Our Journey

Where The OHDSI Community Has Been And Where We Are Going

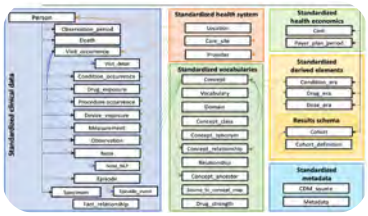
OHDSI

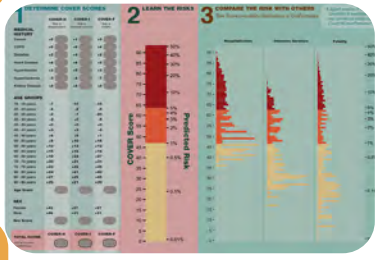
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

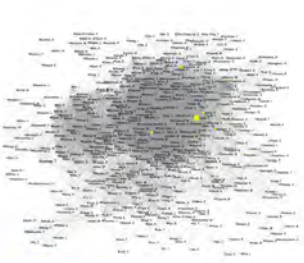
Publication was written and designed by Craig Sachson. Editorial assistance by Patrick Ryan, Kristin Kostka, George Hripcsak, Martijn Schuemie, Marc Suchard, Jody-Ann McLeggon, Jenna Reps, Peter Rijnbeek, Henrik John, Mui Van Zandt.
Photography shared by the OHDSI community unless specifically credited next to image. Printed by ABGPrint.
Thank you to all members of the OHDSI community for all you have done towards improving global healthcare.

Table Of Contents

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







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.

Welcome Letter To The Community

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

Welcome Letter To The Community

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.

George Hripcsak

II.

OHDSI Mission and Values

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.

OHDSI Mission and Values

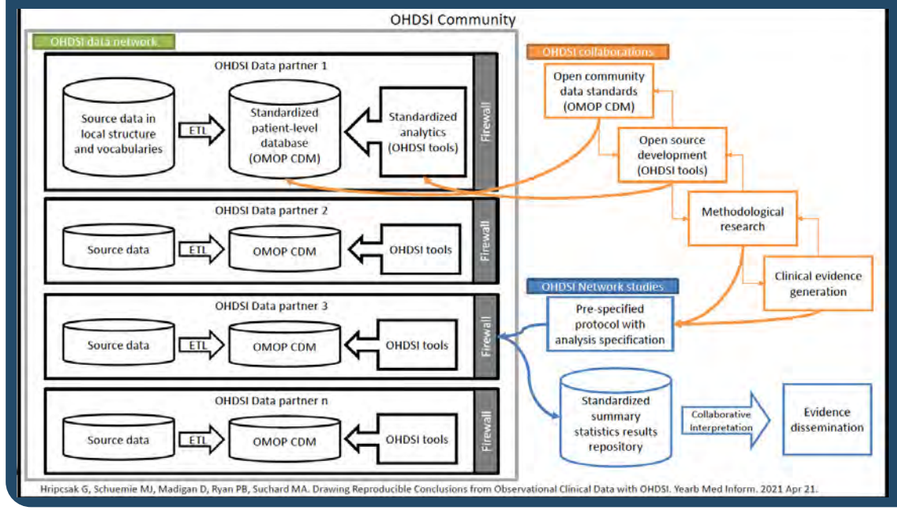
Observational Health Data Sciences and Informatics (OHDSI, pronounced "Odyssey") strives to promote better health decisions and care through globally standardized health data, continuously developing large-scale analytics and a spirit of collaboration though open science.

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



OHDSI Mission and Values


Photo by Odelia Ghodsizadeh/CUIMC

The Department of Biomedical Informatics at Columbia University (DBMI) serves as the coordinating center for the OHDSI community.

Located on the Columbia University Irving Medical Center campus, DBMI is both an academic department and an information services partner to NewYork-Presbyterian Hospital, a major healthcare provider in greater New York.

One of the oldest informatics departments in the nation, faculty and students at DBMI have set the path for design of clinical information systems, methodologies in clinical natural language processing, and machine learning over electronic health record data. Faculty research includes the development and evaluation of innovative information technologies, which has led to enhancements in both health and healthcare.

Both faculty and students work in a highly collaborative environment, applying informatics from the atomic level to global populations.

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!**

III. OHDSI Collaborators

Map of Collaborators

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.

OHDSI By The Numbers

- 2,367 collaborators
- 74 countries
- 21 time zones
- 6 continents
- 1 community

OHDSI COLLABORATORS

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!

#JoinTheJourney 11 OHDSI.org

Organizations Involved With OHDSI

OHDSI is a global community of collaborators. Many of the individuals represent organizations who contribute to and 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.

2Ca-Braga • Aarhus University • AbbVie • Advocate Aurora Health • Agenzia Di Tutela Della Salute Della Provincia Di Bergamo • Ajou University Hospital • Akrivia Health • All Of Us Research Program • Allscripts • AMC Medical Research BV • Amgen • Andrija Štampar School Of Public Health • APDP Diabetes Portugal • Arcadia Inc • ARS Toscana • Asan Hospital • ASCO CancerLinQ • Asociación Instituto De Investigación Sanitaria Biocruces Bizkaia • Assistance Publique - Hopitaux De Paris / Aphp • Assistance Publique Hopitaux De Marseille • Astellas Pharma • AstraZeneca • ASU • AU-EPBRN • AUS Dept of Veterans Affairs • AWS • Az Delta Vzw • Az Klina • Azienda Ospedaliera Nazionale Ss. Antonio E Biagio E Cesare Arrigo Alessandria • Azienda Ospedaliera Universitaria (Aou) Di Modena • Azienda Ospedaliera Universitaria Integrata Verona • Azienda Unità Sanitaria Locale-Irccs In Reggio Emilia • B2I Healthcare • Barts Health NHS Trust • Bayer AG • BCB Medical Oy • Beijing Safe House • Ben-Gurion University • Berlin Institute of Health • Bill & Melinda Gates Foundation • Boehringer Ingelheim • Booz Allen Hamilton • Bordeaux Hospital • Boston Medical Center • Bradford Teaching Hospitals NHS Foundation Trust • Brazilian MOH • Brown University • Bucheon Hospital • Buddhimed Technologies • Caliber • Cancerdatanet Gmbh • Carilion Clinic • Carnegie Melon in Qatar • Case Western CICB • Catholic University of Korea Seoul St. Mary's Hospital • Catholic University of Korea Yeouido St. Mary's Hospital • CDPHP • CEEISCAT (Catalonia) • Cegedim Health Data • Centre Hospitalier Universitaire De Lille • Centre Hospitalier Universitaire De Toulouse • Cerner • Cha University Bundang Medical Center • Charité - Universitätsmedizin Berlin • CHCO (USA) • Cherokee Health Systems • Children's National • CHLA (USA) • Chonnam National University Hospital • CHOP (USA) • CHU Montpellier • Clínica Alemana de Santiago • Clinical Center of Serbia • Clinical Centre of Nis • Cognizant • Columbia University • Columbia University Irving Medical Center • CRHFEI • CSS Denmark • Daegu Catholic University Hospital • Data Integration Centre University Hospital Carl Gustav Carus Dresden • data4life • Databricks • Datasus Ambulatory • DFCI • DHS Los Angeles • DNAnexus • Dongguk University Ilsan Hospital • Dresden University Of Technology • DRG • Drug Safety Research Unit • Duke • Eau Claire Cooperative Health Center • EBMT (EU) • EGCUT • EHDEN • EISBM (Europe) • Eli Lilly & Company • Ephir Inc. • Erasmus MC • European Medicines Agency • Evidera • Evidnet • Ewha Womans University Mokdong Hospital • FIBH12O • FinnGen • Flatiron • Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico • Fondazione IRCCS Istituto Nazionale Dei Tumori • Fondazione IRCCS Policlinico "San Matteo" • Fondazione Poliambulanza • Fred Hutch Cancer Center • Freyr Ltd • Fudan University • Fujitsu • Fundacio Institut D'Inves72tigacions Mèdiques • Fundación Rioja Salud • FUS • GA4GH • Gacheon Gil Hospital • Galilee Medical Center • Gangbuk Samsung Hospital • Gangdong Sacred Heart Hospital • Gangnam Severance Hospital • Geisinger • General Hospital Of Kavala • Georgetown/MedStar Health • Getrude's Children Hospital • Glsmed Learning Health • Google • Great Ormond Street Hospital NHS Foundation Trust • GlaxoSmithKline • Georgia Tech Research Institute • George Washington University • Hanover Medical School (Germany) • Hanyang University Hospital • Harvard • Harvey Walsh Ltd • Hasselt University • HealthVerity • Hebei Mental Health Center • Helix • Helsinki UH CCC Hematology • Hierarchia D.O.O. On Behalf Of University Hospital Centre Zagreb • Health Insurance Review and Assessment Service • HL7 • HM Hospitals • HMAR • Hospital District Of Southwest Finland (Varsinais-Suomen Sairaanhoitopiiri) • Hulafe (Spain) • Hus Datalake Ecareforme Poc • Hwasun Chonnam National University Hospital • IBM T.J. Watson Research Center • Ican School Of Medicine At Mount Sinai • ICON • ICVS (Portugal) • IDIAPJGOL / SIDIAP • Idival • IMASIS • Imperial College Of Science Technology And Medicine • Incheon Sejong Hospital • Indian Society for Clinical Research • Indiana University

School Of Medicine • Inha University Hospital • Innovative Medical Research SA • Inova Health • Institute of Applied Biosciences • Int'l Uni of Health And Welfare • Integraal Kankercentrum Nederland • Intermountain Healthcare • IQVIA • IRST (Italy) • Istanbul Universitesi • Istanbul Universty-Cerrahpasa • Janssen R&D • Janssen Scientific Affairs • Jayne Koskinas Ted Giovanis Foundation • Jiangxi Province • Johns Hopkins University • Johnson & Johnson • Juntendo Uni SOM • Kangwon National University Hospital • Karolinska Institutet • Keck Medicine (USC) • Khoo Teck Puat Hospital • KI Research Institute • King Saud University Medical City • King's College London • Kliničko-Bolnički Centar Zvezdara • Knight Cancer Institute • Konkuk University Hospital • Konyang University 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, Japan • Leeds Teaching Hospitals NHS Trust • Leiden MC • LIH (Luxembourg) • Loyola University (NOLA) • LTS Computing LLC • Lundbeck • Lynxcare Clinical Informatics NV • M2GEN • MaineHealth • Marina Salud S.A. • Mass General Brigham • Mayo Clinic • MDV (Japan) • Medaman BV • mederrata • Medibloc • Merck • Microsoft • MIT • MITRE • Momentum AD • Montefiore/AECOM • MS 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 Ilsan Hospital • National Institute of Public Health (Japan) • National University of Hospital (SG_NUH) • NCQA • Nemours • NHIRD • NICE • Northshore • Northwell Health • Northwestern Med • Novartis • Novo Nordisk Inc. • NYU Langone • Odysseus Data Services • OHSU • Okayama University • Oklahoma U • Optimum Patient Care Limited • OSU Medical Center • Outcomes Insights • Oxford • Pareto Intelligence • Paxata • Pedianet • PEDSnet • Peking Union Medical College Hospital • Penn State • PhysioNet • PicnicHealth • Pirkanmaa 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 • RCGP (UK) • Regeneron • Regenstrief Institute • Reliant Medical Group • Roche • Rush UMC • Rutgers • RWJ Barnabas • Sage Bionetworks • SAIL Databank • Samsung Seoul Hospital • Sanford Health • Sanofi • Saudi FDA • SBU (USA) • Semantix • Semmelweis Egyetem • Seoul National University Bundang Hospital • Seoul National University Hospital • SERMAS & FIIBAP • Severance Hospital • Shuanghe 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 • STIZON • Sydney LHD • 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 Minho • U São Paulo Medical School • U South Australia • U Tartu • U Tsukuba • U Utah • U Witwatersrand • UA-Birmingham • UArkansas • UBuffalo • UColorado Health • UColorado-Anschutz Medical Campus • UCalgary • UChicago • UCincinnati • UCL (UK) • UCLA • UCSF • UFlorida Health • UH Geneva • UHG (USA) • UIO • University of Iowa • UK Biobank • UK-CRIS • UKentucky • UKER • Ulsan University Hospital • U Mass Memorial MC • UMC New Orleans • UMessina • University of Miami • University of Michigan • UMichigan School of Dentistry • University of Minnesota • University of Mississippi MC • UNC Chapel Hill • Unidade Local De Saúde De Matosinhos Epe • Université De Bordeaux • Université De Genève • University College London Hospitals NHS Foundation Trust • University of Pécs • UNMC • UNew Mexico • UNSW Medicine Australia • UPennsylvania • UPittsburgh • URochester • US Department of Veterans Affairs • US Department of Defense • US Food & Drug Administration • US National Cancer Institute • US National Institutes of Health • US National Library of Medicine • USAID • USC (LA) • UTexas-Austin • UTexas-Houston • UTHCS-Houston • UTMC • UVirginia • UWashington (Seattle) • UWisconsin-Madison • Vall D'Hebrón Hospital Campus • Vanderbilt • VCU • Veradigm • Vertex • Vivante Health Software • Vrije Universiteit Amsterdam • Wake Forest • Wanfang Hospital • Washington University • WashU St Louis • Weill Cornell Medical Center • WHO Uppsala Monitoring Centre • Winship Cancer Institute of Emory University • WMichigan USOM • Wonju Severance Hospital • Wonkwang University Hospital • WVU • Yale • Yongin Severance Hospital • Yonsei University • ZOL (Belgium) • ZS Associates

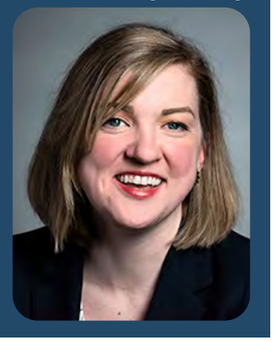
Testimonials From The OHDSI Community

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

OHDSI has six values: Innovation, Reproducibility, Community, Collaboration, Openness, and Beneficence. OHDSI pursues highly reliable evidence through innovative ways and open-science spirit. When I started my journey in 2016, many gurus in OHDSI welcomed and helped me a lot. OHDSI works hard to produce medical evidence for better care for humanity. It is truly impressive that OHDSI is trying to hold its values.

Seng Chan You
Research Assistant Professor
Severance Hospital

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
IQVIA

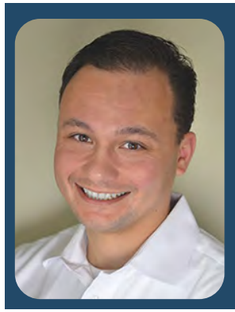
2020 was the year of OHDSI for me. I've always been fascinated by the idea of replicating observational studies internationally, and the more I heard about the open nature of OHDSI, the more I wanted to be involved. I thoroughly enjoy the way the community deals with issues head on and tirelessly aims 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

As a clinical informatician, I'd credit OHDSI for a great deal of the perspective that I am able to bring my own health system in areas of predictive analytics and observational analysis of real world EMR data. I can't understate how important that's been, in terms of my ability to help lead in some of these areas within Alberta Health Services. I've also used OHDSI tools such as ATLAS, Athena, Usagi, and other constructs to help solve various informatics problems for AHS. You don't need a CDM to benefit from the cutting edge work this community does.

As an internist, I've become a better consumer of observational studies, and predictive model studies in particular. That has been very helpful in the COVID era. I was better able to appraise the observational studies that were being published, while OHDSI's own work produced effect estimates that I was able to confidently bring to the bedside. As an example, I was able to cite the work done by this group when justifying restraint regarding hydroxychloroquine in COVID-19 related respiratory disease to patients and their families.

Evan Minty
General Internist and Clinical Assistant Professor O'Brien Institute for Public Health at the University of Calgary



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)

Clair Blacketer

Open-Source Development

2020 - Anthony Sena, Janssen Research and Development

2019 - Pavil Grafkin, Odysseus Data Services

2018 - Christopher Knoll, Janssen Research and Development

Anthony Sena

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

Nicholas Thurin

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

Jenny Lane

Community Collaboration

2020 - Talita Duarte-Salles, IDIAPJGol

2019 - Andrew Williams, Tufts Medical Center

2018 - Kristin Kostka, Deloitte; Mui Van Zandt/IQVIA

Talita Duarte-Salles

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

Daniel Prieto-Alhambra

Community Support

2020 - Erasmus University Medical Center

2019 - James Wiggins, Amazon Web Services

2018 - Lee Evans, LTS Computing LLC

Erasmus University Medical Center

OHDSI + Large Community Initiatives

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.

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.

PIONEER is part of the Innovative Medicine Initiative's (IMI's) "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.

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.

IV.

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!

ATLAS
Current Participants: 56
Lead: Anthony Sena

Clinical Trials
Current Participants: 111
Leads: Mike Hamidi, Lin Zhen

Common Data Model
Current Participants: 261
Lead: Clair Blacketer

Data Quality Dashboard Development
Current Participants: 90
Lead: Clair Blacketer

Early-Stage Researchers
Current Participants: 44
Leads: Faaizah Arshad, Ross Williams

Education
Current Participants: 31
Lead: Nigel Hughes

Electronic Health Record (EHR) ETL
Current Participants: 168
Lead: Melanie Philofsky

Geographic Information System (GIS)
Current Participants: 58
Leads: Robert Miller, Andrew Williams

HADES (Health Analytics Data-to-Evidence Suite)
Current Participants: 120
Lead: Martijn Schuemie

Health Equity
Current Participants: 87
Lead: Jake Gillberg

Latin America
Current Participants: 15
Lead: Jose Posada

Medical Devices
Current Participants: 52
Leads: Vojtech Huser, Asiyah Lin

Natural Language Processing
Current Participants: 228
Lead: Hua Xu

OHDSI Asia-Pacific (APAC)
Current Participants: 46
Lead: Mui Van Zandt

OHDSI APAC Steering Committee
Current Participants: 29
Lead: Mui Van Zandt

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 Steering Committee
Current Participants: 26
Lead: Patrick Ryan

Oncology
Current Participants: 129
Lead: Shilpa Ratwani

Patient-Generated Health Data
Current Participants: 76
Lead: Seng Chan You

Pharmacovigilance Evidence Investigation
Current Participants: 48
Leads: Rich Boyce, Erica Voss

Phenotype Development & Evaluation
Current Participants: 96
Leads: Gowtham Rao

Population-Level Effect Estimation
Current Participants: 164
Lead: Martijn Schuemie, Marc Suchard

Patient-Level Prediction
Current Participants: 164
Lead: Jenna Reps, Peter Rijnbeek

Psychiatry
Current Participants: 66
Lead: Shilpa Ratwani

Registry (formerly UK Biobank)
Current Participants: 57
Lead: Maxim Moinat

Vaccine Safety
Current Participants: 28
Lead: Patrick Ryan

Vaccine Vocabulary
Current Participants: 36
Lead: Adam Black

Women of OHDSI
Current Participants: 97
Lead: Maura Beaton

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: 163
Lead: Hua Xu

Europe
Current Participants: 135
Lead: Peter Rijnbeek

Japan
Current Participants: 19
Lead: Tatsuo Hiramatsu

Korea
Current Participants: 26
Lead: Seng Chan You

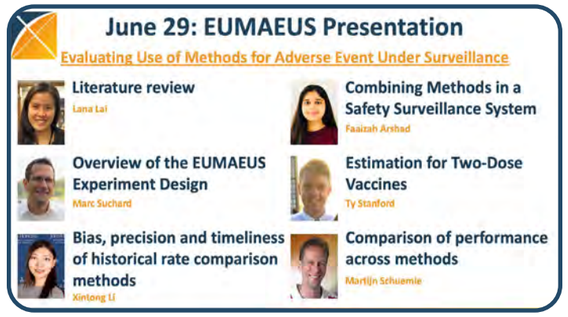
Singapore
Current Participants: 30
Lead: Mengling Feng

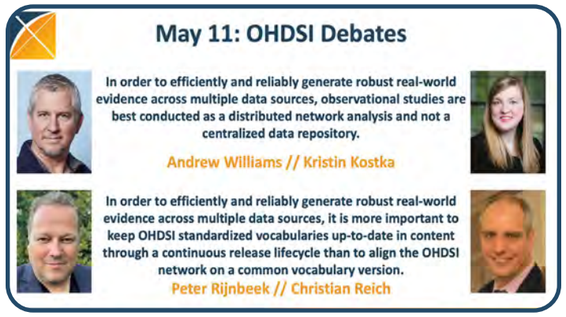
Taiwan
Current Participants: 48
Lead: Jason Hsu

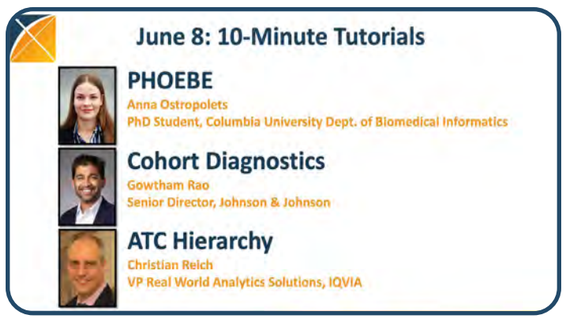
OHDSI Community Calls

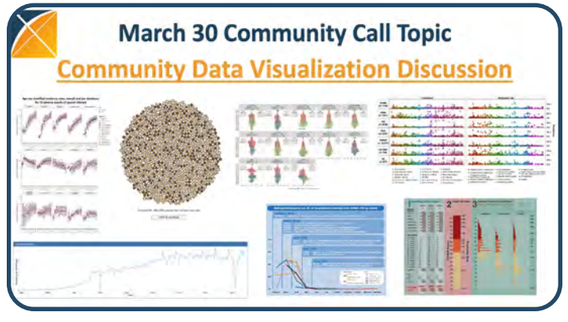
The weekly OHDSI community call is where our global network gathers together to share research, discuss various topics around observational health, keep apprised on community updates, and plenty more. Our weekly calls are led by Craig Sachson, and they are both recorded and posted to both OHDSI.org and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

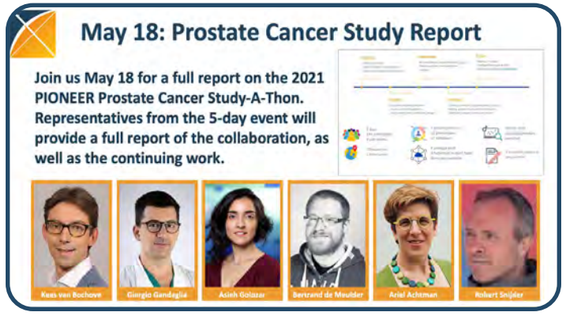
These pages highlight just a few of the meeting topics from 2021; please check out ohdsi.org/ohdsi-community-calls to learn more about these interactive community gatherings.

