the comparative treatment even in the short term based on findings from more than 330,000 years of that combination therapy. This treatment also produced a 15-20% increase of angina/pain and heart failure.

This study, first released on MedRxiv, made significant impacts in the healthcare community. On April 23, 2020, the European Medicines Agency (EMA) cited the study in a warning about the risk of serious side effects with chloroquine and hydroxychloroquine. Two months later, the EMA again highlighted it, among other efforts within the OHDSI community, in its eighth revision of The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCAP) Guide on Methodological Standards in Pharmacoepidemiology.

HCQ, a drug commonly used in the treatment of malaria, lupus and rheumatoid arthritis (RA), gained early attention during the pandemic as a potential COVID-19 treatment. The short-term (>30 days) safety profile did not identify excess risk in any of the 16 adverse event comparisons as compared to a similar RA drug, sulfasalazine (SSZ). The long-term HCQ therapy was associated with a 66% increase in cardiovascular mortality as compared to SSZ.

While there was extensive use of drug repurposing throughout the first 10 months of the COVID-19 pandemic, there was substantial heterogeneity over the types of drugs used for treatment purposes globally. Some drugs, including hydroxychloroquine, saw sharp declines in use, while adjunctive therapies grew into a more relied upon method for patient management.

Often, scientific discovery overturned misconceptions proffered via press conferences and social media.

The OHDSI network study “Use of repurposed and adjuvant drugs in hospital patients with covid-19: multinational network cohort study,” published May 11 by The BMJ, provides a global view of drug utilization in routine practice of more than 330,000 hospitalized patients from China, South Korea, Spain and the United States.

The study highlights the need for future research on the safety and efficacy of the more commonly used treatments.

“The start of the pandemic, when we knew little about COVID-19 and how to treat it, there were many differences between hospitals around the world on how health professionals were treating it,” said study co-lead Albert Prats-Urice, a DPhil candidate and Research Assistant in Clinical Epidemiology at the University of Oxford.

In the campaign in social and traditional media with clearly political intentions,” Prats-Urice said. “This would have taken a long time to counter in the traditional scientific settings. With the work of a community of people around the world producing reliable evidence using observational data, we were able to shift these tendencies and influence decision-making to improve COVID-19 patients.”
Project CHARYBDIS

Within 88 hours of global collaboration through open science, the OHDSI community set the foundation for boundless research possibilities to help inform the response to the deadliest pandemic in more than a century.

“IT takes a village to move the needle,” said Kristin Kostka, a project co-lead and 2018 OHDSI Titan Award recipient. “I use that phrase a lot when it comes to the work we do in OHDSI. It was never more true this year.”

Welcome to CHARYBDIS Village.

Characterizing COVID-19

The CHARYBDIS Project (Characterizing Health Associated Risks, and Your Baseline Disease In SARS-COV-2) had two goals when it was created in the months following the COVID-19 Study-A-Thon in late March:

1) Describe the baseline demographics, clinical characteristics, treatments and symptoms of interest among individuals with COVID-19 overall and stratified by sex, age and other comorbidities;

2) Describe characteristics and outcomes of influenza patients between September 2017 and April 2018 compared to the COVID-19 population.


“Ed’s characterization study is the foundation of how we got to this spot,” Kostka said. “We quickly figured out we needed an overarching frame to put everything into, so a lot of ideas came in could be covered in one protocol.”

It was time to bring in the village.

Early Work

Phenotype development, led by (among many others) Gowtham Rao, Anna Ostropolets, Matthew Spotnitz, Azza Shoaibi, and Patrick Ryan, allowed the CHARYBDIS team to characterize COVID-19 disease natural history by defining diseases and populations of interest so that they could be systematically examined across the OHDSI network. That work carried into the late spring, and coincided with important literature review, led by Lana Lai and Hanieh Razzaghi (again, among others).

Burn and Prats-Uribe worked to develop a code that could generate the most immediate evidence possible on COVID-19, while data partners worked to get their data available to run when the package was available.

Notably, Scott DuVal and Duarte-Salles provided critical leadership with their work around the the IEMAP data, respectively — neither had run an OHDSI network study prior to the pandemic, and now they would provide critical data for its broadest study to date.

By the middle of the summer, more data networks were joining the CHARYBDIS journey, including the first OHDSI study for the University of Washington, and the global community came together to see where they could help inform the COVID-19 response.

“It was a massive work that helped me keep sane during this time by knowing we were helping get information needed to the world, and by collaborating with amazing people and being part of a community,” Prats-Uribe said.

Studies, Studies, Studies

OHDSI collaborators often talk about the inspiration they find in each other. CHARYBDIS meetings, when multiple stakeholders from around the world gathered to discuss their own studies while offering assistance in others, served as great venues for education, inspiration, and the path to generate real-world evidence.

How do you run this many network studies and create robust, reliable and reproducible real-world evidence when the disease itself hadn’t existed a year earlier? Major work went into creating the OHDSI COVID-19 network, which would reach 25 databases from three continents (North America 13, Europe 9 and Asia 3). Within that network, OHDSI collaborators studied:

– more than 16.88 million patients tested for SARS-COV-2
– more than 43,866 obese hospitalized COVID-19 patients.
– more than 25 databases
– more than 52,000 patients diagnosed or loaded positive for COVID-19
– more than 866,000 patients hospitalized with COVID-19

This level of work takes time. Let’s take one for example. The study “Characteristics and outcomes of 63,866 obese hospitalized COVID-19 patients: a community-wide effort to investigate deeper,” published in June, found that obesity is more common amongst COVID-19 than influenza patients, and that obese patients present with more severe forms of COVID-19 with higher hospitalization, intensive services, and fatality than non-obese patients.

The SCYLLA Study

While Project CHARYBDIS studied the natural disease history of COVID-19, the OHDSI community recognized the need for real-world evidence around the different treatments being used around the world. In a world before vaccines, understanding both the safety and effectiveness of these treatments was of critical importance to saving lives.

The SCYLLA (SARS-Cov-2 Large-scale Longitudinal Analyses) Study set out to do that work.

Aided by a grant from the COVID-19 Therapeutics Accelerator, an initiative created by the Bill & Melinda Gates Foundation, Wellcome, and Mastercard, a team of researchers continues to generate evidence to inform the healthcare field in this critical area.

That study, published in July by the International Journal of Obesity was not the first one of its kind to study the impact of obesity around the COVID-19 pandemic. But good luck finding another that includes 207,859 obese patients diagnosed with COVID-19 over three different countries, or 63,866 obese hospitalized COVID-19 patients.

“Moving forward, we are working towards that paper,” Kostka said. “We have more diversity in terms of geography, we have larger sample sizes, we’ve done more curating of reliability of the dataset. The sausage-making may not be exciting, but it’s the OHDSI process that makes the results meaningful.”

Also, top health organizations around the world don’t rely on exciting. The authors of that obesity paper were asked to present their work to the World Health Organization (WHO) European Office, a sign of how these meaningful results were taken seriously by key international organizations.

“IT was an honor to be invited and have the opportunity to present the community work on obesity and COVID-19 at a WHO/Europe expert meeting,” said Duarte-Salles, a 2020 OHDSI Titan Award winner (Community Collaboration). “It is exciting to see the evidence generated in CHARYBDIS being recognized and used by regulatory and public health agencies to help in the design of recommendations to policymakers. I think this is a big accomplishment and we should be very proud of the work we have done as a community this year in the fight against the pandemic.”

Research generated by the HIV study team was presented by a representative from USAID. Regulators have recognized both the clinical and methodological advances made within the OHDSI community over the last year — both in CHARYBDIS and beyond — and those advances are in line with OHDSI’s core mission of improving health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care.
COVID-19 vaccines, as published in BMJ and profiled below.

Largest, Most Extensive Measurement Of Adverse Events Background Rates Can Inform Safety Monitoring Efforts For COVID Vaccines

COVID vaccine surveillance efforts are a global priority, but safety monitoring for vaccines should not reflect a single global population. The largest international network study ever completed on the background rates of adverse events of special interest (AESIs) being monitored in vaccine surveillance efforts identified that these rates vary substantially by age, sex, and database.

There were significant differences in the observed rates of AESIs based on the age groups and sex of more than 136 million people across four continents and 13 total databases in this observational study. Furthermore, differences were observed across people in distinct databases.

This analysis provides historical context for how often outcomes happen in the general population, and can facilitate comparisons with what is observed among those vaccinated. The findings, which suggest caution and adjusted analysis will be needed in vaccine safety analyses to avoid misleading conclusions, can support international efforts aimed at monitoring the safety of COVID vaccines.