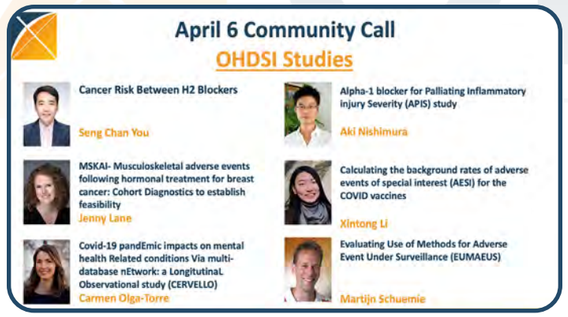


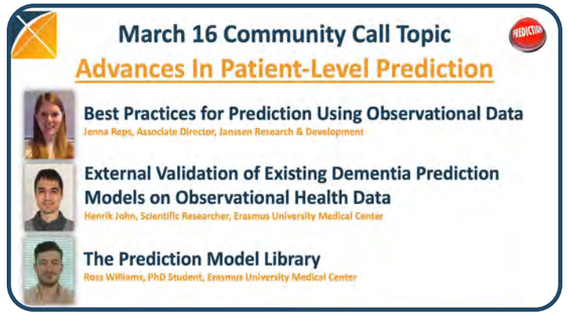










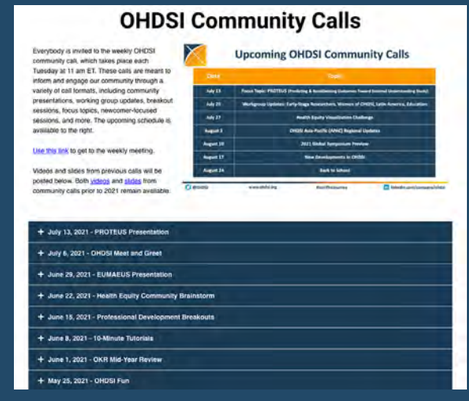


How Can You Join Our Calls?

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



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.

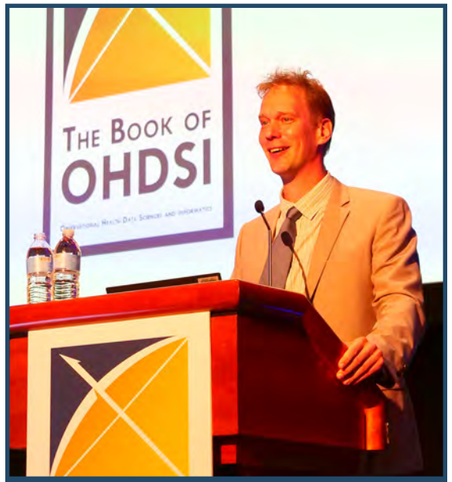
OHDSI collaborators have held multiple study-a-thons on a wide array of topics, including orthopedic surgery, rheumatoid arthritis, colorectal cancer, cardiovascular prediction, prostate cancer, and COVID-19. Each event has demonstrated our collective ability to accomplish in a short time what may be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

The Book of OHDSI

Published in 2019, the Book of OHDSI (book.ohdsi.org) 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 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.



Martijn Schuemie, who co-led the Book of OHDSI development with David Madigan, introduced the book at the 2019 U.S. Symposium.

Members of the OHDSI community collaborated on documentation efforts for the Book of OHDSI at Case Western Reserve Univ. in Cleveland.

Chapters on data quality, clinical validity, software validity, and method validity 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 available for free online in English, Korean and Chinese, and can also be purchased through Amazon (all links on OHDSI.org).

Thank You To Our Book of OHDSI Contributors

Hamed Abedtash	Mustafa Ascha	Mark Beno	Clair Blacketer	David Blatt
Brian Christian	Gino Cloft	Frank DeFalco	Sara Dempster	Jon Duke
Sergio Eslava	Clark Evans	Thomas Falconer	George Hripscak	Vojtech Huser
Mark Khayter	Greg Klebanov	Kristin Kostka	Bob Lanese	Wanda Lattimore
Chun Li	David Madigan	Sindhoosha Malay	Harry Menegay	Akihiko Nishimura
Ellen Palmer	Nirav Patil	Jose Posada	Nicole Pratt	Dani Prieto-Alhambra
Christian Reich	Jenna Reps	Peter Rijnbeek	Patrick Ryan	Craig Sachson
Izzy Saridakis	Paola Saroufim	Martijn Schuemie	Sarah Seager	Anthony Sena
Sunah Song	Matthew Spotnitz	Marc Suchard	Joel Swerdel	Devin Tian
Don Torok	Kees van Bochove	Mui Van Zandt	Erica Voss	Kristin Waite
Mike Warfe	Jamie Weaver	James Wiggins	Andrew Williams	Seng Chan You

What Will You Find in The Book of OHDSI?

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.

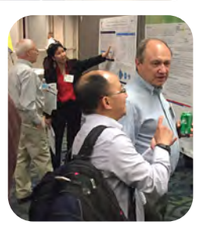
Take a look at some images from past symposia, and we hope to all return together in 2022 and celebrate the incredible work we have done together.

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.

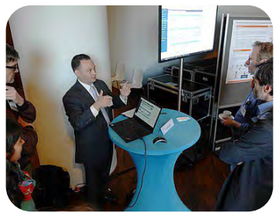



Oct. 11-13, 2018 • Bethesda, Md.




Mar. 29-30, 2019 • Rotterdam, Neth.



Oct. 20, 2019 • Guangzhou, China



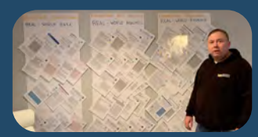
Sept. 15-17, 2019 • Bethesda, Md.




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.

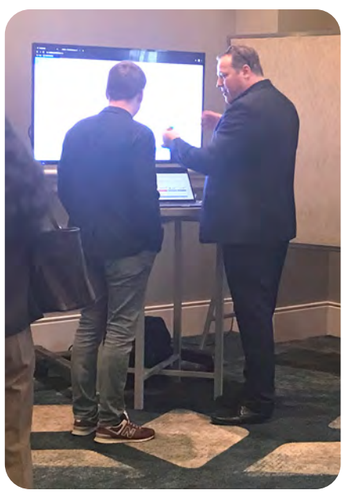
A few memories are below!




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.

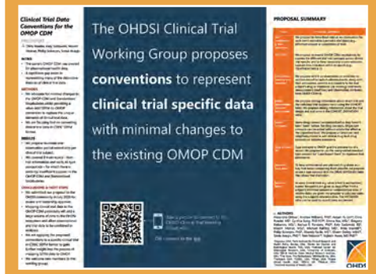
Collaborator showcase research is shared beyond the symposium. OHDSI posts each presentation on both Twitter and LinkedIn as part of the #OHDSISocialShowcase series. Each submission since 2019 is also posted on OHDSI.org.



2020 Showcase Awards

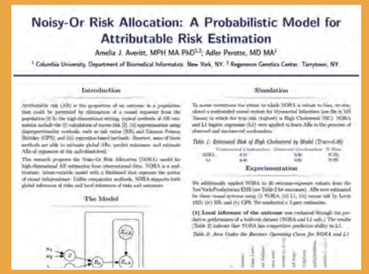
The community votes on top awards within OHDSI's four major categories of research each year. Below are the 2020 honorees.

Observational Data Standards and Management



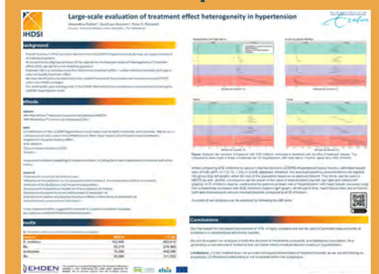
Clinical trial data conventions for the OMOP Common Data Model (Chris Roeder, Katy Sadowski, Maxim Moinat, Philip Solovyev, Sonia Araujo)

Methodological Research



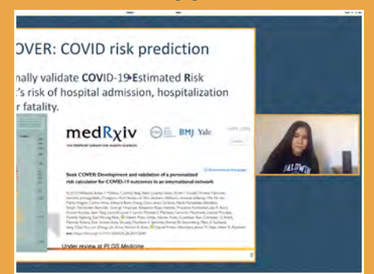
Noisy-Or Risk Allocation: A Probabilistic Model for Attributable Risk Estimation (Amelia Averitt, Adler Perotte)

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

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

Courses In Development

- Characterisation
- Citizen and Patient Group Training
- Estimation Library I HADES
- Data Quality Assessment & Reporting
- HTA & RWD
- EHDEN Platform Training
- USAGI
- Drug Utilisation Studies

EHDEN
EUROPEAN HEALTH DATA & EVIDENCE NETWORK

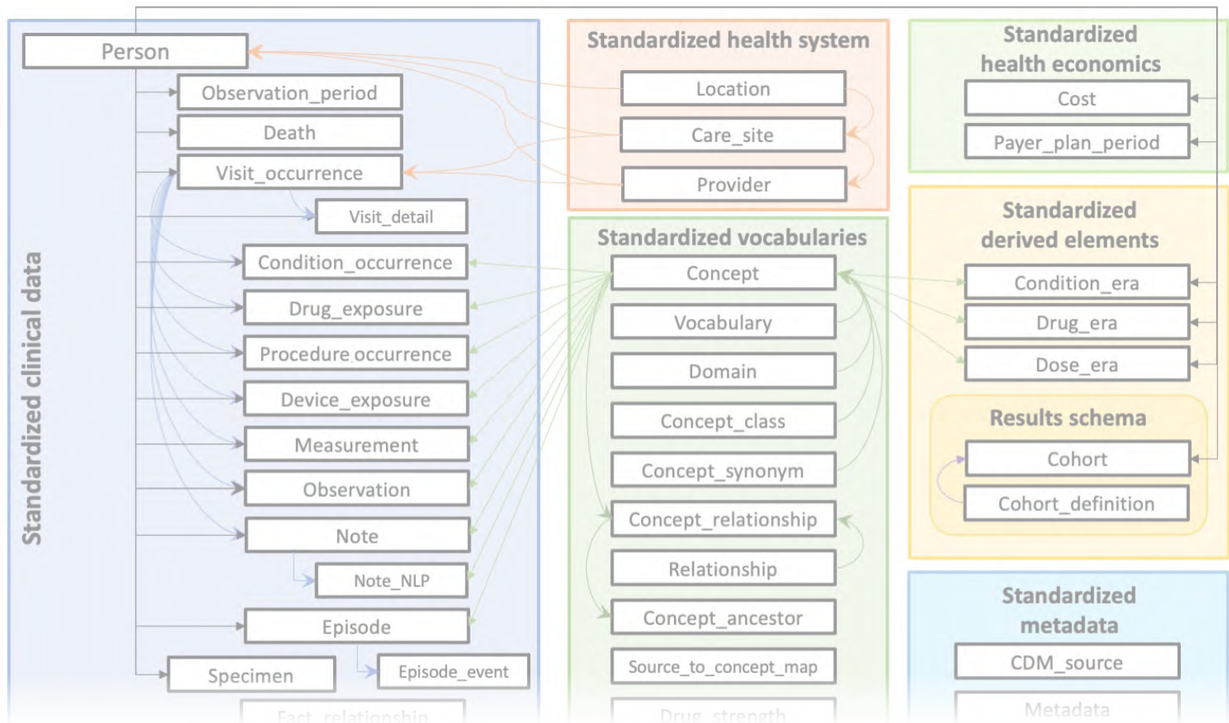
The European Health Data & Evidence Network (EHDEN) aspires to be the trusted observational research ecosystem to enable better health decisions, outcomes and care.

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.

V.

Data Standards



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."

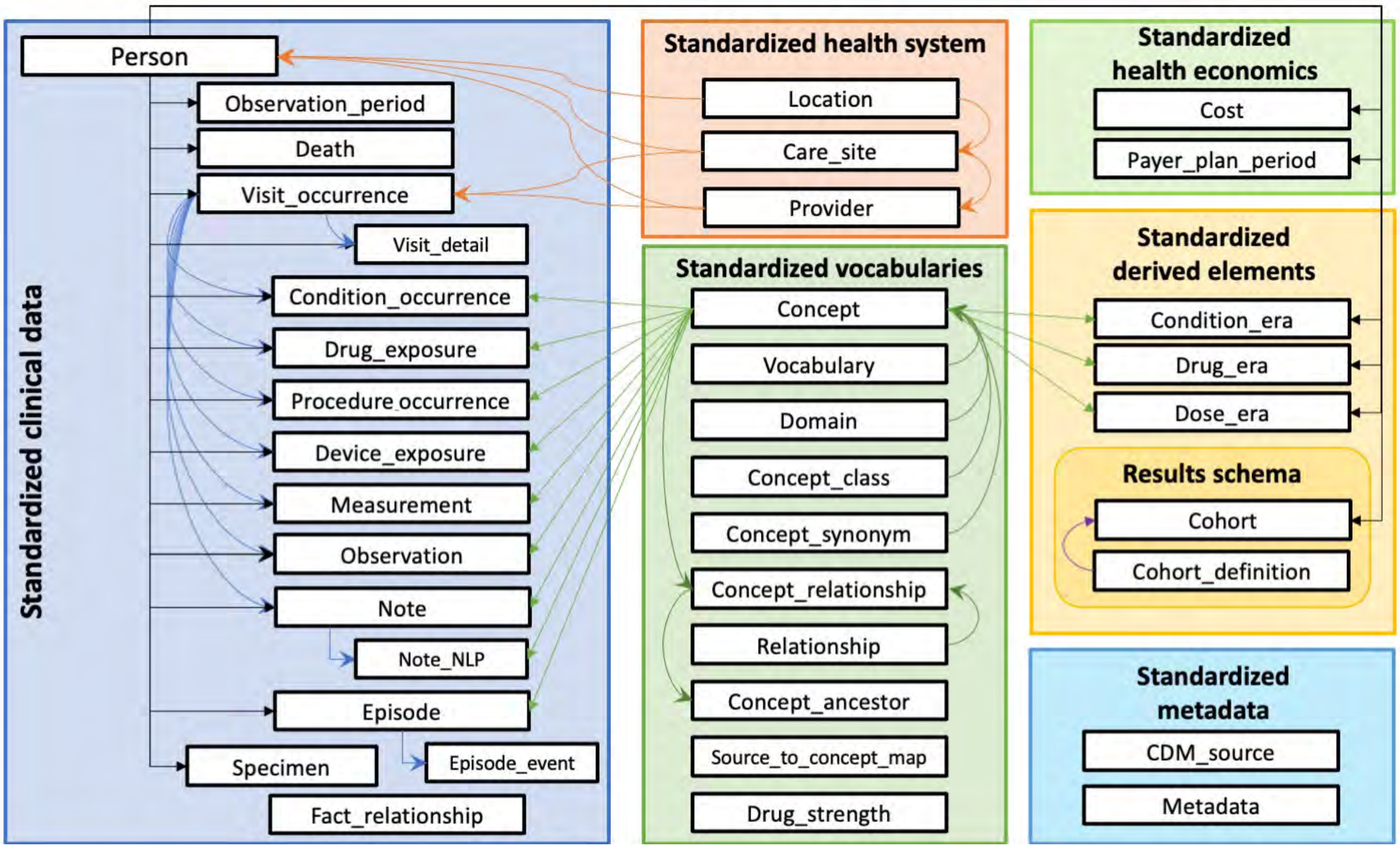
- Clair Blacketer
2020 Titan Award for Data Standards recipient

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



OHDSI Data Partners

What does it take to be an OHDSI data partner? Anyone with access to observational data can standardize their database in the OMOP Common Data Model, apply OHDSI's open-source tools, and participate in collaborative research.

Who has already joined the journey and adopted the OMOP CDM? There are currently 331 databases, including 284 electronic health records and 28 administrative claims sources, that come from 34 different countries. Together, these databases represent more than 810 million unique patient records, approximately 11% of the world's population.