“We knew regulators would be monitoring a long list of events for the surveillance of COVID vaccines safety,” said co-senior author Dari Phiyo-Alhambra MD MSc PhD, Professor of Pharmacoepidemiology at the University of Oxford. “To do this, they need robust estimates of the background rates of these events in historical data. These results can be used as benchmarks for the monitoring of these potential safety events and for any upcoming COVID-19 vaccines.”

There were 15 prespecified adverse events studied, matching those being monitored by the U.S. Food and Drug Administration and similar to those used by other top regulatory agencies, which include heart attack, stroke, and blood clotting. Incidence rates were classified by age groupings and gender across the 13 databases, though the outcomes of those specific groupings would even vary by database.

“For example, stroke is a very rare (<1/10,000) outcome for a 24-year-old female, but it was a common (<1/10 to ≥1/100) outcome for an 88-year-old male. The research team believes that the observed and expected rates comparison should also be conducted within the same health database whenever possible,” Li added. “While we understand that is not possible for all surveillance systems or vaccine safety studies, choosing a similar population and stratifying or standardizing by age and sex is highly recommended.”

Heart attack, for example, was observed as a very rare (<1/10,000) outcome for a 24-year-old female, but it was a common (<1/10 to ≥1/100) one for an 88-year-old male. The research team believes that the observed and expected rates comparison should be analyzed separately from populations in much lower risk groups.

“If a vaccinated population is older than an unvaccinated population, and we do not adjust for it, we may see a false increased risk of events following vaccination,” said co-lead author Anna Ostropolets MD, a PhD student and Clarendon scholar at the University of Oxford. “If we compare these rates regardless of age or sex group, we may either find a false signal or neglect a real safety signal while monitoring vaccine surveillance.

“The observed and expected rates comparison should also be conducted within the same health database whenever possible,” Li added. “While we understand that is not possible for all surveillance systems or vaccine safety studies, choosing a similar population and stratifying or standardizing by age and sex is highly recommended.”

COVID-19 COnTribuTiOns

OHDSI Work Around Vaccine Surveillance

The OHDSI community is collaborating with both the European Medicines Agency and the U.S. Food and Drug Administration to assist in monitoring the safety and effectiveness of COVID-19 vaccines.

OHDSI has undertaken a large-scale methodological research experiment to evaluate the performance of methods considered for use in vaccine safety surveillance. The EUMAEUS (Evaluating Use of Methods for Adverse Event Under Surveillance) study has provided a reference benchmark to compare comparative cohort, case-control and self-controlled designs when applied to historical vaccine exposures and negative control outcomes, and has generated results that inform study design for future COVID-19 vaccine surveillance activities.

Another major OHDSI collaborative activity in 2021 was characterizing the background rates of adverse events of special interest, which provides context when evaluating emerging safety data on COVID-19 vaccines, as published in BMJ and profiled below.

X. Join The Journey

Largest, Most Extensive Measurement Of Adverse Events Background Rates Can Inform Safety Monitoring Efforts For COVID Vaccines

OHDSI.org

84 #JoinTheJourney

85 #JoinTheJourney
Cheers, From The OHDSI Community!

2020 threatened to pull people apart, but the OHDSI community came closer together. Volunteer researchers from around the globe joined forces to study COVID-19 and other critical healthcare concerns. Collaboration in the spirit of open science drove us to do far more together than anybody could have done alone.

We also had a lot of fun in the process. To close our 2020 Global Symposium, we created a virtual “cheers” to celebrate our shared successes. To all of you who have done so much for the community, and to those of you who will join our future endeavors, CHEERS!
As a child, I knew I wanted to be an engineer someday. I loved math and science. I loved the pursuit of truth, the satisfaction of solving hard problems and getting the right answer. I learned how to program and got my first taste of statistics during college, and I found out how much fun it was to play with data.

I also grew up thinking that healthcare was the most important sector to work in, because it touches every single person in the world. Some dream of becoming doctors. Many in our community followed that dream and directly impact the lives of their patients every day. I knew that wasn’t my path. I wanted to be an engineer and I wanted to be in healthcare.

I just didn’t know where someone like me could fit, or if I could actually make a difference.