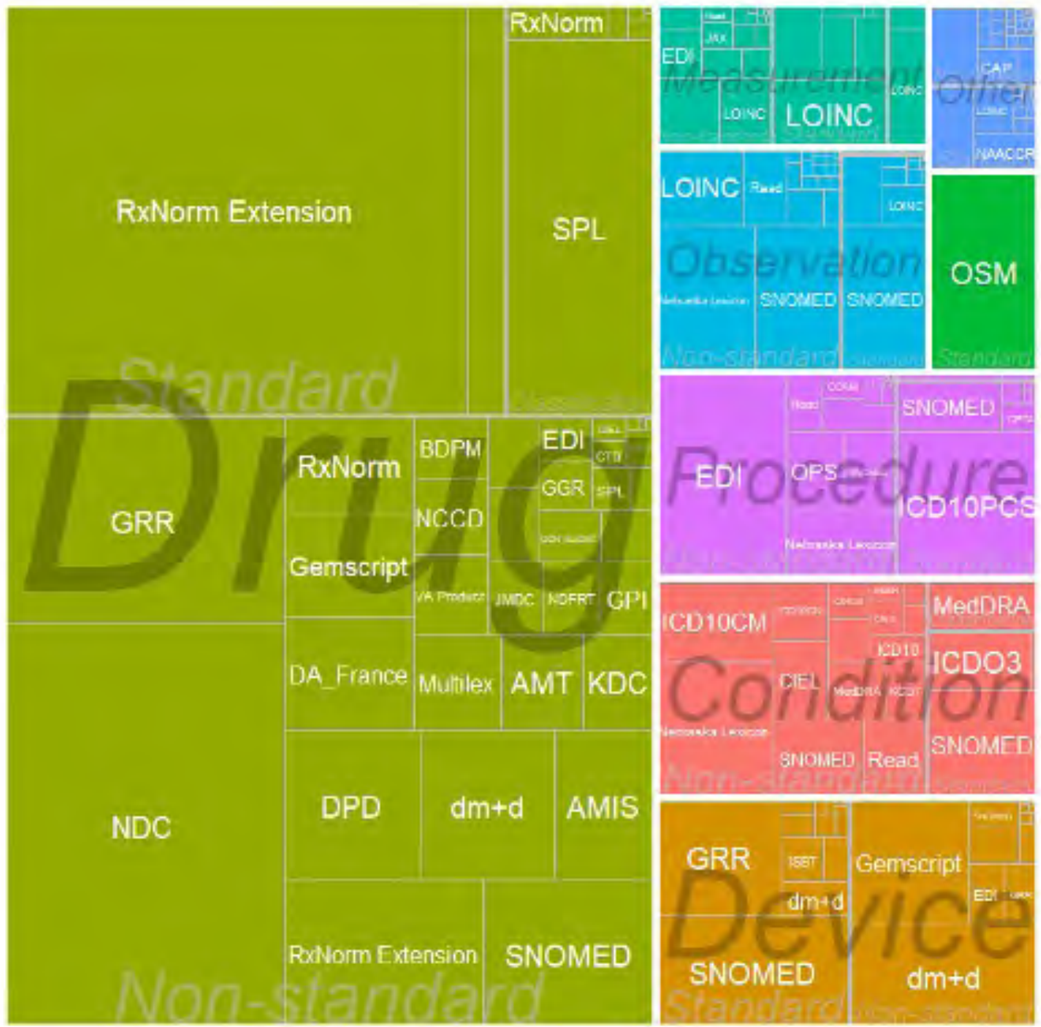
Aarhus University Hospital Database (Claims; Denmark)
Advocate Aurora Health & University of Madison Health Non-Muscle Invasive Bladder Cancer (EHR; USA)
Advocate Aurora Health / U of Madison Bladder Cancer (EHR; USA)
Agenzia regionale di sanità della Toscana (ARS) (Claims; Italy)
Ajou University Hospital (EHR; South Korea)
Ajou University Hospital Bio-signals (ICU + EHR; South Korea)
Akrivia Health (EHR; UK)
All of Us Research Program (EHR, Survey; USA)
ALTAMED (University of Southern California) (EHR; USA)
Amsterdam UMC (EHR; Netherlands)
APDP (EHR; Portugal)
APHP-EDS (CDW; France)
Asan Medical Center (EHR; South Korea)
Assistance Publique - Hopitaux de Marseille (EHR; France)
ATS Bergamo (Regional Dataset; Italy)
AU-ePBRN (Australian Electronic practice based research network) (Claims; Australia)
Australian Electronic practice based research network (EHR; Australia)
AZ Delta (EHR; Belgium)
AZ Klina (EHR; Belgium)
Azienda Ospedaliera SS Antonio e Biagio e Cesare Arrigo (CDW; Italy)
Azienda Ospedaliera Universitaria Integrata Verona (EHR; Italy)
Barts Health NHS Trust (EHR; UK)
BCB Medical Ltd. (EHR; Finland)
Beijing Anding Psychiatry Hospital (EHR; China)
BIOCRUCES BIZKAIA HEALTH RESEARCH INSTITUTE (EHR; Spain)
Blue Health Intelligence (Claims; USA)
Bordeaux hospital (EHR; France)
Bordeaux PharmacoEpi (EHR; France)
Boston Medical Center (EHR; USA)
Brown University - Rhode Island HIE (EHR; USA)
Bucheon Sejong Hospital (EHR; South Korea)
Buddhimed Technologies (EHR; India)
CALIBER (EHR; UK)
CancerDataNet GmbH (EHR; Germany)
Carilion Clinic (Claims; USA)
Case Western (EHR; USA)
CEGEDIM HEALTH DATA (Registry; France)
Center for Surgical Science (CSS) (EHR; Denmark)
Centre Hospitalier Universitaire de Montpellier (EHR; France)
Centro Clínico Academico a Braga, Associaçiao (2CA-Braga) (EHR; Portugal)
Centro Clínico Academico a Braga, Associaçiao (EHR; South Korea)
Cerner (EHR; USA)
Charité - Universitätsmedizin Berlin (EHR; Germany)
Cherokee Health Systems (EHR; USA)
Children's Hospital of Colorado (EHR; USA)
Children's Hospital of Los Angeles (EHR; USA)
Children's Hospital of Philadelphia (EHR; USA)
Children's National (EHR; USA)
Chonnam National University Hospital (EHR; South Korea)
Chonnam National University Hwasun Hospital (EHR; South Korea)
CHU de Toulouse (EHR; France)
Chungnam National University Hospital (EHR; South Korea)
Clinical Center of Serbia (EHR; Serbia)
Clinical centre of Nis (EHR; Serbia)
Clinical Hospital Dubrava (EHR; Croatia)
Clinical Practice Research Datalink (CPRD) (EHR; UK)
Columbia University Irving Medical Center (EHR; USA)
Connected Bradford (EHR; UK)
Consorci Mar Parc de Salut de Barcelona (PSMAR) (EHR; Spain)
CRHFEI (EHR; USA)
Daegu Catholic University Medical Center (CDW; South Korea)
Dankook University Hospital (EHR; South Korea)
DARTNet Institute: CER2 Study (EHR; USA)
DataSUS Ambulatory (EHR; Brazil)
Decision Resources Group (DRG) (EHR; USA)
Department of Health Services - Los Angeles (CDW; USA)
Dongguk University Ilsan Hospital (Claims; South Korea)
Duke University (Claims, EHR ; USA)
Eau Claire Cooperative Health Center (Claims; USA)
EBMT: The European Society for Blood and Marrow Transplantation (EHR; Netherlands)
Estonian Genome Center at the University of Tartu (EGCUT) (EHR; Estonia)
European Society for Blood and Marrow Transplantation (Registry; Finland)
Ewha Womans University Medical Center Mokdong (EHR; South Korea)
Finnish Hematology Registry/ HUS (Biobank; Finland)
Flatiron - OSCER (EHR; USA)
Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico (EHR; Italy)
Fondazione IRCCS Istituto Neurologico Carlo Besta (EHR; Italy)
Fondazione IRCCS Policlinico San Matteo (EHR; Italy)
Fondazione Istituto Nazionale dei Tumori (EHR; Italy)
Fondazione Poliambulanza Istituto Ospedaliero (EHR; Italy)
Fundació Institut d'Investigacions Mèdiques (FIMIM) (EHR; Spain)
Fundacion de Investigacion Biomedica del Hospital Universitario 12 de Octubre (CDW; Spain)
FUNDACION PARA LA INVESTIGACION DEL HOSPITAL UNIVERSITARIO LA FE DE LA COMUNIDAD VALENCIANA (HULAFE) (EHR; Spain)
Fundacion Para La Investigacion del Hospital Universitario La Fe de la Comunidad Valenciana (HULAFE) (EHR; Spain)
Gachon University Gil Medical Center (EHR; South Korea)
Gangnam Severance Hospital (EHR; South Korea)
Geisinger Health System (CDW; USA)
GENERAL HOSPITAL OF KAVALA (EHR; Greece)
General Hospital of Kavala (EHR; USA)
Geneva Cancer Registry (EHR; Switzerland)
Georgetown University ARIA (EHR; USA)
GeriOMOP (Registry; USA)
GOSH (EHR; UK)
Great Ormond Street Hospital NHS Foundation Trust (GOSH) (EHR; South Korea)
Hanover Medical School, Germany (EHR; Germany)
Harvard University Mass General Brigham (Nursing home + drug; USA)
Harvey Walsh Ltd (CDW; UK)
Health Data Hub (EHR; France)
Health Informatics Centre (HIC) (EHR; UK)
Health Insurance Review & Assessment Service (National Dataset; South Korea)
Healthcare Cost and Utilization Project, Nationwide Inpatient Sample (HCUP/NIS) (Hospital Billing; China)
HealthVerity (EHR; USA)
Hierarchia d.o.o. & University Hospital Centre Zagreb (EHR; Croatia)
HM Hospitals (Claims; Spain)
Hospital da Luz Learning Health (Claims; Portugal)
Hospital del Mar (HMAR) (EHR; Spain)
Hospital District of Helsinki and Uusimaa (EHR; Finland)
Hospital District of Southwest Finland (EHR; Finland)
HUG and SCQM (EHR; Switzerland)
HUS Datalake eCareforMe POC (EHR; Finland)
IBM CED (EHR; USA)
IBM MarketScan CCAE + MDCR (EHR; USA)
IBM(R) MarketScan(R) Commercial Claims (CCAE) (Registry; USA)
IBM(R) MarketScan(R) Medicare Supplemental Database (MDCR) (EHR; USA)
IBM(R) MarketScan(R) Multi-State Medicaid Database (MDCD (Claims, EHR ; USA)
Ican School of Medicine at Mount Sinai (Claims; USA)
IKNL (Claims; Netherlands)
Incheon Sejong Hospital (Claims; South Korea)
Indiana University School of Medicine / Regenstrief Institute (Claims; USA)
INFOBANCO12 (EHR; Spain)
Information System of Parc de Salut Mar (IMASIS) (CDW; Spain)
Inha University Hospital (EHR; South Korea)
Innovative Medical Research SA (EHR; Greece)
Inova Health System (EHR; USA)
Integrated Primary Care Information (IPCI) (EHR; Netherlands)
IQVIA Australia LPD (EHR; Australia)
IQVIA Belgium LPD (EHR; Belgium)
IQVIA Brazil (EHR; Brazil)
IQVIA France DA (EHR; France)
IQVIA France LPD (Hospital Billing; France)
IQVIA Germany DA (EHR; Germany)
IQVIA Hospital CDM (EHR; USA)
IQVIA HTI (EHR; UK)
IQVIA Italy LPD (Claims; Italy)
IQVIA Japan HIS (EHR; Japan)
IQVIA Japanese Claims (EHR; Japan)
IQVIA LPD Australia (EHR; Australia)
IQVIA OncoEMR (EHR; USA)
IQVIA Spain LPD (EHR; Spain)
IQVIA US Ambulatory EMR (EHR; USA)
IQVIA US Hospital Charge Data Master (CDM) (Claims; USA)
IQVIA US Oncology EMR (Claims; USA)
IQVIA US Open Claims (EHR; USA)
IQVIA US Pharmetrics Plus (PMTX+) (EHR; USA)
IRCCS Policlinico San Donato (EHR; Italy)
Istanbul University Istanbul Faculty of Medicine (Claims; Turkey)
IUC Cerrahpaşa TIP Fakületesi (EHR; Turkey)
Japan Medical Data Center (JMDC) (EHR; Japan)

Jeonbuk National University Hospital (EHR; South Korea)
Jiangsu Province People's Hospital (EHR; China)
Johns Hopkins University (EHR; USA)
Kangbuk Samsung Hospital (EHR; South Korea)
Kangdong Sacred Heart Hospital (EHR; South Korea)
Kangwon National University Hospital (EHR; South Korea)
Keck Medicine of University of Southern California (EHR; USA)
Khoo Teck Puat Hospital - T2DM Cohort (SG_T2DM) (EHR; Singapore)
Khoo Teck Puat Hospital (SG_KTPH) (EHR; Singapore)
Klinično-bolnički centar Zvezdara (Clinical-hospital center Zvezdara) (EHR; Serbia)
Klinično-bolnički centar Zvezdara (EHR; UK)
Konkuk University Medical Center (EHR; South Korea)
Konyang University Hospital (EHR; South Korea)
Korea University Anam Hospital (EHR; South Korea)
Korea University Ansan Hospital (EHR, Genomics; South Korea)
Korea University Guro Hospital (EHR; South Korea)
KTPH Diabetes Data (EHR; Singapore)
Kyung Hee University Hospital at Gangdong (EHR; South Korea)
Kyung Hee University Medical Center (EHR; South Korea)
Kyungpook National University Chilgok Hospital (EHR; South Korea)
Kyungpook National University Hospital (EHR; South Korea)
Leeds Teaching Hospitals (EHR; UK)
Lille University Hospital (EHR; France)
Loyola University New Orleans (EHR; USA)
LynxCare (EHR; Belgium)
Maine Medical Center (EHR; USA)
Marina Salud S.A. (Claims; Spain)
Mayo Clinic (National Dataset; USA)
MDV (Medical Data Vision) (EHR; Japan)
MEB KI (EHR; Sweden)
Medaman (EHR; Belgium)
Medical University of South Carolina (Claims; USA)
Medical University of Vienna (EHR, Registry; Austria)
Medicare Research Identifiable Files (EHR; USA)
Memorial Sloan Kettering Cancer Center (EHR; USA)
Modena Oncology Center - Azienda Ospedaliera Modena (EHR; Italy)
Momentum AD (EHR; USA)
Montefiore Medical Center (Albert Einstein College of Medicine) (EHR; USA)
MS Forschungs- und Projektentwicklungs-gGmbH (EHR; Germany)
Myongji Hospital (EHR; South Korea)
Nanfang Hospital COVID-19 Research Database (NFHCRD) (EHR; China)
National Cancer Center (Registry; South Korea)
National Health Insurance Service Ilsan Hospital (Registry; South Korea)
National Intensive Care Evaluation foundation (EHR; Netherlands)
National Scientific Programme "E-Health in Bulgaria" (EHR; Bulgaria)
National University of Hospital (SG_NUH) (Claims; Singapore)
Nemours Children's Health System (EHR; USA)
NHIRD (EHR; Taiwan)
NorthShore University HealthSystem (EHR; USA)
Northwestern Medicine Enterprise Data Warehouse (NMEDW) (EHR; USA)
NYC-CDRN (EHR; USA)
NYU Langone (EHR; USA)
OCHIN (Oregon Community Health Information Network) (EHR; USA)
Ochsner Medical Center (EHR; USA)
Oklahoma University (EHR; USA)
Optimum Patient Care Limited (EHR; UK)
Optum© De-Identified Clinformatics(R) Data Mart Database (Claims; USA)
Optum© De-Identified Clinformatics(R) Data Mart Database - SES & DOD (EHR; USA)
Optum© de-identified Electronic Health Record Dataset (PANTHER) (EHR; USA)
Oregon Health & Science University (; USA)
Parc Sanitari Sant Joan de Déu (EHR; Spain)
Pareto Intelligence (EHR; USA)
Pedianet (EHR; Italy)
PEDSnet (Claims; USA)
Penn State (EHR; USA)
Pharmaceutical Benefits Scheme 10% extract (Hospital Billing; Australia)
Pharmo (EHR; Netherlands)
Premier Healthcare Database (PHD) (EHR; USA)
Primary Care GP data (Patron) (Registry; Australia)
Pusan National University Hospital (EHR; South Korea)
Queen Mary University of London (Registry; UK)
QueensCare - Los Angeles (EHR; USA)
Registre National du Cancer du Luxembourg (EHR; Luxembourg)
Reliant Medical Group (EHR; USA)
Rioja Salud (EHR; Spain)
Royal College of General Practitioners Research and Surveillance Centre (EHR; UK)
Rush University Medical Center (EHR; USA)
Rutgers (EHR; USA)
SAIL Databank (Claims, EHR ; UK)
Samsung Medical Center (EHR; South Korea)
Saudi FDA (EHR; Saudi Arabia)
Saudi Pharmacoepidemiology Database (EHR; Saudi Arabia)
Semmelweis University (EHR; Hungary)
Seoul National University Bundang Hospital (EHR; South Korea)
Seoul National University Hospital (CDW; South Korea)
Servicio Cántabro de Salud and IDIVAL (EHR; Spain)
Servicio Madrileño de Salud (EHR; Spain)
Severance Hospital (EHR; South Korea)
SIMG, Società Italiana di Medicina Generale e delle cure Primarie (Italian College of General Practice (EHR; Italy)
Società Italiana di Medicina Generale e delle cure Primarie (EHR; Taiwan)
Soon Chun Hyang University Hospital Bucheon (EHR; South Korea)
Soon Chun Hyang University Hospital Cheonan (EHR; South Korea)
Soon Chun Hyang University Hospital Gumi (EHR; South Korea)
Soon Chun Hyang University Hospital Seoul (EHR; South Korea)
STAnford medicine Research data Repository (STARR) OMOP (EHR; USA)
Stony Brook (EHR; USA)
Surveillance, Epidemiology, and End Results Program (SEER) (Claims, Registry; Netherlands)
Surveillance, Epidemiology, and End Results Program (SEER): B-Cell (EHR; USA)
Sydney Local Health District (LHD) (; Australia)
Taipei Medical University Hospital (EHR; Taiwan)
TCC - Los Angeles (EHR; USA)
The Catholic University of Korea Seoul ST. Mary's Hospital (EHR; South Korea)
The Catholic University of Korea Yeouido ST. Mary's Hospital (EHR; South Korea)
The Directorate of Government Medical Centers at the Israeli Ministry Of Health (EHR; Israel)
The Healthcare Cost and Utilization Project (HCUP), Nationwide Inpatient Sample (NIS) (EHR; USA)
The Information System for Research in Primary Care – Hospitalization Linked Data (SIDIAP-H) (Registry; Spain)
The Information System for Research in Primary Care (EHR; Spain)
The Information System for Research in Primary Care (SIDIAP) (EHR; Spain)
The Information System for Research in Primary Care-Hospitalization (SIDIAP-H) (EHR; USA)
The National Health and Nutrition Examination Survey (NHANES) (EHR; USA)
THIN BE (EHR; Belgium)
THIN FR (Survey Data; France)
THIN RO (EHR; Romania)
THIN UK (EHR; UK)
Tianjin Anding Psychiatry Hospital (EHR; China)
Tufts MC Research Data Warehouse (TRDW) (EHR; USA)
Tulane (EHR; USA)
UCL (EHR; UK)
UK Biobank (EHR; UK)
UK Integrated Medical Record Database (IMRD) (EHR; UK)
UK National Neonatal Research Database (EHR; UK)
UKCRIS (EHR, Survey; UK)
UKER (EHR; Germany)
Ulsan University Hospital (EHR; South Korea)
ULSM (EHR; Portugal)
UMass Memorial Medical Center (EHR; USA)
UNC Chapel Hill (EHR; USA)
University College London Hospitals NHS Foundation Trust (EHR; UK)
University Medical Center New Orleans (EHR; USA)
University Medicine Dresden (EHR; Germany)
University MS Center (EHR; Belgium)
University of Alabama Birmingham (EHR; USA)
University of Arkansas (EHR; USA)
University of Buffalo (EHR; USA)
University of California, Davis (EHR; USA)
University of California, Irvine (EHR; USA)
University of California, Los Angeles (EHR; USA)
University of California, Riverside (EHR; USA)
University of California, San Diego (EHR; USA)
University of California, San Francisco (EHR; USA)
University of Chicago (EHR; USA)
University of Cincinnati (EHR; USA)
University of Colorado (EHR; USA)
University of Colorado Anschutz (EHR; USA)
University of Edinburgh (EHR; UK)
University of Illinois Chicago (EHR; USA)
University of Iowa (EHR; USA)
University of Kentucky (EHR; USA)
University of Miami (EHR; USA)
University of Michigan (EHR; USA)
University of Minnesota (EHR; USA)
University of Mississippi Medical Center (EHR; USA)
University of Nebraska Medical Center (EHR; USA)
University of Oslo PharmaSafe (National Dataset; Hungary)
University of Oslo, Department of Pharmacy, Pharmacoepidemiology and Drug Safety Research Group (EHR; Norway)
University of Pittsburgh - Banner (EHR; USA)
University of Pittsburgh (EHR; USA)
University of Rochester (EHR; USA)
University of Tartu (EHR; Estonia)
University of Texas Houston (EHR; USA)
University of Texas Medical Branch (EHR; USA)
University of Utah (EHR; USA)
University of Virginia (EHR; USA)
University of Washington Medicine COVID Research Dataset (EHR; USA)
University of Washington Medicine COVID Research Dataset (UWM-CRD) (EHR; USA)
US Department of Defense OMOP (EHR; USA)
US Department of Veterans Affairs (EHR; USA)
Vall d'hebron Hospital Campus (EHR; Spain)
Vanderbilt University (Claims, EHR ; USA)
Veradigm Health Insights Data - Allscripts (EHR; USA)
Veradigm Health Insights Data - Practice Fusion (EHR; USA)
Virginia Commonwealth University (EHR; USA)
Wake Forest University (EHR; USA)
Wanfang Hospital (EHR; Taiwan)
WashU St Louis (EHR; USA)
Weill Cornell Medical Center - Epic (EHR; USA)
Weill Cornell Medicine/NewYork-Presbyterian Hospital (East Campus) (EHR; USA)
West Virginia University (EHR; USA)
Winship Cancer Institute of Emory University (EHR; USA)
Wonju Severance Christian Hospital (EHR; South Korea)
Wonkwang University Hospital (EHR; South Korea)
Yongin Severance Hospital (EHR, Claims, Clinical Trials; South Korea)
Ziekenhuis Oost-Limburg (EHR; Belgium)

The EHDEN-led initiative to certify small-and-medium enterprises to assist in mapping data across Europe has helped bring 98 partners into the OHDSI data network.

OHDSI Vocabularies

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).

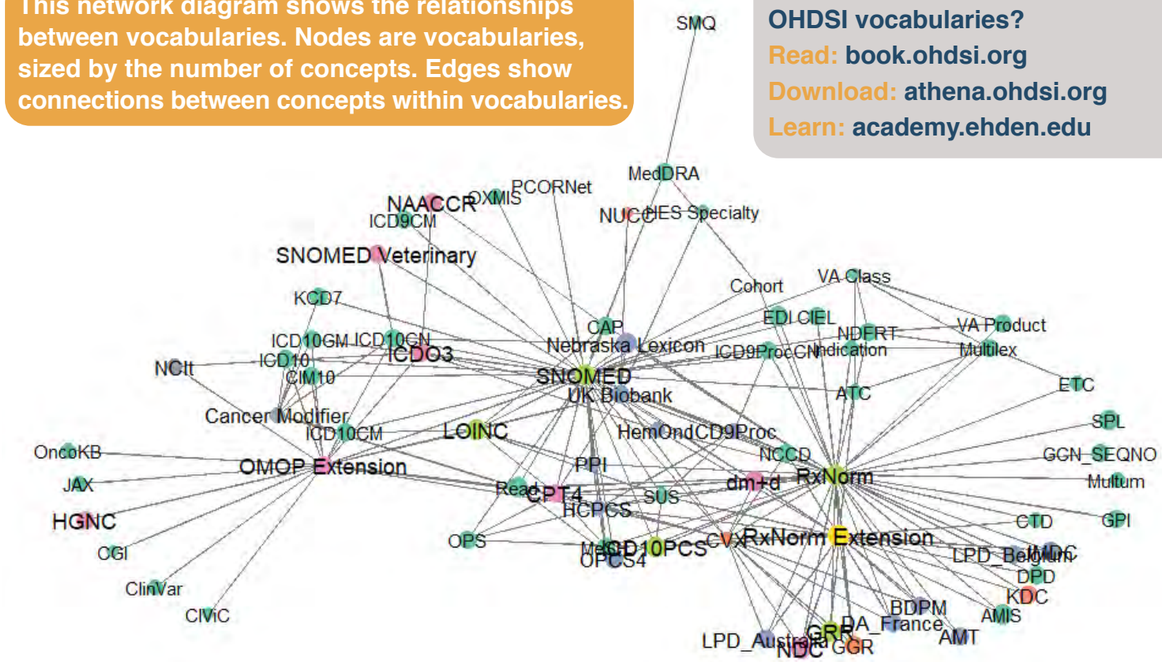
OHDSI Vocabularies By The Numbers

as of v5.0 • July 13, 2021

- 9,833,611 concepts
  - 3,392,214 standard concepts
  - 701,277 classification concepts
- 133 vocabularies
- 40 domains
- 75,164,214 concept relationships
- 87,392,704 ancestral relationships
- 2,703,706 concept synonyms

1 Shared Resource to Enable Data Standards

This network diagram shows the relationships between vocabularies. Nodes are vocabularies, sized by the number of concepts. Edges show connections between concepts within vocabularies.



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.

OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

VI.

Open-Source Software

HADES
HEALTH ANALYTICS DATA-TO-EVIDENCE SUITE

CohortMethod New-user cohort studies using large-scale regression for propensity and outcome models. Learn more…	**SelfControlledCaseSeries** Self-Controlled Case Series analysis using few or many predictors, includes splines for age and seasonality. Learn more…	**Cyclops** Highly efficient implementation of regularized logistic, Poisson and Cox regression. Learn more…	**DatabaseConnector** Connect directly to a wide range of database platforms, including SQL Server, Oracle, and PostgreSQL. Learn more…	**SqlRender** Generate SQL on the fly for the various SQL dialects. Learn more…
SelfControlledCohort A self-controlled cohort design, where time preceding exposure is used as control. Learn more…	**EvidenceSynthesis** Routines for combining causal effect estimates and study diagnostics across multiple data sites in a distributed study. Learn more…	**ParallelLogger** Support for parallel computation with logging to console, disk, or e-mail. Learn more…	**FeatureExtraction** Automatically extract large sets of features for user-specified cohorts using data in the CDM. Learn more…	**Andromeda** Storing very large data objects on a local drive, while still making it possible to manipulate the data in an efficient manner. Learn more…
PatientLevelPrediction Build and evaluate predictive models for user-specified outcomes, using a wide array of machine learning algorithms. Learn more…	**EmpiricalCalibration** Use negative control exposure-outcome pairs to profile and calibrate a particular analysis design. Learn more…	**BigKnn** A large scale k-nearest neighbor classifier using the Lucene search engine. Learn more…	**ROhdsiWebApi** Interact with OHDSI WebAPI web services. Learn more…	**OhdsiSharing** Securely sharing (large) files between OHDSI collaborators. Learn more…
MethodEvaluation	**CohortDiagnostics**	**Hydra**	**Eunomia**	**CirceR**

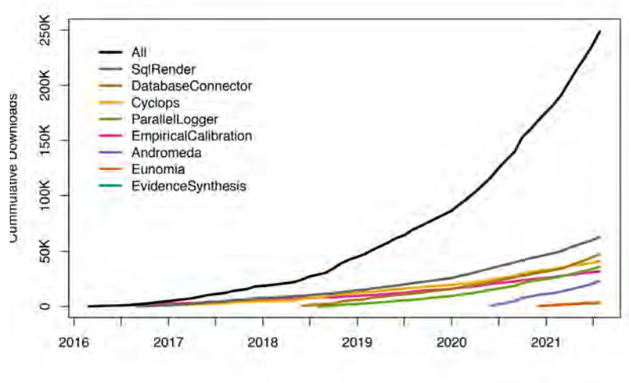
HADES

Health Analytics Data-to-Evidence Suite

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

One critical factor in OHDSI's ability to perform rigorous, ground-breaking analyses lies under the surface, but it holds an equally important role in the overall community mission.

A core foundation for OHDSI is open-source software development, and a small group of community collaborators, led by Martijn Schuemie, has generated a collection of analytics tools that enable research both in and out of the OHDSI community.



HADES — the Health Analytics Data-to-Evidence Suite — is a set of 20 open-source R packages for large scale analytics, including population characterization, 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 large 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, CHARYBDIS, hydroxychloroquine safety, the ongoing work with COVID AESI characterization, and many more. All packages are developed and released as open-source tools at github.com/OHDSI. Amongst the HADES ecosystem, eight packages have matured to be additionally released on CRAN (The Comprehensive R Archive Network, a public repository for all R users).

"Our community, and observational researchers in general, owe an enormous debt of gratitude to Martijn and the HADES team for leading this effort," said Provost and Senior Vice President for Academic Affairs at Northeastern University David Madigan, who is leading efforts around the new OHDSI Center at the Roux Institute. "Open-source development within the OHDSI community is the quiet force that is impacting important evidence that can save lives, and it shouldn't be taken for granted."

These 8 HADES packages have matured to be released on CRAN and have been downloaded more than 250,000 times (see graphic above).

Beyond network studies, HADES (formerly known as the OHDSI Methods Library) allows researchers to conduct analyses locally. It supports best practices for use of observational data as learned from previous and ongoing research; for example, the population-level estimation methods have been extensively evaluated using the OHDSI Methods Benchmark, as published in the Harvard Data Science Review.

Researchers can learn how to use HADES through documentation found in the Book of OHDSI (ohdsi.github.io/TheBookOfOhdsi/).

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

OHDSI's reach has expanded over the last year, including its recent role supporting the FDA BEST program in vaccine surveillance, as well as informing best practices in the most recent EMA revision of its guidelines. Researchers continue to join the 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 aided the growth of OHDSI, every small contribution from the community can aid the advancement of HADES.

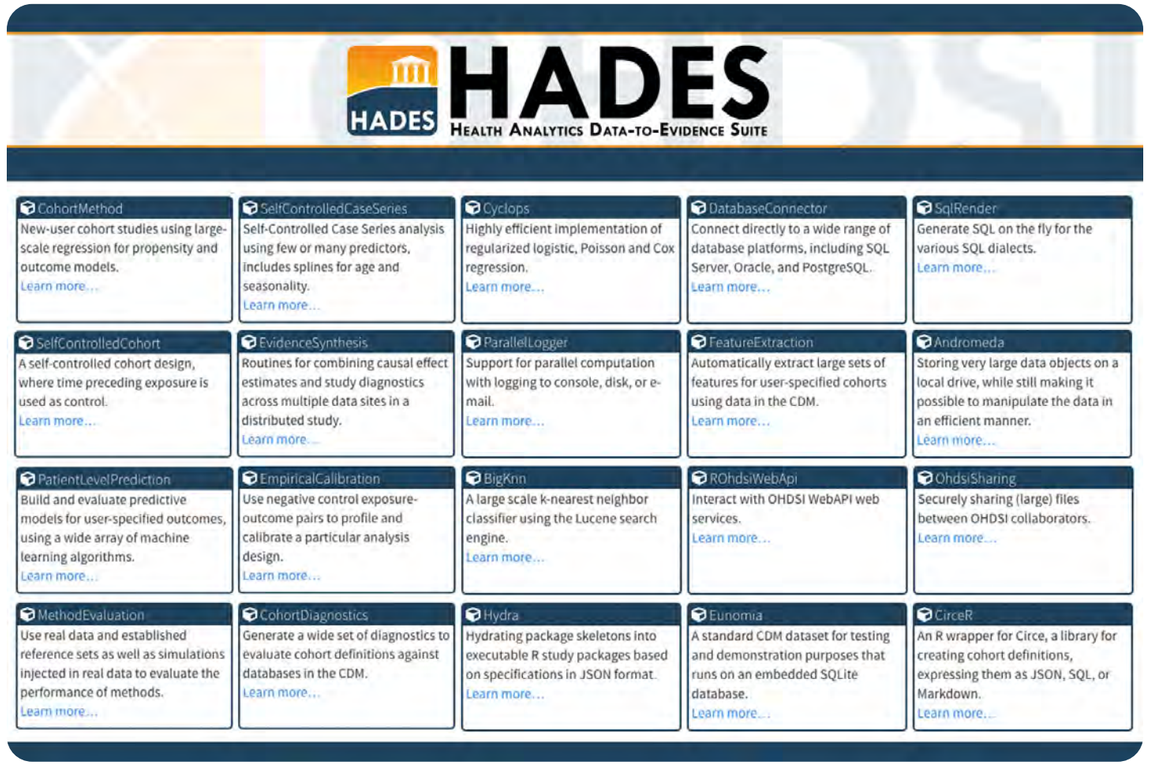
"Open-source development within the HADES ecosystem has been critical to our growth and success as a community," said George Hripcsak, Chair and Vivian Beaumont Allen Professor of Biomedical Informatics at Columbia, the coordinating center for OHDSI. "Martijn and the HADES team have done extraordinary 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."

Open-Source Software

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

How can you help?

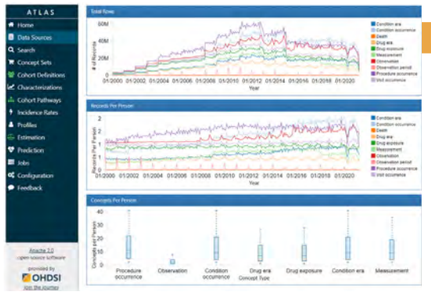
- **developers can contribute by helping develop and test code**
- **users of the tools can help with testing, user documentation and other training resources**
- **those with the means can provide financial support to help pay for developers specifically focused on open-source development**
- **anybody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



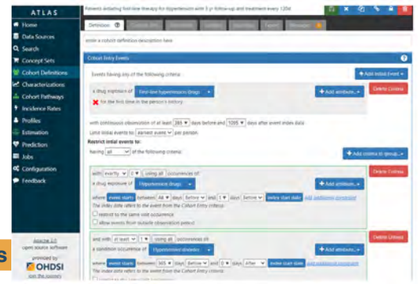
ATLAS

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



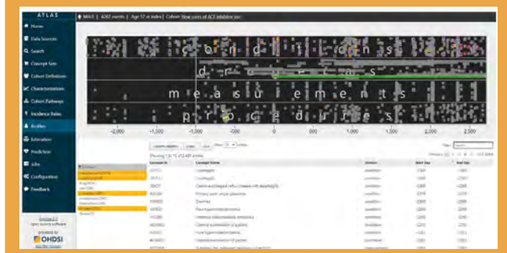
Explore Data

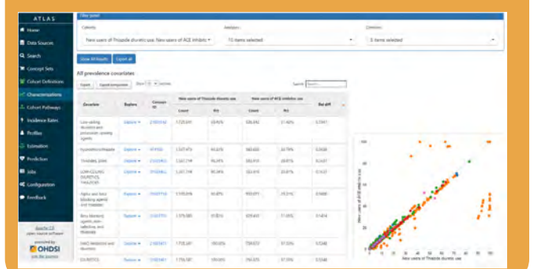


Design Analyses

Produce Code

Generate 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/

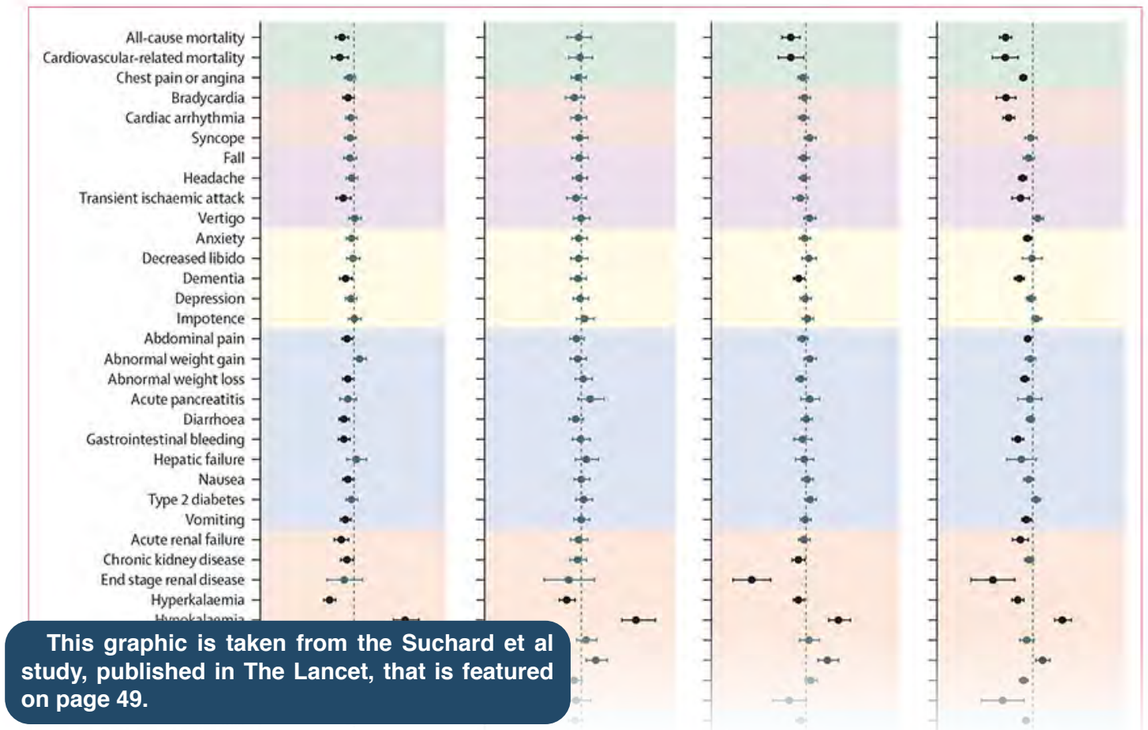
Download: github.com/ohdsi/atlas

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.

Research Article

Statistics in Medicine

Received 12 November 2012, Accepted 3 July 2013 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.5925

Interpreting observational studies: why empirical calibration is needed to correct *p*-values

Martijn J. Schuemie,[a,b*†] Patrick B. Ryan,[b,c] William DuMouchel,[b,d] Marc A. Suchard[b,e] and David Madigan[b,f]

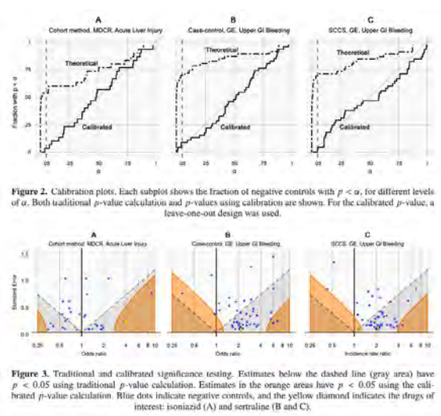


Figure 2. Calibration plots. Each subplot shows the fraction of negative controls with $p < \alpha$, for different levels of α. Both traditional p-value calculation and p-values using calibration are shown. For the calibrated p-value, a leave-one-out design was used.

Figure 3. Traditional and calibrated significance testing. Estimates below the dashed line (gray area) have $p < 0.05$ using traditional p-value calculation. Estimates in the orange areas have $p < 0.05$ using the calibrated p-value calculation. Blue dots indicate negative controls, and the yellow diamond indicates the drugs of interest: isoniazid (A) and sertraline (B and C).

PNAS

Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data

Martijn J. Schuemie[a,b,1], George Hripcsak[a,c,d], Patrick B. Ryan[a,b,c], David Madigan[a,e], and Marc A. Suchard[a,f,g,h]

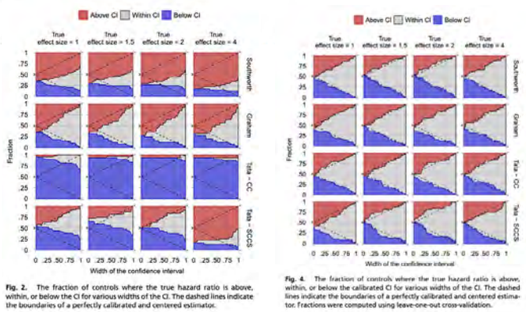


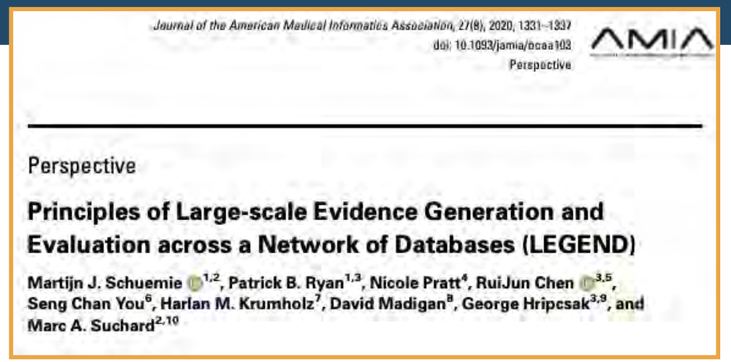
Fig. 2. The fraction of controls where the true hazard ratio is above, within, or below the CI for various widths of the CI. The dashed lines indicate the boundaries of a perfectly calibrated and centered estimator.

Fig. 4. The fraction of controls where the true hazard ratio is above, within, or below the calibrated CI for various widths of the CI. The dashed lines indicate the boundaries of a perfectly calibrated and centered estimator. Fractions were computed using leave-one-out cross-validation.

LEGEND in Principle

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.



Journal of the American Medical Informatics Association, 27(8), 2020, 1331–1337
doi: 10.1093/jamia/ocaa103
Perspective

AMIA

Perspective

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

Martijn J. Schuemie[1,2], Patrick B. Ryan[1,3], Nicole Pratt[4], RuiJun Chen[3,5], Seng Chan You[6], Harlan M. Krumholz[7], David Madigan[8], George Hripcsak[3,9], and Marc A. Suchard[2,10]

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. All 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.

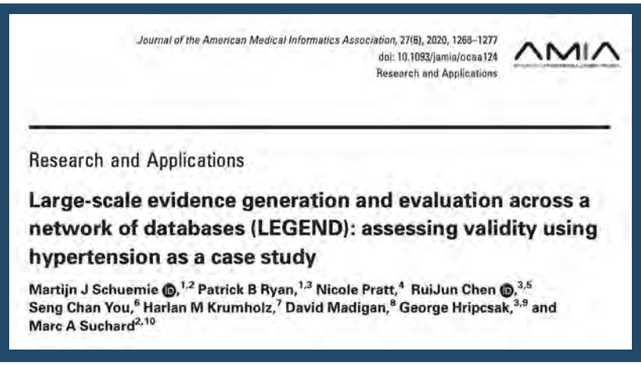
9. LEGEND will generate evidence across a network of multiple databases. Multiple heterogeneous databases (different data capture processes, health-care systems, and populations) will be used to generate the evidence to allow an assessment of the replicability of findings across sites. **Aim:** Enhances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

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.