Reflecting back on my own personal journey, I appreciate how challenges that seemed like obstacles actually created opportunities that brought me to where I am today. It was two decades ago when I joined the University of Arizona Arthritis Center right after they had installed their first electronic health records system, and I was challenged to figure out how we could use it for both clinical care and clinical research. A few years thereafter, while working at GlaxoSmithKline, the Chief Medical Officer posed a challenge to me: “When we need to make decisions about the safety of our medicines, we need high quality evidence right away. Isn’t there something more we can do with observational data?”

A while later, he asked me to attend a meeting where leaders from multiple companies lamented how industry and regulators alike were all struggling with the same problems, and I was challenged to consider how collaborative research could be part of the solution.

A logistical challenge to conducting the methodological experiments we dreamed up while planning the Observational Medical Outcomes Partnership was a lack of data standardization, which led to the development of the OMOP CDM. When I joined Janssen R&D, I was challenged to build an analytics team that could respond to the immediate clinical needs of the organization, while also contributing to long-term ambitions of advancing the science of epidemiology.
Every step along the way, I learned, I experimented, I failed, and I persisted. I felt like I was making progress, but also like something was missing.

OHDSI was what I was missing. OHDSI has become a home where my background and skills can allow me to contribute, and where — together with the contributions of others — I feel like I can be a part of making a difference. And I hope it remains a place where everyone — no matter your background, your education, your affiliation, your location — feels belonging and motivation created by legitimate opportunities to have a meaningful impact on health.

When we started OHDSI in 2014, we knew there were hard problems to solve in healthcare and thought that proper analysis of observational healthcare data could be part of the solution. We enjoyed working together, and we figured that if we created a safe space that focused on doing good science, free from bureaucracy and blind to organizational allegiance, that others would enjoy working together too. We valued community and innovation and thought principles of open, reproducible science could be a guiding light. We didn’t know if anyone would join the journey with us, but we wanted to give it a try.

I couldn’t be prouder of how OHDSI has so richly expanded into such a diverse, inclusive and talented community of collaborators all around the world. I am awe of the scientific and technical innovations that continue to be produced year after year, but also of the servant leadership and the willingness of so many to give of themselves for the community. I am gratified by the growth of our collaborative and heartened by the major impacts that newcomers make on a regular basis.

OHDSI has become a place where acquaintances become collaborators, collaborators become friends, and friends become family. The connections we have established are far deeper than any ETL conversion to the CDM, any block of code committed to GitHub or any publication in a journal. It is the shared sense of purpose, a mutual respect and admiration for our collective efforts that makes working in OHDSI humbling and inspiring.

Despite our tremendous progress, the journey is far from over. There remain major challenges that present exciting opportunities. Still today:

- most data from healthcare experiences of patients around the world are captured in a way that makes it challenging to use to inform future care
- most healthcare data that are standardized are not actually used in any analyses
- most analyses using available data are time-consuming and resource-intensive, and yet still may yield unreliable results
- most questions that patients, providers, and policymakers have about the effects of medical interventions remain unanswered
- most health decisions are not informed by reliable evidence, either because the evidence doesn’t yet exist or the evidence is not readily accessible when it is needed

Our future should be one where every health decision can be made confidently together by patients and providers. It should be directly informed by real-time, personalized evidence, guided by the real-world experiences of those patients who came before, and with empirical proof that the evidence is indeed reliable. We need to engineer a learning health system accessible to all stakeholders and make it a commonplace expectation that it be used by everyone to promote better health decision and better care.

The journey from ‘where we are now’ to ‘where we want to be’ might feel overwhelming, like a destination a million miles away.

So what are our next steps together along this journey?

1) We should commit to consistently apply open community data standards within our datasets and across our network, following shared conventions and adopting data quality procedures that assure data are ‘fit for use’ for our evidence needs.
JOIN THE JOURNEY

2) We should support and hold each other accountable for adhering to community best practices for network research, including study pre-specification, open-source and fully reproducible analyses, and transparent reporting of all diagnostics and results. We should continually evolve those best practices through methodological research.

3) We should learn from our successful collaborations during the COVID-19 pandemic, and apply the same sense of urgency to other important public health issues, whether it be applying LEGEND principles to study type 2 diabetes treatments, generating evidence to promote health equity, or tackling other clinical questions raised by the community.

4) We should lean into the notion of evidence-at-scale. We should develop open-source tools which allow us generate characterization results across a wide range of target cohorts and outcomes for questions like ‘how often?’ and train patient-level prediction models at scale to answer ‘what will happen to me?’. We should build an international medical product safety surveillance system that provides all stakeholders access to evidence about the incidence and risk of all outcomes associated with all exposures.