Journal of the American Medical Informatics Association, 27(8), 2020, 1268–1277
doi: 10.1093/jamia/ocaa124
Research and Applications

AMIA

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 Schuemie,[1,2] Patrick B Ryan,[1,3] Nicole Pratt,[4] RuiJun Chen,[3,5] Seng Chan You,[6] Harlan M Krumholz,[7] David Madigan,[8] George Hripcsak,[3,9] and Marc A Suchard[2,10]

JAMA Internal Medicine | Original Investigation

Comparison of Cardiovascular and Safety Outcomes of Chlorthalidone vs Hydrochlorothiazide to Treat Hypertension

George Hripcsak, MD, MS; Marc A. Suchard, MD, PhD; Steven Shea, MD; RuiJun Chen, MD; Seng Chan You, MD; Nicole Pratt, PhD; David Madigan, PhD; Harlan M. Krumholz, MD, SM; Patrick B. Ryan, PhD; Martijn J. Schuemie, PhD

IMPORTANCE Chlorthalidone is currently recommended as the preferred thiazide diuretic to treat hypertension, but no trials have directly compared risks and benefits.

OBJECTIVE To compare the effectiveness and safety of chlorthalidone and hydrochlorothiazide as first-line therapies for hypertension in real-world practice.

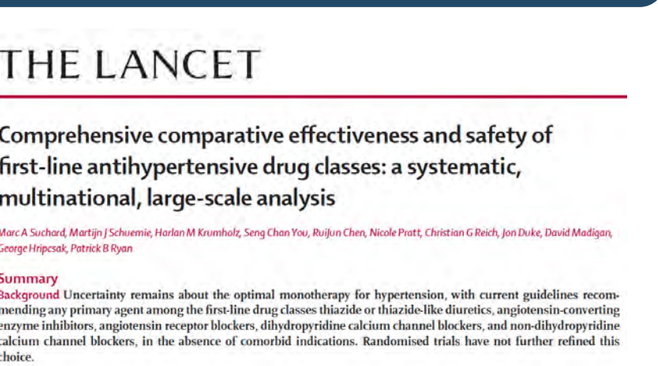
DESIGN, SETTING, AND PARTICIPANTS This is a Large-Scale Evidence Generation and Evaluation in a Network of Databases (LEGEND) observational comparative cohort study with large-scale propensity score stratification and negative-control and synthetic positive-control calibration on databases spanning January 2001 through December 2018. Outpatient and inpatient care episodes of first-time users of antihypertensive monotherapy in the United States based on 2 administrative claims databases and 1 collection of electronic health records were analyzed. Analysis began June 2018.

EXPOSURES Chlorthalidone and hydrochlorothiazide.

MAIN OUTCOMES AND MEASURES The primary outcomes were acute myocardial infarction, hospitalization for heart failure, ischemic or hemorrhagic stroke, and a composite cardiovascular disease outcome including the first 3 outcomes and sudden cardiac death. Fifty-one safety outcomes were measured.

RESULTS Of 730 225 individuals (mean [SD] age, 51.5 [13.3] years; 450 100 women [61.6%]), 36 918 were dispensed or prescribed chlorthalidone and had 149 composite outcome events, and 693 337 were dispensed or prescribed hydrochlorothiazide and had 3089 composite outcome events. No significant difference was found in the associated risk of myocardial infarction, hospitalized heart failure, or stroke, with a calibrated hazard ratio for the composite cardiovascular outcome of 1.00 for chlorthalidone compared with hydrochlorothiazide (95% CI, 0.85-1.17). Chlorthalidone was associated with a significantly higher risk of hypokalemia (hazard ratio [HR], 2.72; 95% CI, 2.38-3.12), hyponatremia (HR, 1.31; 95% CI, 1.16-1.47), acute renal failure (HR, 1.37; 95% CI, 1.15-1.63), chronic kidney disease (HR, 1.24; 95% CI, 1.09-1.42), and type 2 diabetes mellitus (HR, 1.21; 95% CI, 1.12-1.30). Chlorthalidone was associated with a significantly lower risk of diagnosed abnormal weight gain (HR, 0.73; 95% CI, 0.61-0.86).

CONCLUSIONS AND RELEVANCE This study found that chlorthalidone use was not associated with significant cardiovascular benefits when compared with hydrochlorothiazide, while its use was associated with greater risk of renal and electrolyte abnormalities. These findings do not support current recommendations to prefer chlorthalidone vs hydrochlorothiazide for hypertension treatment in first-time users was found. We used advanced methods, sensitivity analyses, and diagnostics, but given the possibility of residual confounding and the limited length of observation periods, further study is warranted.



THE LANCET

Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis

Marc A Suchard, Martijn J Schuemie, Harlan M Krumholz, Seng Chan You, RuiJun Chen, Nicole Pratt, Christian G Reich, Jon Duke, David Madigan, George Hripcsak, Patrick B Ryan

Summary

Background Uncertainty remains about the optimal monotherapy for hypertension, with current guidelines recommending any primary agent among the first-line drug classes thiazide or thiazide-like diuretics, angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, dihydropyridine calcium channel blockers, and non-dihydropyridine calcium channel blockers, in the absence of comorbid indications. Randomised trials have not further refined this choice.

Hypertension

Comparative First-Line Effectiveness and Safety of ACE (Angiotensin-Converting Enzyme) Inhibitors and Angiotensin Receptor Blockers: A Multinational Cohort Study

RuiJun Chen, Marc A. Suchard, Harlan M. Krumholz, Martijn J. Schuemie, Steven Shea, Jon Duke, Nicole Pratt, Christian G. Reich, David Madigan, Seng Chan You, Patrick B. Ryan, George Hripcsak

Comprehensive Comparative Effectiveness and Safety of First-Line β-Blocker Monotherapy in Hypertensive Patients

A Large-Scale Multicenter Observational Study

Seng Chan You, Harlan M. Krumholz, Marc A. Suchard, Martijn J. Schuemie, George Hripcsak, RuiJun Chen, Steven Shea, Jon Duke, Nicole Pratt, Christian G. Reich, David Madigan, Patrick B. Ryan, Rae Woong Park, Sungha Park

Starting On The Most Popular Hypertension Drug Isn't Most Effective, Per OHDSI's LEGEND Study

Thiazide diuretics demonstrate better effectiveness and cause fewer side effects than ACE inhibitors as first-line antihypertensive drugs, according to a report published Oct. 24, 2019, in The Lancet. The study factors insurance claim data and electronic health records from 4.9 million patients across nine observational databases, making it the most comprehensive one ever on first-line antihypertensives, and it provides additional context to the 2017 guidelines for high blood pressure treatment developed by the American College of Cardiology (ACC) and American Heart Association (AHA).

Collaborators within the OHDSI network produced the paper **"Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis"** as part of the collaborative's ongoing LEGEND (Large-Scale Evidence Generation and Evaluation across a Network of Databases) project, which applies high-level analytics to perform observational research on hundreds of millions of patient records within OHDSI's international database network.

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."

The paper also reported that non-dihydropyridine calcium channel blockers proved inferior to the four other first-line antihypertensive drug classes recommended in the 2017 guidelines; other classes included are angiotensin receptor blockers and dihydropyridine calcium channel blockers.

A LEGEND-ary Approach To Observational Science

"LEGEND is a unique, sophisticated approach to using observational data in a way that is reliable, rich and relevant," Suchard said. "With the availability of existing health data available, we can start to answer important clinical questions in a reproducible manner."

The LEGEND Hypertension project used state-of-the-art causal methods to address both observed confounding and residual bias. Covering patients from July 1996 to March 2018, the study filled in evidence gaps that were unavailable for the 2017 ACC/AHA guidelines. The RCTs from those guidelines factored approximately 31,000 users of either thiazide diuretics or ACE inhibitors, far fewer than the approximately 3.2 million new-users available in the LEGEND project.

"LEGEND is a novel approach that could transform the way we use real-world evidence in healthcare," said senior author Patrick Ryan, PhD, Adjunct Assistant Professor of Biomedical Informatics (Columbia University). "Rather than inefficiently conducting bespoke analyses one-question-one-method-one-database-at-a-time, leaving us vulnerable to various threats to scientific validity, LEGEND provides a systematic framework that can reproducibly generate evidence by applying advanced analytics across a network of disparate databases for a wide array of exposures and outcomes."

"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."

- Marc Suchard

2018 Titan Award recipient for Methodological Research

METHODS RESEARCH

The Journey To Reliable Evidence With Patient-Level Prediction

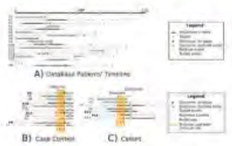
Clear specification of the prediction task:

- **Target Population: patients at risk**
- **Outcome: medical event to predict**
- **Time-at-risk (TAR): interval to predict if outcome will occur**

The patient-level prediction journey is more than just classification…

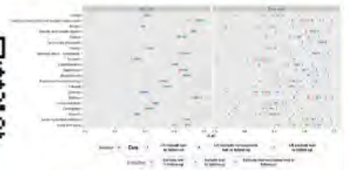
Design and Extraction

Study design



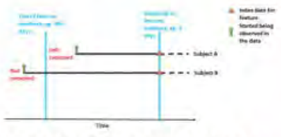
Case-control prone to misclassification and should be avoided; use cohort design

Loss to follow-up



Excluding non-outcomes lost to follow-up can bias the data

Feature extraction



Feature lookback can make an impact on model performance if it is too short (<180 days)

Model Development

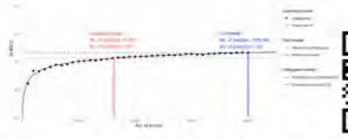
Learning across datasets

Models can be learned across datasets while maintaining privacy

Test/Train split

The design used to pick hyper-parameters and evaluate internal validity matters, even with big p and big n data.

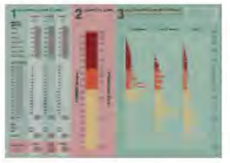
Sample size



Learning curves provide a way for model developers to determine whether they have sufficiently sized data

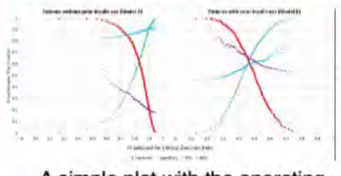
Model Evaluation

Model usability



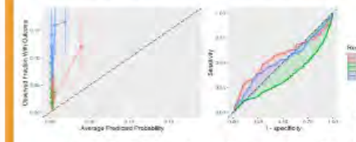
Simple score-based models are easier to apply and can be benchmarked against large-scale models

Visualizing performance

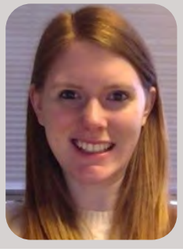


A simple plot with the operating characteristics for all cut-offs informs model usefulness

Network validation



The OHDSI network enables large-scale external validation and improves our understanding of models



"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

Join The PLP Journey

PLP GitHub: github.com/OHDSI/PatientLevelPrediction

#JoinTheJourney 51 OHDSI.org

OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

VIII. OHDSI Publications

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.

PNAS

Characterizing treatment pathways at scale using the OHDSI network

George Hripcsak[a,b,c,1], Patrick B. Ryan[c,d], Jon D. Duke[c,e], Nigam H. Shah[c,f], Rae Woong Park[c,g], Vojtech Huser[c,h], Marc A. Suchard[c,i,j,k], Martijn J. Schuemie[c,d], Frank J. DeFalco[c,d], Adler Perotte[a,c], Juan M. Banda[c,f], Christian G. Reich[c,l], Lisa M. Schilling[c,m], Michael E. Matheny[c,n,o], Daniella Meeker[c,p,q], Nicole Pratt[c,r], and David Madigan[c,s]

[a]Department of Biomedical Informatics, Columbia University Medical Center, New York, NY 10032; [b]Medical Informatics Services, NewYork-Presbyterian Hospital, New York, NY 10032; [c]Observational Health Data Sciences and Informatics, New York, NY 10032; [d]Epidemiology Analytics, Janssen Research and Development, Titusville, NJ 08560; [e]Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, IN 46205; [f]Center for Biomedical Informatics Research, Stanford University, CA 94305; [g]Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, South Korea, 443-380; [h]Lister Hill National Center for Biomedical Communications (National Library of Medicine), National Institutes of Health, Bethesda, MD 20894; [i]Department of Biomathematics, University of California, Los Angeles, CA 90095; [j]Department of Biostatistics, University of California, Los Angeles, CA 90095; [k]Department

JAMA | Original Investigation

Association of Ticagrelor vs Clopidogrel With Net Adverse Clinical Events in Patients With Acute Coronary Syndrome Undergoing Percutaneous Coronary Intervention

Seng Chan You, MD, MS; Yeunsook Rho, PhD; Behnood Bikdeli, MD, MS; Jiwoo Kim, MS; Anastasios Siapos, MSc; James Weaver, MSc; Ajit Londhe, MPH; Jaehyeong Cho, BS; Jimyung Park, BS; Martijn Schuemie, PhD; Marc A. Suchard, MD, PhD; David Madigan, PhD; George Hripcsak, MD, MS; Aakriti Gupta, MD, MS; Christian G. Reich, MD; Patrick B. Ryan, PhD; Rae Woong Park, MD, PhD; Harlan M. Krumholz, MD, SM

Risk of hydroxychloroquine alone and in combination with azithromycin in the treatment of rheumatoid arthritis: a multinational, retrospective study

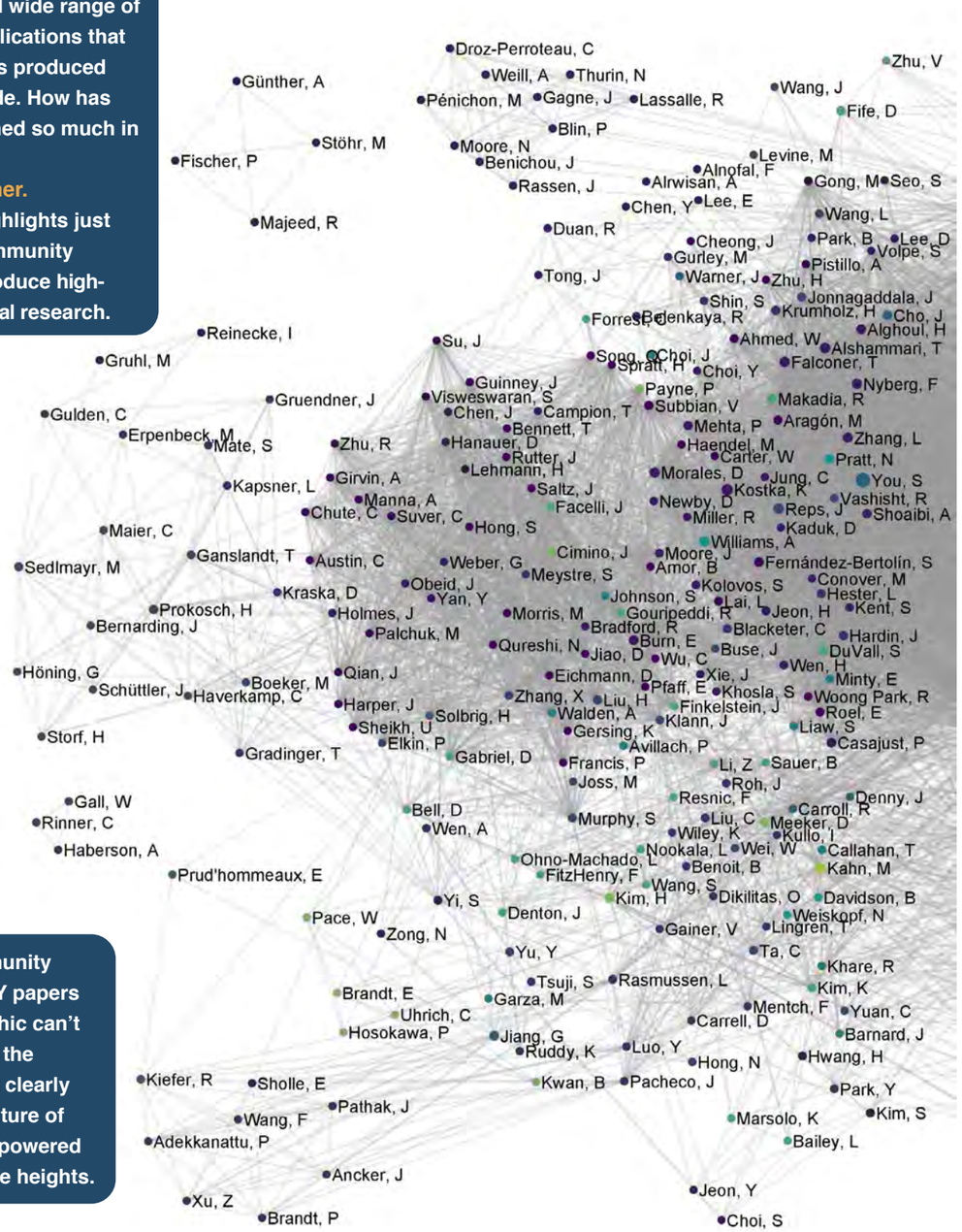
Jennifer C E Lane, James Weaver*, Kristin Kostka, Talita Duarte-Salles, Maria Tereza F Abrahao, Heba Alghoul, Osaid Alser, Thamir M Alshammari, Patricia Biedermann, Juan M Banda, Edward Burn, Paula Casajust, Mitchell M Conover, Aedin C Culhane, Alexander Davydov, Scott L DuVall, Dmitry Dymshyts, Sergio Fernandez-Bertolin, Kristina Fišter, Jill Hardin, Laura Hester, George Hripcsak, Benjamin Skov Kaas-Hansen, Seamus Kent, Sajan Khosla, Spyros Kolovos, Christophe G Lambert, Johan van der Lei, Kristine E Lynch, Rupa Makadia, Andrea Maier, Michael Morales, Andrew*

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?

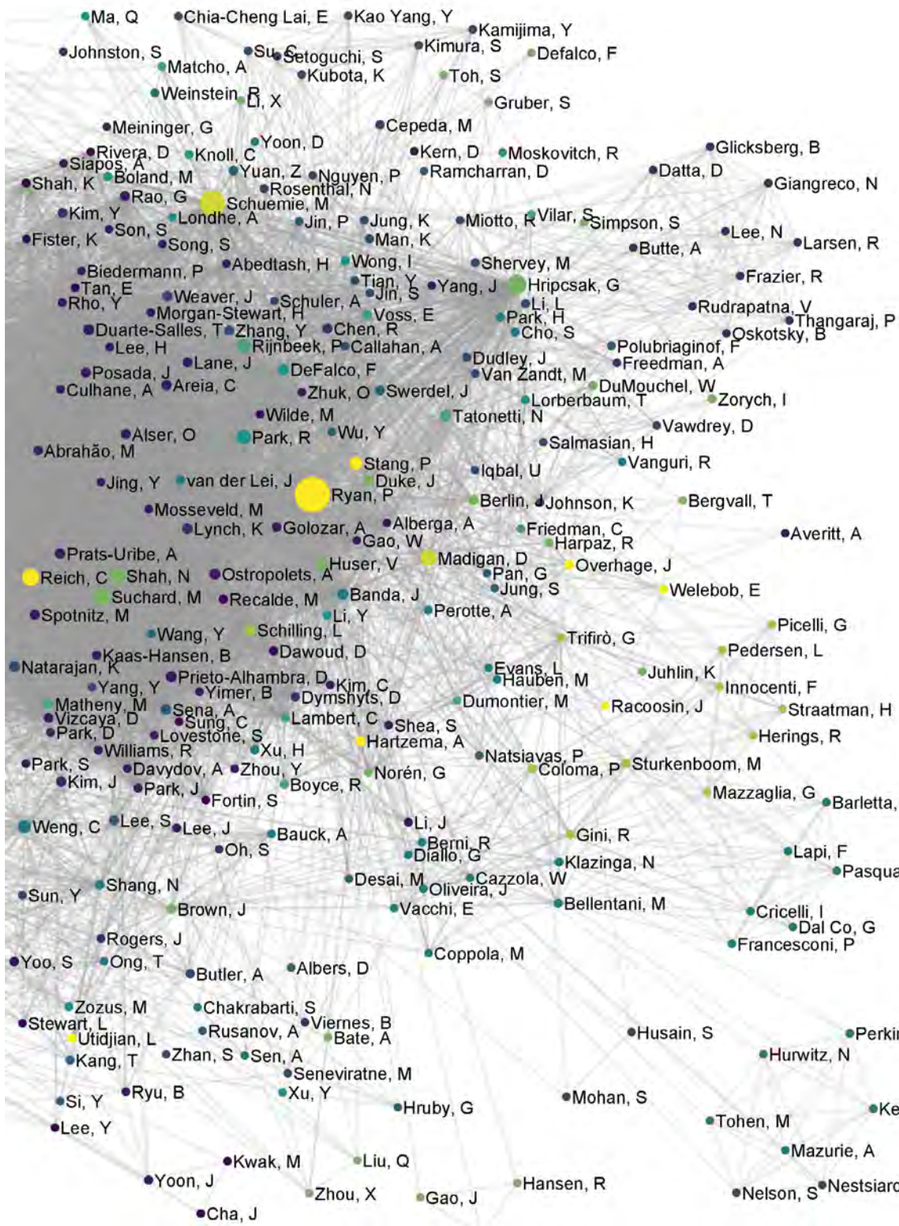
We work together.

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

OHDSI Publications

Legend

Orange.....Data Standards
Red.....Methods Research
Green.....Clinical Applications
Black.....Other

1. Ryan P, Welebob E, Hartzema AG, Stang P, Overhage JM. **Surveying US observational data sources and characteristics for drug safety needs.** Pharmaceutical Medicine. 2010;24(4):231-8. Active Surveillance

2. Stang PE, Ryan PB, Racoosin JA, Overhage JM, Hartzema AG, Reich C, Welebob E, Scarnecchia T, Woodcock J. **Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership.** Ann Intern Med. 2010;153(9):600-6. Epub 2010/11/03. doi: 10.7326/0003-4819-153-9-201011020-00010. PubMed PMID: 21041580. Foundation

3. Madigan D, Ryan P. **What can we really learn from observational studies?: the need for empirical assessment of methodology for active drug safety surveillance and comparative effectiveness research.** Epidemiology. 2011;22(5):629-31. Epub 2011/08/04. doi: 10.1097/EDE.0b013e318228ca1d. PubMed PMID: 21811110. Opinion

4. Schuemie MJ. **Methods for drug safety signal detection in longitudinal observational databases: LGPS and LEOPARD.** Pharmacoepidemiol Drug Saf. 2011;20(3):292-9. Epub 2010/10/15. doi: 10.1002/pds.2051. PubMed PMID: 20945505. Estimation Methods

5. Overhage JM, Ryan PB, Reich CG, Hartzema AG, Stang PE. **Validation of a common data model for active safety surveillance research.** J Am Med Inform Assoc. 2012;19(1):54-60. Epub 2011/11/01. doi: amiajnl-2011-000376 [pii] 10.1136/amiajnl-2011-000376. PubMed PMID: 22037893; PubMed Central PMCID: PMC3240764. Active Surveillance

6. Schuemie MJ, Coloma PM, Straatman H, Herings RM, Trifirò G, Matthews JN, Prieto-Merino D, Molokhia M, Pedersen L, Gini R. **Using electronic health care records for drug safety signal detection: a comparative evaluation of statistical methods.** Medical care. 2012;50(10):890-7. Estimation Methods; Active Surveillance

7. Kahn MG, Raebel MA, Glanz JM, Riedlinger K, Steiner JF. **A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research.** Medical care. 2012;50. Data Quality

8. Carnahan RM, Moores KG. **Mini-Sentinel's systematic reviews of validated methods for identifying health outcomes using administrative and claims data: methods and lessons learned.** Pharmacoepidemiol Drug Saf. 2012;21 Suppl 1:82-9. Epub 2012/01/25. doi: 10.1002/pds.2321. PubMed PMID: 22262596. Phenotyping

9. Stang PE, Ryan PB, Dusetzina SB, Hartzema AG, Reich C, Overhage JM, Racoosin JA. **Health outcomes of interest in observational data: issues in identifying definitions in the literature.** Health Outcomes Research in Medicine. 2012;3(1):e37-e44. Phenotyping

10. Reich C, Ryan PB, Stang PE, Rocca M. **Evaluation of alternative standardized terminologies for medical conditions within a network of observational healthcare databases.** J Biomed Inform. 2012;45(4):689-96. Epub 2012/06/12. doi: S1532-0464(12)00069-X [pii] 10.1016/j.jbi.2012.05.002. PubMed PMID: 22683994. Vocabulary

11. Ryan PB, Madigan D, Stang PE, Overhage JM, Racoosin JA, Hartzema AG. **Empirical assessment of methods for risk identification in healthcare data: results from the experiments of the Observational Medical Outcomes Partnership.** Stat Med. 2012;31(30):4401-15. Epub 2012/09/28. doi: 10.1002/sim.5620. PubMed PMID: 23015364. Estimation Methods

12. Kahn MG, Batson D, Schilling LM. **Data model considerations for clinical effectiveness researchers.** Medical care. 2012;50. ETL

13. Stang PE, Ryan PB, Overhage JM, Schuemie MJ, Hartzema AG, Welebob E. **Variation in choice of study design: findings from the Epidemiology Design Decision Inventory and Evaluation (EDDIE) survey.** Drug Saf. 2013;36 Suppl 1:S15-25. Epub 2013/11/06. doi: 10.1007/s40264-013-0103-1. PubMed PMID: 24166220. Estimation Methods

14. Li X, Hui S, Ryan P, Rosenman M, Overhage M. **Statistical visualization for assessing performance of methods for safety surveillance using electronic databases.** Pharmacoepidemiol Drug Saf. 2013;22(5):503-9. Epub 2013/02/15. doi: 10.1002/pds.3419. PubMed PMID: 23408560. Active Surveillance

15. Schilling LM, Kwan BM, Drolshagen CT, Hosokawa PW, Brandt E, Pace WD, Uhrich C, Kamerick M, Bunting A, Payne PR. **Scalable Architecture for Federated Translational Inquiries Network (SAFTINet) technology infrastructure for a distributed data network.** Egems. 2013;1(1). ETL

16. Schuemie MJ, Gini R, Coloma PM, Straatman H, Herings RM, Pedersen L, Innocenti F, Mazzaglia G, Picelli G, van der Lei J, Sturkenboom MC. **Replication of the OMOP experiment in Europe: evaluating methods for risk identification in electronic health record databases.** Drug Saf. 2013;36 Suppl 1:S159-69. Epub 2013/11/06. doi: 10.1007/s40264-013-0109-8. PubMed PMID: 24166232. Foundation

17. Harpaz R, DuMouchel W, LePendu P, Bauer-Mehren A, Ryan P, Shah NH. **Performance of pharmacovigilance signal-detection algorithms for the FDA adverse event reporting system.** Clin Pharmacol Ther. 2013;93(6):539-46. Epub 2013/04/11. doi: clpt201324 [pii] 10.1038/clpt.2013.24. PubMed PMID: 23571771; PubMed Central PMCID: PMC3857139. Estimation Methods

18. Simpson SE, Madigan D, Zorych I, Schuemie MJ, Ryan PB, Suchard MA. **Multiple self-controlled case series for large-scale longitudinal observational databases.** Biometrics. 2013;69(4):893-902. Epub 2013/10/15. doi: 10.1111/biom.12078. PubMed PMID: 24117144. Estimation Methods

19. Ryan PB, Madigan D, Stang PE, Schuemie MJ, Hripcsak G. **Medication-wide association studies.** CPT Pharmacometrics Syst Pharmacol. 2013;2(9):e76. Epub 2014/01/23. doi: 10.1038/psp.2013.52. PubMed PMID: 24448022; PubMed Central PMCID: PMCPMC4026636. Estimation Methods

20. Suchard MA, Simpson SE, Zorych I, Ryan P, Madigan D. **Massive parallelization of serial inference algorithms for a complex generalized linear model.** ACM Trans Model Comput Simul. 2013;23(1). Epub 2013/01/01. doi: 10.1145/2414416.2414791. PubMed PMID: 25328363; PubMed Central PMCID: PMCPMC4201181. Foundation

21. Hartzema AG, Reich CG, Ryan PB, Stang PE, Madigan D, Welebob E, Overhage JM. **Managing data quality for a drug safety surveillance system.** Drug Saf. 2013;36 Suppl 1:S49-58. Epub 2013/11/06. doi: 10.1007/s40264-013-0098-7. PubMed PMID: 24166223. Data Quality

22. Ryan P, Suchard MA, Schuemie M, Madigan D. **Learning from epidemiology: interpreting observational database studies for the effects of medical products.** Statistics in Biopharmaceutical Research. 2013;5(3):170-9. Estimation Methods

23. Reich CG, Ryan PB, Suchard MA. **The impact of drug and outcome prevalence on the feasibility and performance of analytical methods for a risk identification and analysis system.** Drug Saf. 2013;36 Suppl 1:S195-204. Epub 2013/11/06. doi: 10.1007/s40264-013-0112-0. PubMed PMID: 24166235. Foundation

24. Ogunyemi OI, Meeker D, Kim HE, Ashish N, Farzaneh S, Boxwala A. **Identifying appropriate reference data models for comparative effectiveness research (CER) studies based on data from clinical information systems.** Med Care. 2013;51(8 Suppl 3):S45-52. Epub 2013/06/19. doi: 10.1097/MLR.0b013e31829b1e0b. PubMed PMID: 23774519. Vocabulary

25. Hansen RA, Gray MD, Fox BI, Hollingsworth JC, Gao J, Zeng P. **How well do various health outcome definitions identify appropriate cases in observational studies?** Drug safety. 2013;36(1):27-32. Phenotyping

26. Zhou X, Murugesan S, Bhullar H, Liu Q, Cai B, Wentworth C, Bate A. **An evaluation of the THIN database in the OMOP Common Data Model for active drug safety surveillance.** Drug Saf. 2013;36(2):119-34. Epub 2013/01/19. doi: 10.1007/s40264-012-0009-3. PubMed PMID: 23329543. Active Surveillance, ETL

≤ 2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

OHDSI Publications

27. DuMouchel W, Ryan PB, Schuemie MJ, Madigan D. **Evaluation of disproportionality safety signaling applied to healthcare databases.** Drug Saf. 2013;36 Suppl 1:S123-32. Epub 2013/11/06. doi: 10.1007/s40264-013-0106-y. PubMed PMID: 24166229. Estimation Methods

28. Madigan D, Ryan PB, Schuemie M, Stang PE, Overhage JM, Hartzema AG, Suchard MA, DuMouchel W, Berlin JA. **Evaluating the impact of database heterogeneity on observational study results.** Am J Epidemiol. 2013;178(4):645-51. Epub 2013/05/08. doi: 10.1093/aje/kwt010. PubMed PMID: 23648805; PubMed Central PMCID: PMCPMC3736754. Estimation Methods

29. Ryan PB, Schuemie MJ. **Evaluating performance of risk identification methods through a large-scale simulation of observational data.** Drug Saf. 2013;36 Suppl 1:S171-80. Epub 2013/11/06. doi: 10.1007/s40264-013-0110-2. PubMed PMID: 24166233. Estimation Methods

30. Suchard MA, Zorych I, Simpson SE, Schuemie MJ, Ryan PB, Madigan D. **Empirical performance of the self-controlled case series design: lessons for developing a risk identification and analysis system.** Drug Saf. 2013;36 Suppl 1:S83-93. Epub 2013/11/06. doi: 10.1007/s40264-013-0100-4. PubMed PMID: 24166226. Estimation Methods

31. Madigan D, Schuemie MJ, Ryan PB. **Empirical performance of the case-control method: lessons for developing a risk identification and analysis system.** Drug Saf. 2013;36 Suppl 1:S73-82. Epub 2013/11/06. doi: 10.1007/s40264-013-0105-z. PubMed PMID: 24166225. Estimation Methods

32. Norén GN, Bergvall T, Ryan PB, Juhlin K, Schuemie MJ, Madigan D. **Empirical performance of the calibrated self-controlled cohort analysis within temporal pattern discovery: lessons for developing a risk identification and analysis system.** Drug Saf. 2013;36 Suppl 1:S107-21. Epub 2013/11/06. doi: 10.1007/s40264-013-0095-x. PubMed PMID: 24166228. Estimation Methods

33. Schuemie MJ, Madigan D, Ryan PB. **Empirical performance of LGPS and LEOPARD: lessons for developing a risk identification and analysis system.** Drug Saf. 2013;36 Suppl 1:S133-42. Epub 2013/11/06. doi: 10.1007/s40264-013-0107-x. PubMed PMID: 24166230. Estimation Methods

34. Ryan PB, Schuemie MJ, Madigan D. **Empirical performance of a self-controlled cohort method: lessons for developing a risk identification and analysis system.** Drug Saf. 2013;36 Suppl 1:S95-106. Epub 2013/11/06. doi: 10.1007/s40264-013-0101-3. PubMed PMID: 24166227. Estimation Methods

35. Ryan PB, Schuemie MJ, Gruber S, Zorych I, Madigan D. **Empirical performance of a new user cohort method: lessons for developing a risk identification and analysis system.** Drug Saf. 2013;36 Suppl 1:S59-72. Epub 2013/11/06. doi: 10.1007/s40264-013-0099-6. PubMed PMID: 24166224. Estimation Methods

36. Madigan D, Ryan PB, Schuemie M. **Does design matter? Systematic evaluation of the impact of analytical choices on effect estimates in observational studies.** Ther Adv Drug Saf. 2013;4(2):53-62. Epub 2013/04/01. doi: 10.1177/2042098613477445. PubMed PMID: 25083251; PubMed Central PMCID: PMCPMC4110833. Estimation Methods

37. Zorych I, Madigan D, Ryan P, Bate A. **Disproportionality methods for pharmacovigilance in longitudinal observational databases.** Stat Methods Med Res. 2013;22(1):39-56. Epub 2011/09/01. doi: 10.1177/0962280211403602. PubMed PMID: 21878461. Estimation Methods

38. Overhage JM, Ryan PB, Schuemie MJ, Stang PE. **Desideratum for evidence based epidemiology.** Drug Saf. 2013;36 Suppl 1:S5-14. Epub 2013/11/06. doi: 10.1007/s40264-013-0102-2. PubMed PMID: 24166219. Foundation

39. Huser V, Cimino JJ. **Desiderata for healthcare integrated data repositories based on architectural comparison of three public repositories.** AMIA Annu Symp Proc. 2013;2013:648-56. Epub 2014/02/20. PubMed PMID: 24551366; PubMed Central PMCID: PMC3900207. Vocabulary. ETL

40. Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. **Defining a reference set to support methodological research in drug safety.** Drug Saf. 2013;36 Suppl 1:S33-47. Epub 2013/11/06. doi: 10.1007/s40264-013-0097-8. PubMed PMID: 24166222. Foundation

41. Brown J, Kahn M, Toh S. **Data quality assessment for comparative effectiveness research in distributed data networks.** Medical care. 2013;51(8 0 3):S22. Data Quality

42. Ryan PB, Stang PE, Overhage JM, Suchard MA, Hartzema AG, DuMouchel W, Reich CG, Schuemie MJ, Madigan D. **A comparison of the empirical performance of methods for a risk identification system.** Drug Saf. 2013;36 Suppl 1:S143-58. Epub 2013/11/06. doi: 10.1007/s40264-013-0108-9. PubMed PMID: 24166231. Estimation Methods

43. Gini R, Ryan PB, Brown JS, Vacchi E, Coppola M, Cazzola W, Coloma PM, Berni R, Diallo G, Avillach P. **Comparison Among EU-ADR, OMOP, Mini-Sentinel And MATRICE Strategies For Data Extraction And Management.** Pharmacoepidemiology and Drug Safety. 2013;22:07030-5774. ETL

44. Katz AJ, Ryan PB, Racoosin JA, Stang PE. **Assessment of case definitions for identifying acute liver injury in large observational databases.** Drug Saf. 2013;36(8):651-61. Epub 2013/05/15. doi: 10.1007/s40264-013-0060-8. PubMed PMID: 23670723. Phenotyping

45. Defalco FJ, Ryan PB, Soledad Cepeda M. **Applying standardized drug terminologies to observational healthcare databases: a case study on opioid exposure.** Health Serv Outcomes Res Methodol. 2013;13(1):58-67. Epub 2013/02/12. doi: 10.1007/s10742-012-0102-1 102 [pii]. PubMed PMID: 23396660; PubMed Central PMCID: PMC3566397. Vocabulary, Estimation Study

46. Reich CG, Ryan PB, Schuemie MJ. **Alternative outcome definitions and their effect on the performance of methods for observational outcome studies.** Drug Saf. 2013;36 Suppl 1:S181-93. Epub 2013/11/06. doi: 10.1007/s40264-013-0111-1. PubMed PMID: 24166234. Phenotyping; Estimation Methods

47. Lian Duan L, Khoshneshin M, Street WN, Liu M. **Adverse drug effect detection.** IEEE J Biomed Health Inform. 2013;17(2):305-11. Epub 2013/11/16. doi: 10.1109/TITB.2012.2227272. PubMed PMID: 24235108. Estimation Methods

48. Schuemie MJ, Ryan PB, DuMouchel W, Suchard MA, Madigan D. **Interpreting observational studies: why empirical calibration is needed to correct p-values.** Stat Med. 2014;33(2):209-18. Epub 2013/08/01. doi: 10.1002/sim.5925. PubMed PMID: 23900808; PubMed Central PMCID: PMCPMC4285234. Foundation; Estimation Methods

49. Makadia R, Ryan PB. **Transforming the Premier Perspective Hospital Database into the Observational Medical Outcomes Partnership (OMOP) Common Data Model.** EGEMS (Wash DC). 2014;2(1):1110. Epub 2014/01/01. doi: 10.13063/2327-9214.1110 egems1110 [pii]. PubMed PMID: 25848597; PubMed Central PMCID: PMC4371500. ETL

50. Vilar S, Ryan PB, Madigan D, Stang PE, Schuemie MJ, Friedman C, Tatonetti NP, Hripcsak G. **Similarity-based modeling applied to signal detection in pharmacovigilance.** CPT Pharmacometrics Syst Pharmacol. 2014;3:e137. Epub 2014/09/25. doi: psp201435 [pii] 10.1038/psp.2014.35. PubMed PMID: 25250527; PubMed Central PMCID: PMC4211266. Estimation Methods

51. Ohno-Machado L, Agha Z, Bell DS, Dahm L, Day ME, Doctor JN, Gabriel D, Kahlon MK, Kim KK, Hogarth M. **pSCANNER: patient-centered scalable national network for effectiveness research.** Journal of the American Medical Informatics Association. 2014;21(4):621-6. ETL

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

51. Forrest CB, Margolis PA, Bailey LC, Marsolo K, Del Beccaro MA, Finkelstein JA, Milov DE, Vieland VJ, Wolf BA, Yu FB. **PEDSnet: a national pediatric learning health system.** Journal of the American Medical Informatics Association. 2014;21(4):602-6. ETL

53. Matcho A, Ryan P, Fife D, Reich C. **Fidelity assessment of a clinical practice research datalink conversion to the OMOP common data model.** Drug Saf. 2014;37(11):945-59. Epub 2014/09/05. doi: 10.1007/s40264-014-0214-3. PubMed PMID: 25187016; PubMed Central PMCID: PMC4206771. ETL

54. Fife D, Zhu V, Voss E, Levy-Clarke G, Ryan P. **Exposure to oral fluoroquinolones and the risk of retinal detachment: retrospective analyses of two large healthcare databases.** Drug Saf. 2014;37(3):171-82. Epub 2014/02/15. doi: 10.1007/s40264-014-0138-y. PubMed PMID: 24526267; PubMed Central PMCID: PMCP-MC3936132. Estimation Study

55. Hansen RA, Zeng P, Ryan P, Gao J, Sonawane K, Teeter B, Westrich K, Dubois RW. **Exploration of heterogeneity in distributed research network drug safety analyses.** Res Synth Methods. 2014;5(4):352-70. Epub 2015/06/09. doi: 10.1002/jrsm.1121. PubMed PMID: 26052957. Estimation Methods

56. Stang P, Ryan P, Hartzema AG, Madigan D, Marc Overhage J, Welebob E, Reich CG, Scarnecchia T. **Development and evaluation of infrastructure and analytic methods for systematic drug safety surveillance: lessons and resources from the observational medical outcomes partnership.** Mann's Pharmacovigilance. 2014:453-61.Active Surveillance

57. Pace WD, Fox CH, Turner White DG, Schilling LM, West DR. **The DARTNet Institute: seeking a sustainable support mechanism for electronic data enabled research networks.** Egems. 2014;2(2). ETL

58. Rijnbeek PR. **Converting to a common data model: what is lost in translation?: Commentary on "fidelity assessment of a clinical practice research datalink conversion to the OMOP common data model".** Drug Saf. 2014;37(11):893-6. Epub 2014/09/05. doi: 10.1007/s40264-014-0221-4. PubMed PMID: 25187018. ETL

59. Trifiro G, Coloma PM, Rijnbeek PR, Romio S, Mosseveld B, Weibel D, Bonhoeffer J, Schuemie M, van der Lei J, Sturkenboom M. **Combining multiple healthcare databases for postmarketing drug and vaccine safety surveillance: why and how?** J Intern Med. 2014;275(6):551-61. Epub 2014/03/19. doi: 10.1111/joim.12159. PubMed PMID: 24635221. ETL