Today, it is possible (and even status quo) for one researcher or team to get access to one dataset and march through one bespoke analysis to test one hypothesis and publish one paper that contributes to the current evidence base. It’s hard, it’s time-consuming, it’s only one drop in a bucket and it may not necessarily be reliable, but it’s possible. Compare that to the possibilities that exist when thousands of researchers collaborate on the world’s largest observational data network, systematically execute scientific best practices through highly efficient analytics tools that allow for simultaneous evaluation of millions of research questions. Imagine the impact that we can have on the lives of the patients we serve: our parents, children, loved ones, neighbors, and friends. That’s also possible, as long as we work together.

‘Join The Journey’ is more than just a catchy hashtag, it truly is a call for collaboration and a call to action. I’m excited to be together with you on this journey, and can’t wait to see what happens next.

Patrick Ryan
Our community has set both the foundation and the highest of standards for global collaboration around observational research. We are making a difference in healthcare, and we are doing it through transparent and reproducible science. We also recognize that there is so much more to be done, so much more that we can do.

If you are inspired by what you read in this book, if you want to learn more about methods research or open-source development, if you have a clinical question you believe needs answering, or if you just want to join a community of people dedicated to the team sport of observational health data sciences and informatics, we have a place ready for you.

How can you get started?

Step 1: Join The OHDSI Forums
Connect with other OHDSI collaborators on our community forums (forums.ohdsi.org) and start discussing how you can help us inform medical decision-making, or simply follow discussions that are interesting to you and learn about the work happening within our global community.

Step 2: Join The OHDSI MS Teams Environment
Collaborate with us globally on our Microsoft Teams environment. We have a main OHDSI team, as well as many others focused on specific workgroups, studies, regional chapters, etc. You can get access to our Teams environment by filling this form out, and then use this form to let us know what workgroups, studies and/or chapters you wish to join. Forms to join are available on ohdsi.org.

Step 3: Join Our Community Calls Or Workgroup Calls
Interact with members of our community weekly during our OHDSI Community Call, held Tuesdays at 11 am ET within the Teams environment. Following weekly updates, we have a variety of call formats, including breakout discussions, research presentations, workgroup updates, and calls dedicated to welcoming newcomers. These calls are recorded, and you can access them (as well as the meeting link) at our Community Calls page (ohdsi.org/ohdsi-community-calls).
Our workgroups meet regularly to discuss a broad variety of specific topics of interest in the community. We keep an updated schedule at our Workgroup Calls page (ohdsi.org/upcoming-working-group-calls), and we invite you to join these calls and collaborate with our community.

Step 4: Continue To Learn About OHDSI
Learn about OHDSI tools and research processes in a variety of ways.

- The Book of OHDSI (book.ohdsi.org) is a community-developed resource with information for every step of your journey. It is also translated into both Chinese and Korean; both are also on our homepage.

- We collaborate with our friends at the EHDEN Consortium (ehden.eu) to develop the EHDEN Academy (academy.ehden.eu), a set of free, on-demand training and development courses. These are open to anybody, and we encourage new OHDSI collaborators to use this resource to learn about best practices towards our mission of improving health by empowering a community to collaboratively generate evidence that promotes better health decisions and better care. Courses are still being developed for the EHDEN Academy.

- Our OHDSI News & Updates (ohdsi.org/ohdsi-news-updates) page keeps you informed of recent publications, upcoming studies and more, while also profiling collaborators and providing other updates about our global efforts.

- Our social platforms provide consistent updates on publications, upcoming meetings, and more, while also highlighting all the work that comes from our collaborator showcase. Follow us on Twitter (@OHDSI) and LinkedIn (search OHDSI), and check out our YouTube site, which is accessible from our homepage and includes all presentations from our weekly calls, symposia, and more.

Your journey with OHDSI has already started. Your interest in this global collaboration is a great step in making a real impact in global health. There is no limit to the contributions you can make in our community. We invite you to search our website, post to the forum, join us in Teams, check out our Github (github.com/OHDSI), or reach out over email (contact@ohdsi.org). Thank you for Joining The Journey with OHDSI!
Join The Journey

@OHDSI  OHDSI.org  OHDSI