60. Boyce RD, Ryan PB, Noren GN, Schuemie MJ, Reich C, Duke J, Tatonetti NP, Trifiro G, Harpaz R, Overhage JM, Hartzema AG, Khayter M, Voss EA, Lambert CG, Huser V, Dumontier M. **Bridging islands of information to establish an integrated knowledge base of drugs and health outcomes of interest.** Drug Saf. 2014;37(8):557-67. Epub 2014/07/06. doi: 10.1007/s40264-014-0189-0. PubMed PMID: 24985530; PubMed Central PMCID: PMC4134480. Methods Other

61. McCormick TH, Ferrell R, Karr AF, Ryan PB. **Big data, big results: Knowledge discovery in output from large-scale analytics.** Statistical Analysis and Data Mining: the ASA Data Science Journal. 2014;7(5):404-12. Estimation Methods

62. Kahn MG, Brown JS, Chun AT, Davidson BN, Meeker D, Ryan PB, Schilling LM, Weiskopf NG, Williams AE, Zozus MN. **Transparent reporting of data quality in distributed data networks.** EGEMS (Wash DC). 2015;3(1):1052. Epub 2015/05/21. doi: 10.13063/2327-9214.1052. PubMed PMID: 25992385; PubMed Central PMCID: PMCPMC4434997. Data Quality

63. Voss EA, Ryan PB, Stang PE, Hough D, Alphs L. **Switching from risperidone long-acting injectable to paliperidone long-acting injectable or oral antipsychotics: analysis of a Medicaid claims database.** Int Clin Psychopharmacol. 2015;30(3):151-7. Epub 2015/03/03. doi: 10.1097/yic.0000000000000068. PubMed PMID: 25730525; PubMed Central PMCID: PMCPMC4383368. Estimation Study

64. Ling W. **Study on application of data organization and common data model of the observational medical outcomes partnership in US.** Chinese Journal of Pharmacovigilance. 2015;12(6):341. ETL

65. Sun H, Depraetere K, De Roo J, Mels G, De Vloed B, Twagirumukiza M, Colaert D. **Semantic processing of EHR data for clinical research.** J Biomed Inform. 2015;58:247-59. Epub 2015/10/31. doi: S1532-0464(15)00231-2 [pii] 10.1016/j.jbi.2015.10.009. PubMed PMID: 26515501.Referring to

66. Shahn Z, Ryan P, Madigan D. **Predicting health outcomes from high-dimensional longitudinal health histories using relational random forests.** Statistical Analysis and Data Mining: the ASA Data Science Journal. 2015;8(2):128-36. Prediction Methods

67. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, Suchard MA, Park RW, Wong IC, Rijnbeek PR, van der Lei J, Pratt N, Noren GN, Li YC, Stang PE, Madigan D, Ryan PB. **Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers.** Stud Health Technol Inform. 2015;216:574-8. Epub 2015/08/12. PubMed PMID: 26262116; PubMed Central PMCID: PMC4815923. Foundation

68. Li Y, Ryan PB, Wei Y, Friedman C. **A Method to Combine Signals from Spontaneous Reporting Systems and Observational Healthcare Data to Detect Adverse Drug Reactions.** Drug Saf. 2015;38(10):895-908. Epub 2015/07/15. doi: 10.1007/s40264-015-0314-8 10.1007/s40264-015-0314-8 [pii]. PubMed PMID: 26153397; PubMed Central PMCID: PMC4579260. Estimation Methods

69. Vilar S, Lorberbaum T, Hripcsak G, Tatonetti NP. **Improving Detection of Arrhythmia Drug-Drug Interactions in Pharmacovigilance Data through the Implementation of Similarity-Based Modeling.** PLoS One. 2015;10(6):e0129974. Epub 2015/06/13. doi: 10.1371/journal.pone.0129974 PONE-D-15-04363 [pii]. PubMed PMID: 26068584; PubMed Central PMCID: PMC4466327. Estimation Methods

70. Voss EA, Ma Q, Ryan PB. **The impact of standardizing the definition of visits on the consistency of multi-database observational health research.** BMC Med Res Methodol. 2015;15:13. Epub 2015/04/19. doi: 10.1186/s12874-015-0001-6 10.1186/s12874-015-0001-6 [pii]. PubMed PMID: 25887092; PubMed Central PMCID: PMC4369827.Vocabulary

71. Voss EA, Makadia R, Matcho A, Ma Q, Knoll C, Schuemie M, DeFalco FJ, Londhe A, Zhu V, Ryan PB. **Feasibility and utility of applications of the common data model to multiple, disparate observational health databases.** J Am Med Inform Assoc. 2015;22(3):553-64. Epub 2015/02/12. doi: ocu023 [pii] 10.1093/jamia/ocu023. PubMed PMID: 25670757; PubMed Central PMCID: PMC4457111.Foundation

72. Boland MR, Tatonetti NP, Hripcsak G. **Development and validation of a classification approach for extracting severity automatically from electronic health records.** J Biomed Semantics. 2015;6:14. Epub 2015/04/08. doi: 10.1186/s13326-015-0010-8 10 [pii]. PubMed PMID: 25848530; PubMed Central PMCID: PMC4386082. Prediction Study

73. FitzHenry F, Resnic FS, Robbins SL, Denton J, Nookala L, Meeker D, Ohno-Machado L, Matheny ME. **Creating a Common Data Model for Comparative Effectiveness with the Observational Medical Outcomes Partnership.** Appl Clin Inform. 2015;6(3):536-47. Epub 2015/10/09. doi: 10.4338/ACI-2014-12-CR-0121. PubMed PMID: 26448797; PubMed Central PMCID: PMC4586341. ETL

Legend

Orange.....Data Standards
Red.....Methods Research
Green.....Clinical Applications
Black.....Other

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

74. Xu Y, Zhou X, Suehs BT, Hartzema AG, Kahn MG, Moride Y, Sauer BC, Liu Q, Moll K, Pasquale MK, Nair VP, Bate A. **A Comparative Assessment of Observational Medical Outcomes Partnership and Mini-Sentinel Common Data Models and Analytics: Implications for Active Drug Safety Surveillance.** Drug Saf. 2015;38(8):749-65. Epub 2015/06/10. doi: 10.1007/s40264-015-0297-5 10.1007/s40264-015-0297-5 [pii]. PubMed PMID: 26055920. Foundation, Active Surveillance

75. Khalilia M, Choi M, Henderson A, Iyengar S, Braunstein M, Sun J. **Clinical Predictive Modeling Development and Deployment through FHIR Web Services.** AMIA Annu Symp Proc. 2015;2015:717-26. Epub 2015/01/01. PubMed PMID: 26958207; PubMed Central PMCID: PMC4765683. FHIR

76. Boland MR, Shahn Z, Madigan D, Hripcsak G, Tatonetti NP. **Birth month affects lifetime disease risk: a phenome-wide method.** J Am Med Inform Assoc. 2015;22(5):1042-53. Epub 2015/06/05. doi: ocv046 [pii] 10.1093/jamia/ocv046. PubMed PMID: 26041386; PubMed Central PMCID: PMC4986668. Estimation Study

77. Garbe E, Pigeot I. **[Benefits of large healthcare databases for drug risk research] Der Nutzen grosser Gesundheitsdatenbanken fur die Arzneimittelrisikoforschung.** Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2015;58(8):829-37. Epub 2015/06/21. doi: 10.1007/s00103-015-2185-7 10.1007/s00103-015-2185-7 [pii]. PubMed PMID: 26092163.Referring to

78. Boland MR, Jacunski A, Lorberbaum T, Romano JD, Moskovitch R, Tatonetti NP. **Systems biology approaches for identifying adverse drug reactions and elucidating their underlying biological mechanisms.** Wiley Interdiscip Rev Syst Biol Med. 2016;8(2):104-22. Epub 2015/11/13. doi: 10.1002/wsbm.1323. PubMed PMID: 26559926; PubMed Central PMCID: PMC4760887. Estimation Methods

79. Weinstein RB, Schuemie MJ, Ryan PB, Stang PE. **Seasonality in acute liver injury? Findings in two health care claims databases.** Drug Healthc Patient Saf. 2016;8:39-48. Epub 2016/04/22. doi: 10.2147/dhps.S95399. PubMed PMID: 27099532; PubMed Central PMCID: PMCPMC4824282. Characterization Study

80. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. **Robust empirical calibration of p-values using observational data.** Stat Med. 2016;35(22):3883-8. Epub 2016/09/07. doi: 10.1002/sim.6977. PubMed PMID: 27592566; PubMed Central PMCID: PMC5108459. Estimation Methods

81. Boyce RD, Handler SM, Karp JF, Perera S, Reynolds CF, 3rd. **Preparing Nursing Home Data from Multiple Sites for Clinical Research - A Case Study Using Observational Health Data Sciences and Informatics.** EGEMS (Wash DC). 2016;4(1):1252. Epub 2016/11/29. doi: 10.13063/2327-9214.1252 egems1252 [pii]. PubMed PMID: 27891528; PubMed Central PMCID: PMC5108634. ETL

82. Huser V, DeFalco FJ, Schuemie M, Ryan PB, Shang N, Velez M, Park RW, Boyce RD, Duke J, Khare R, Utidjian L, Bailey C. **Multisite Evaluation of a Data Quality Tool for Patient-Level Clinical Data Sets.** EGEMS (Wash DC). 2016;4(1):1239. Epub 2017/02/06. doi: 10.13063/2327-9214.1239 egems1239 [pii]. PubMed PMID: 28154833; PubMed Central PMCID: PMC5226382. Data Quality

83. Gruber S, Tchetgen Tchetgen E. **Limitations of empirical calibration of p-values using observational data.** Stat Med. 2016;35(22):3869-82. Epub 2016/03/13. doi: 10.1002/sim.6936. PubMed PMID: 26970249; PubMed Central PMCID: PMC5012943. Estimation Methods

84. Agarwal V, Podchiyska T, Banda JM, Goel V, Leung TI, Minty EP, Sweeney TE, Gyang E, Shah NH. **Learning statistical models of phenotypes using noisy labeled training data.** Journal of the American Medical Informatics Association. 2016;23(6):1166-73. Phenotyping

85. Samwald M, Xu H, Blagec K, Empey PE, Malone DC, Ahmed SM, Ryan P, Hofer S, Boyce RD. **Incidence of Exposure of Patients in the United States to Multiple Drugs for Which Pharmacogenomic Guidelines Are Available.** PLoS One. 2016;11(10):e0164972. Epub 2016/10/21. doi: 10.1371/journal.pone.0164972. PubMed PMID: 27764192; PubMed Central PMCID: PMCPMC5072717 adherence to PLOS ONE policies on sharing data and materials. Characterization Study

86. Lambert CG, Mazurie AJ, Lauve NR, Hurwitz NG, Young SS, Obenchain RL, Hengartner NW, Perkins DJ, Tohen M, Kerner B. **Hypothyroidism risk compared among nine common bipolar disorder therapies in a large US cohort.** Bipolar disorders. 2016;18(3):247-60. Estimation Study

87. Shaddox TR, Ryan PB, Schuemie MJ, Madigan D, Suchard MA. **Hierarchical Models for Multiple, Rare Outcomes Using Massive Observational Healthcare Databases.** Stat Anal Data Min. 2016;9(4):260-8. Epub 2017/05/16. doi: 10.1002/sam.11324. PubMed PMID: 28503249; PubMed Central PMCID: PMCPMC5423675. Estimation Methods

88. Kahn MG, Callahan TJ, Barnard J, Bauck AE, Brown J, Davidson BN, Estiri H, Goerg C, Holve E, Johnson SG. **A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data.** Egems. 2016;4(1). Data Quality

89. Sen A, Chakrabarti S, Goldstein A, Wang S, Ryan PB, Weng C. **GIST 2.0: A scalable multi-trait metric for quantifying population representativeness of individual clinical studies.** J Biomed Inform. 2016;63:325-36. Epub 2016/10/25. doi: S1532-0464(16)30115-0 [pii] 10.1016/j.jbi.2016.09.003. PubMed PMID: 27600407; PubMed Central PMCID: PMC5077682. Estimation Methods

90. Kim H, Choi J, Jang I, Quach J, Ohno-Machado L. **Feasibility of Representing Data from Published Nursing Research Using the OMOP Common Data Model.** AMIA Annu Symp Proc. 2016;2016:715-23. Epub 2017/03/09. PubMed PMID: 28269868; PubMed Central PMCID: PMC5333244. Vocabulary

91. Hauben M, Aronson JK, Ferner RE. **Evidence of Misclassification of Drug-Event Associations Classified as Gold Standard 'Negative Controls' by the Observational Medical Outcomes Partnership (OMOP).** Drug Saf. 2016;39(5):421-32. Epub 2016/02/18. doi: 10.1007/s40264-016-0392-2 10.1007/s40264-016-0392-2 [pii]. PubMed PMID: 26879560. Estimation Methods

92. Garza M, Del Fiol G, Tenenbaum J, Walden A, Zozus MN. **Evaluating common data models for use with a longitudinal community registry.** J Biomed Inform. 2016;64:333-41. Epub 2016/12/19. doi: S1532-0464(16)30153-8 [pii] 10.1016/j.jbi.2016.10.016. PubMed PMID: 27989817; PubMed Central PMCID: PMC6810649.Vocabulary

93. Schuemie MJ, Trifirò G, Coloma PM, Ryan PB, Madigan D. **Detecting adverse drug reactions following long-term exposure in longitudinal observational data: The exposure-adjusted self-controlled case series.** Stat Methods Med Res. 2016;25(6):2577-92. Epub 2014/04/02. doi: 10.1177/0962280214527531. PubMed PMID: 24685766. Estimation Methods

94. Gruber S, Chakravarty A, Heckbert SR, Levenson M, Martin D, Nelson JC, Psaty BM, Pinheiro S, Reich CG, Toh S, Walker AM. **Design and analysis choices for safety surveillance evaluations need to be tuned to the specifics of the hypothesized drug-outcome association.** Pharmacoepidemiol Drug Saf. 2016;25(9):973-81. Epub 2016/07/16. doi: 10.1002/pds.4065. PubMed PMID: 27418432. Estimation Methods

95. Gini R, Schuemie M, Brown J, Ryan P, Vacchi E, Coppola M, Cazzola W, Coloma P, Berni R, Diallo G, Oliveira JL, Avillach P, Trifiro G, Rijnbeek P, Bellentani M, van Der Lei J, Klazinga N, Sturkenboom M. **Data Extraction and Management in Networks of Observational Health Care Databases for Scientific Research: A Comparison of EU-ADR, OMOP, Mini-Sentinel and MATRICE Strategies.** EGEMS (Wash DC). 2016;4(1):1189. Epub 2016/03/26. doi: 10.13063/2327-9214.1189 egems1189 [pii]. PubMed PMID: 27014709; PubMed Central PMCID: PMC4780748. Referring to

96. Banda JM, Evans L, Vanguri RS, Tatonetti NP, Ryan PB, Shah NH. **A curated and standardized adverse drug event resource to accelerate drug safety research.** Sci Data. 2016;3:160026. Epub 2016/05/20. doi: 10.1038/sdata.2016.26. PubMed PMID: 27193236; PubMed Central PMCID: PMCPMC4872271. ETL; NLP

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

97. Yoon D, Ahn EK, Park MY, Cho SY, Ryan P, Schuemie MJ, Shin D, Park H, Park RW. **Conversion and Data Quality Assessment of Electronic Health Record Data at a Korean Tertiary Teaching Hospital to a Common Data Model for Distributed Network Research.** Healthc Inform Res. 2016;22(1):54-8. Epub 2016/02/20. doi: 10.4258/hir.2016.22.1.54. PubMed PMID: 26893951; PubMed Central PMCID: PMC4756059. ETL, Data Quality

98. Hripcsak G, Ryan PB, Duke JD, Shah NH, Park RW, Huser V, Suchard MA, Schuemie MJ, DeFalco FJ, Perotte A, Banda JM, Reich CG, Schilling LM, Matheny ME, Meeker D, Pratt N, Madigan D. **Characterizing treatment pathways at scale using the OHDSI network.** Proc Natl Acad Sci U S A. 2016;113(27):7329-36. Epub 2016/06/09. doi: 1510502113 [pii] 10.1073/pnas.1510502113. PubMed PMID: 27274072; PubMed Central PMCID: PMC4941483. Characterization Methods

99. Gini R, Schuemie MJ, Mazzaglia G, Lapi F, Francesconi P, Pasqua A, Bianchini E, Montalbano C, Roberto G, Barletta V, Cricelli I, Cricelli C, Dal Co G, Bellentani M, Sturkenboom M, Klazinga N. **Automatic identification of type 2 diabetes, hypertension, ischaemic heart disease, heart failure and their levels of severity from Italian General Practitioners' electronic medical records: a validation study.** BMJ open. 2016;6(12):e012413. Epub 2016/12/13. doi: bmjopen-2016-012413 [pii] 10.1136/bmjopen-2016-012413. PubMed PMID: 27940627; PubMed Central PMCID: PMC5168667. Referring to

100. Harpaz R, DuMouchel W, Schuemie M, Bodenreider O, Friedman C, Horvitz E, Ripple A, Sorbello A, White RW, Winnenburg R. **Toward multimodal signal detection of adverse drug reactions.** Journal of biomedical informatics. 2017;76:41-9. Active Surveillance

101. Park RW. **Sharing clinical big data while protecting confidentiality and security: observational health data sciences and informatics.** Healthcare informatics research. 2017;23(1):1-3. ETL

102. Reisinger S, McDonald L, Carroll R, O'Hara D, Anstatt D, Ramagopalan S. **A Robust, Reproducible Method For Evaluating The Suitability of Disparate Observational Databases for Pooled Analysis, Using The Omop Common Data Model.** Value in Health. 2017;20(9):A776. Characterization Methods

103. Yuan Z, Voss EA, DeFalco FJ, Pan G, Ryan PB, Yannicelli D, Nessel C. **Risk Prediction for Ischemic Stroke and Transient Ischemic Attack in Patients Without Atrial Fibrillation: A Retrospective Cohort Study.** J Stroke Cerebrovasc Dis. 2017;26(8):1721-31. Epub 2017/04/11. doi: 10.1016/j.jstrokecerebrovasdis.2017.03.036. PubMed PMID: 28392100. Prediction Study

104. Duke JD, Ryan PB, Suchard MA, Hripcsak G, Jin P, Reich C, Schwalm MS, Khoma Y, Wu Y, Xu H, Shah NH, Banda JM, Schuemie MJ. **Risk of angioedema associated with levetiracetam compared with phenytoin: Findings of the observational health data sciences and informatics research network.** Epilepsia. 2017;58(8):e101-e6. Epub 2017/07/07. doi: 10.1111/epi.13828. PubMed PMID: 28681416; PubMed Central PMCID: PMC6632067. Estimation Study

105. Rosenbloom ST, Carroll RJ, Warner JL, Matheny ME, Denny JC. **Representing Knowledge Consistently Across Health Systems.** Yearb Med Inform. 2017;26(1):139-47. Epub 2017/10/25. doi: 10.15265/IY-2017-018. PubMed PMID: 29063555; PubMed Central PMCID: PMC6239235. Referring to

106. Callahan T, Barnard J, Helmkamp L, Maertens J, Kahn M. **Reporting data quality assessment results: identifying individual and organizational barriers and solutions.** Egems. 2017;5(1). Data Quality

107. Ceusters W, Blaisure J. **A Realism-Based View on Counts in OMOP's Common Data Model.** Stud Health Technol Inform. 2017;237:55-62. Epub 2017/05/10. PubMed PMID: 28479543. Data Quality

108. Moskovitch R, Polubriaginof F, Weiss A, Ryan P, Tatonetti N. **Procedure prediction from symbolic Electronic Health Records via time intervals analytics.** J Biomed Inform. 2017;75:70-82. Epub 2017/08/22. doi: S1532-0464(17)30178-8 [pii] 10.1016/j.jbi.2017.07.018. PubMed PMID: 28823923. Prediction Methods

109. Si Y, Weng C. **An OMOP CDM-Based Relational Database of Clinical Research Eligibility Criteria.** Stud Health Technol Inform. 2017;245:950-4. Epub 2018/01/04. PubMed PMID: 29295240; PubMed Central PMCID: PMC5893219.NLP,Trial

110. Gini R, Schuemie MJ, Pasqua A, Carlini E, Profili F, Cricelli I, Dazzi P, Barletta V, Francesconi P, Lapi F, Donatini A, Dal Co G, Visca M, Bellentani M, Sturkenboom M, Klazinga N. **Monitoring compliance with standards of care for chronic diseases using healthcare administrative databases in Italy: Strengths and limitations.** PLoS One. 2017;12(12):e0188377. Epub 2017/12/13. doi: 10.1371/journal.pone.0188377 PONE-D-16-51291 [pii]. PubMed PMID: 29232365; PubMed Central PMCID: PMC5726627. Estimation Study

111. Dixon BE, Duke J, Grannis S. **Measuring and Improving the Quality of Data Used for Syndromic Surveillance.** Online Journal of Public Health Informatics. 2017;9 (1). Data Quality

112. Khare R, Utidjian L, Ruth BJ, Kahn MG, Burrows E, Marsolo K, Patibandla N, Razzaghi H, Colvin R, Ranade D. **A longitudinal analysis of data quality in a large pediatric data research network.** Journal of the American Medical Informatics Association. 2017;24(6):1072-9.Data Quality

113. Knowledge Base workgroup of the Observational Health Data S, Informatics c. **Large-scale adverse effects related to treatment evidence standardization (LAERTES): an open scalable system for linking pharmacovigilance evidence sources with clinical data.** J Biomed Semantics. 2017;8(1):11. Epub 2017/03/09. doi: 10.1186/s13326-017-0115-3 10.1186/s13326-017-0115-3 [pii]. PubMed PMID: 28270198; PubMed Central PMCID: PMC5341176. Estimation Methods

114. Chakrabarti S, Sen A, Huser V, Hruby GW, Rusanov A, Albers DJ, Weng C. **An Interoperable Similarity-based Cohort Identification Method Using the OMOP Common Data Model version 5.0.** J Healthc Inform Res. 2017;1(1):1-18. Epub 2017/08/05. doi: 10.1007/s41666-017-0005-6. PubMed PMID: 28776047; PubMed Central PMCID: PMC5536903. Phenotyping

115. Wang Y, Desai M, Ryan PB, DeFalco FJ, Schuemie MJ, Stang PE, Berlin JA, Yuan Z. **Incidence of diabetic ketoacidosis among patients with type 2 diabetes mellitus treated with SGLT2 inhibitors and other antihyperglycemic agents.** Diabetes Res Clin Pract. 2017;128:83-90. Epub 2017/04/28. doi: 10.1016/j.diabres.2017.04.004. PubMed PMID: 28448895.Estimation Study

116. Kang T, Zhang S, Tang Y, Hruby GW, Rusanov A, Elhadad N, Weng C. **EliIE: An open-source information extraction system for clinical trial eligibility criteria.** J Am Med Inform Assoc. 2017;24(6):1062-71. Epub 2017/04/06. doi: 3098256 [pii] 10.1093/jamia/ocx019. PubMed PMID: 28379377; PubMed Central PMCID: PMC6259668. Trial

117. Banda JM, Halpern Y, Sontag D, Shah NH. **Electronic phenotyping with APHRODITE and the Observational Health Sciences and Informatics (OHDSI) data network.** AMIA Jt Summits Transl Sci Proc. 2017;2017:48-57. Epub 2017/08/18. PubMed PMID: 28815104; PubMed Central PMCID: PMC5543379. Phenotyping

118. Ong TC, Kahn MG, Kwan BM, Yamashita T, Brandt E, Hosokawa P, Uhrich C, Schilling LM. Dynamic. **ETL: a hybrid approach for health data extraction, transformation and loading.** BMC Med Inform Decis Mak. 2017;17(1):134. Epub 2017/09/15. doi: 10.1186/s12911-017-0532-3 10.1186/s12911-017-0532-3 [pii]. PubMed PMID: 28903729; PubMed Central PMCID: PMC5598056. ETL

Legend

Orange	Data Standards
Red	Methods Research
Green	Clinical Applications
Black	Other

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

OHDSI PUBLICATIONS

119. You SC, Lee S, Cho SY, Park H, Jung S, Cho J, Yoon D, Park RW. **Conversion of National Health Insurance Service-National Sample Cohort (NHIS-NSC) Database into Observational Medical Outcomes Partnership-Common Data Model (OMOP-CDM).** Stud Health Technol Inform. 2017;245:467-70. Epub 2018/01/04. PubMed PMID: 29295138. ETL

120. Schwalm M, Raoul T, Chu D, Shah U, Potdar M, Van Zandt M, Coffin G, Jouaville S. **Conversion of a French Electronic Medical Record (Emr) Database into the Observational Medical Outcomes Partnership Common Data Model.** Value in Health. 2017;20(9):A741. ETL

121. Jiang G, Kiefer RC, Sharma DK, Prud'hommeaux E, Solbrig HR. **A Consensus-Based Approach for Harmonizing the OHDSI Common Data Model with HL7 FHIR.** Stud Health Technol Inform. 2017;245:887-91. Epub 2018/01/04. PubMed PMID: 29295227; PubMed Central PMCID: PMC5939955. FHIR

122. Callahan TJ, Bauck AE, Bertoch D, Brown J, Khare R, Ryan PB, Staab J, Zozus MN, Kahn MG. **A Comparison of Data Quality Assessment Checks in Six Data Sharing Networks.** EGEMS (Wash DC). 2017;5(1):8. Epub 2018/06/09. doi: 10.5334/egems.223. PubMed PMID: 29881733; PubMed Central PMCID: PMCPMC5982846. Data Quality

123. Weinstein RB, Ryan P, Berlin JA, Matcho A, Schuemie M, Swerdel J, Patel K, Fife D. **Channeling in the Use of Nonprescription Paracetamol and Ibuprofen in an Electronic Medical Records Database: Evidence and Implications.** Drug Saf. 2017;40(12):1279-92. Epub 2017/08/07. doi: 10.1007/s40264-017-0581-7. PubMed PMID: 28780741; PubMed Central PMCID: PMCPMC5688206. Estimation Methods

124. Jiang G, Kiefer R, Prud'hommeaux E, Solbrig HR. **Building Interoperable FHIR-Based Vocabulary Mapping Services: A Case Study of OHDSI Vocabularies and Mappings.** Stud Health Technol Inform. 2017;245:1327. Epub 2018/01/04. PubMed PMID: 29295408; PubMed Central PMCID: PMC5939959.FHIR

125. Ryan PB, Schuemie MJ, Ramcharran D, Stang PE. **Atypical Antipsychotics and the Risks of Acute Kidney Injury and Related Outcomes Among Older Adults: A Replication Analysis and an Evaluation of Adapted Confounding Control Strategies.** Drugs Aging. 2017;34(3):211-9. Epub 2017/01/27. doi: 10.1007/s40266-016-0430-x. PubMed PMID: 28124262. Estimation Study

126. Ramcharran D, Qiu H, Schuemie MJ, Ryan PB. **Atypical Antipsychotics and the Risk of Falls and Fractures Among Older Adults: An Emulation Analysis and an Evaluation of Additional Confounding Control Strategies.** J Clin Psychopharmacol. 2017;37(2):162-8. Epub 2017/02/23. doi: 10.1097/jcp.0000000000000647. PubMed PMID: 28225746. Estimation Study

127. Voss EA, Boyce RD, Ryan PB, van der Lei J, Rijnbeek PR, Schuemie MJ. **Accuracy of an automated knowledge base for identifying drug adverse reactions.** J Biomed Inform. 2017;66:72-81. Epub 2016/12/21. doi: S1532-0464(16)30179-4 [pii] 10.1016 j.jbi.2016.12.005. PubMed PMID: 27993747; PubMed Central PMCID: PMC5316295. Estimation Methods

128. Klann JG, Phillips LC, Herrick C, Joss MAH, Wagholikar KB, Murphy SN. **Web services for data warehouses: OMOP and PCORnet on i2b2.** J Am Med Inform Assoc. 2018;25(10):1331-8. Epub 2018/08/08. doi: 5061849 [pii] 10.1093/jamia/ocy093. PubMed PMID: 30085008; PubMed Central PMCID: PMC6188504. ETL

129. Boland MR, Parhi P, Li L, Miotto R, Carroll R, Iqbal U, Nguyen PA, Schuemie M, You SC, Smith D, Mooney S, Ryan P, Li YJ, Park RW, Denny J, Dudley JT, Hripcsak G, Gentine P, Tatonetti NP. **Uncovering exposures responsible for birth season - disease effects: a global study.** J Am Med Inform Assoc. 2018;25(3):275-88. Epub 2017/10/17. doi: 10.1093/jamia/ocx105. PubMed PMID: 29036387; PubMed Central PMCID: PMCPMC7282503. Characterization Study

130. Maier C, Lang L, Storf H, Vormstein P, Bieber R, Bernarding J, Herrmann T, Haverkamp C, Horki P, Laufer J, Berger F, Honing G, Fritsch HW, Schuttler J, Ganslandt T, Prokosch HU, Sedlmayr M. **Towards Implementation of OMOP in a German University Hospital Consortium.** Appl Clin Inform. 2018;9(1):54-61. Epub 2018/01/25. doi: 10.1055/s-0037-1617452. PubMed PMID: 29365340; PubMed Central PMCID: PMC5801887. ETL

131. Pacaci A, Gonul S, Sinaci AA, Yuksel M, Laleci Erturkmen GB. **A Semantic Transformation Methodology for the Secondary Use of Observational Healthcare Data in Postmarketing Safety Studies.** Front Pharmacol. 2018;9:435. Epub 2018/05/16. doi: 10.3389/fphar.2018.00435. PubMed PMID: 29760661; PubMed Central PMCID: PMC5937227. ETL

132. Yuan Z, DeFalco FJ, Ryan PB, Schuemie MJ, Stang PE, Berlin JA, Desai M, Rosenthal N. **Risk of lower extremity amputations in people with type 2 diabetes mellitus treated with sodium-glucose co-transporter-2 inhibitors in the USA: A retrospective cohort study.** Diabetes Obes Metab. 2018;20(3):582-9. Epub 2017/09/13. doi: 10.1111/dom.13115. PubMed PMID: 28898514; PubMed Central PMCID: PMCPMC5836890. Estimation Study

133. Hong N, Zhang N, Wu H, Lu S, Yu Y, Hou L, Lu Y, Liu H, Jiang G. **Preliminary exploration of survival analysis using the OHDSI common data model: a case study of intrahepatic cholangiocarcinoma.** BMC Med Inform Decis Mak. 2018;18(Suppl 5):116. Epub 2018/12/12. doi: 10.1186/s12911-018-0686-7 10.1186/s12911-018-0686-7 [pii]. PubMed PMID: 30526572; PubMed Central PMCID: PMC6284277. Estimation Study

134. Kubota K, Kamijima Y, Kao Yang YH, Kimura S, Chia-Cheng Lai E, Man KKC, Ryan P, Schuemie M, Stang P, Su CC, Wong ICK, Zhang Y, Setoguchi S. **Penetration of new antidiabetic medications in East Asian countries and the United States: A cross-national comparative study.** PLoS One. 2018;13(12):e0208796. Epub 2018/12/13. doi: 10.1371/journal.pone.0208796. PubMed PMID: 30540837; PubMed Central PMCID: PMCPMC6291148. Characterization Study

135. Prokosch HU, Acker T, Bernarding J, Binder H, Boeker M, Boerries M, Daumke P, Ganslandt T, Hesser J, Honing G, Neumaier M, Marquardt K, Renz H, Rothkotter HJ, Schade-Brittinger C, Schmucker P, Schuttler J, Sedlmayr M, Serve H, Sohrabi K, Storf H. **MIRACUM: Medical Informatics in Research and Care in University Medicine.** Methods Inf Med. 2018;57(S 01):e82-e91. Epub 2018/07/18. doi: 10.3414/ME17-02-0025. PubMed PMID: 30016814; PubMed Central PMCID: PMC6178200.Referring to

136. Huser V, Kahn MG, Brown JS, Gouripeddi R. **Methods for examining data quality in healthcare integrated data repositories.** Pac Symp Biocomput. 2018;23:628-33. Epub 2017/12/09. doi: 9789813235533_0059 [pii]. PubMed PMID: 29218922. Data Quality

137. Levine ME, Albers DJ, Hripcsak G. **Methodological variations in lagged regression for detecting physiologic drug effects in EHR data.** J Biomed Inform. 2018;86:149-59. Epub 2018/09/03. doi: S1532-0464(18)30173-4 [pii] 10.1016/j.jbi.2018.08.014. PubMed PMID: 30172760; PubMed Central PMCID: PMC6207533. Estimation Methods

138. Mower J, Subramanian D, Cohen T. **Learning predictive models of drug side-effect relationships from distributed representations of literature-derived semantic predications.** J Am Med Inform Assoc. 2018;25(10):1339-50. Epub 2018/07/17. doi: 5052182 [pii] 10.1093/jamia/ocy077. PubMed PMID: 30010902; PubMed Central PMCID: PMC6454491. Prediction Methods

139. Matcho A, Ryan P, Fife D, Gifkins D, Knoll C, Friedman A. **Inferring pregnancy episodes and outcomes within a network of observational databases.** PLoS One. 2018;13(2):e0192033. Epub 2018/02/02. doi: 10.1371/journal.pone.0192033. PubMed PMID: 29389968; PubMed Central PMCID: PMCPMC5794136. Phenotyping

140. Schuemie MJ, Ryan PB, Hripcsak G, Madigan D, Suchard MA. **Improving reproducibility by using high-throughput observational studies with empirical calibration.** Philos Trans A Math Phys Eng Sci. 2018;376(2128). Epub 2018/08/08. doi: rsta.2017.0356 [pii] 10.1098/rsta.2017.0356. PubMed PMID: 30082302; PubMed Central PMCID: PMC6107542. Foundation

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

OHDSI Publications

Legend

Orange ..... Data Standards
Red ..... Methods Research
Green ..... Clinical Applications
Black ..... Other

141. Seneviratne MG, Banda JM, Brooks JD, Shah NH, Hernandez-Boussard TM. **Identifying Cases of Metastatic Prostate Cancer Using Machine Learning on Electronic Health Records.** AMIA Annu Symp Proc. 2018;2018:1498-504. Epub 2019/03/01. PubMed PMID: 30815195; PubMed Central PMCID: PMC6371284.Prediction Study

142. Hripcsak G, Albers DJ. **High-fidelity phenotyping: richness and freedom from bias.** Journal of the American Medical Informatics Association. 2018;25(3):289-94.Phenotyping

143. Cepeda MS, Reps J, Ryan P. **Finding factors that predict treatment-resistant depression: Results of a cohort study.** Depress Anxiety. 2018;35(7):668-73. Epub 2018/05/23. doi: 10.1002/da.22774. PubMed PMID: 29786922; PubMed Central PMCID: PMCPMC6055726. Prediction Study

144. Cho S, Mohan S, Husain SA, Natarajan K. **Expanding transplant outcomes research opportunities through the use of a common data model. American Journal of Transplantation.** 2018;18(6):1321-7. ETL

145. Yang Y, Zhou X, Gao S, Lin H, Xie Y, Feng Y, Huang K, Zhan S. **Evaluation of Electronic Healthcare Databases for Post-Marketing Drug Safety Surveillance and Pharmacoepidemiology in China.** Drug Saf. 2018;41(1):125-37. Epub 2017/08/18. doi: 10.1007/s40264-017-0589-z 10.1007/s40264-017-0589-z [pii]. PubMed PMID: 28815480. Vocabulary

146. Tian Y, Schuemie MJ, Suchard MA. **Evaluating large-scale propensity score performance through real-world and synthetic data experiments.** International journal of epidemiology. 2018;47(6):2005-14. Estimation Methods

147. Schuemie MJ, Hripcsak G, Ryan PB, Madigan D, Suchard MA. **Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data.** Proc Natl Acad Sci U S A. 2018;115(11):2571-7. Epub 2018/03/14. doi: 1708282114 [pii] 10.1073/pnas.1708282114. PubMed PMID: 29531023; PubMed Central PMCID: PMC5856503. Estimation Methods

148. Hripcsak G, Levine ME, Shang N, Ryan PB. **Effect of vocabulary mapping for conditions on phenotype cohorts.** J Am Med Inform Assoc. 2018;25(12):1618-25. Epub 2018/11/06. doi: 5159502 [pii] 10.1093/jamia/ocy124. PubMed PMID: 30395248; PubMed Central PMCID: PMC6289550. Phenotyping

149. Polubriaginof FCG, Vanguri R, Quinnies K, Belbin GM, Yahi A, Salmasian H, Lorberbaum T, Nwankwo V, Li L, Shervey MM, Glowe P, Ionita-Laza I, Simmerling M, Hripcsak G, Bakken S, Goldstein D, Kiryluk K, Kenny EE, Dudley J, Vawdrey DK, Tatonetti NP. **Disease Heritability Inferred from Familial Relationships Reported in Medical Records.** Cell. 2018;173(7):1692-704 e11. Epub 2018/05/22. doi: S0092-8674(18)30525-7 [pii] 10.1016/j.cell.2018.04.032. PubMed PMID: 29779949; PubMed Central PMCID: PMC6015747. Characterization Study, Phenotyping

150. Reps J, Hsiao C, Johnston S. **Development and Validation of a Model to Predict Cessation of Antihyperglycemic Medication after Laparoscopic Bariatric Surgery Among Patients with Type 2 Diabetes.** Value in Health. 2018;21:S249-S50. Prediction Study

151. Reps JM, Schuemie MJ, Suchard MA, Ryan PB, Rijnbeek PR. **Design and implementation of a standardized framework to generate and evaluate patient-level prediction models using observational healthcare data.** J Am Med Inform Assoc. 2018;25(8):969-75. Epub 2018/05/03. doi: 10.1093/jamia/ocy032. PubMed PMID: 29718407; PubMed Central PMCID: PMCPMC6077830. Foundation; Prediction Methods; Prediction Study

152. Sun YX, Pei ZC, Zhan SY. **[Data harmonization and sharing in study cohorts of respiratory diseases].** Zhonghua Liu Xing Bing Xue Za Zhi. 2018;39(2):233-9. Epub 2018/03/02. doi: 10.3760 cma.j.issn.0254-6450.2018.02.019. PubMed PMID: 29495212. Vocabulary

153. Butler A, Wei W, Yuan C, Kang T, Si Y, Weng C. **The Data Gap in the EHR for Clinical Research Eligibility Screening.** AMIA Jt Summits Transl Sci Proc. 2018;2017:320-9. Epub 2018/06/12. PubMed PMID: 29888090; PubMed Central PMCID: PMC5961795. NLP, Vocabulary

154. Nestsiarovich A, Mazurie AJ, Hurwitz NG, Kerner B, Nelson SJ, Crisanti AS, Tohen M, Krall RL, Perkins DJ, Lambert CG. **Comprehensive comparison of monotherapies for psychiatric hospitalization risk in bipolar disorders.** Bipolar disorders. 2018;20(8):761-71.Estimation Study

155. Ryan PB, Buse JB, Schuemie MJ, DeFalco F, Yuan Z, Stang PE, Berlin JA, Rosenthal N. **Comparative effectiveness of canagliflozin, SGLT2 inhibitors and non-SGLT2 inhibitors on the risk of hospitalization for heart failure and amputation in patients with type 2 diabetes mellitus: A real-world meta-analysis of 4 observational databases (OBSERVE-4D).** Diabetes Obes Metab. 2018;20(11):2585-97. Epub 2018/06/26. doi: 10.1111/dom.13424. PubMed PMID: 29938883; PubMed Central PMCID: PMCPMC6220807.Estimation Study

156. Liyanage H, Liaw ST, Jonnagaddala J, Hinton W, de Lusignan S. **Common Data Models (CDMs) to Enhance International Big Data Analytics: A Diabetes Use Case to Compare Three CDMs.** Stud Health Technol Inform. 2018;255:60-4. Epub 2018/10/12. PubMed PMID: 30306907. Referring to

157. Ta CN, Dumontier M, Hripcsak G, Tatonetti NP, Weng C. **Columbia Open Health Data, clinical concept prevalence and co-occurrence from electronic health records.** Sci Data. 2018;5:180273. Epub 2018/11/28. doi: sdata2018273 [pii] 10.1038/sdata.2018.273. PubMed PMID: 30480666; PubMed Central PMCID: PMC6257042. ETL

158. Rinner C, Gezgin D, Wendl C, Gall W. **A Clinical Data Warehouse Based on OMOP and i2b2 for Austrian Health Claims Data.** Stud Health Technol Inform. 2018;248:94-9. Epub 2018/05/05. PubMed PMID: 29726424. ETL

159. Elkin PL, Mullin S, Sakilay S. **Biomedical Informatics Investigator.** Stud Health Technol Inform. 2018;255:195-9. Epub 2018/10/12. PubMed PMID: 30306935; PubMed Central PMCID: PMC7847179. Vocabulary

160. Vashisht R, Jung K, Schuler A, Banda JM, Park RW, Jin S, Li L, Dudley JT, Johnson KW, Shervey MM, Xu H, Wu Y, Natrajan K, Hripcsak G, Jin P, Van Zandt M, Reckard A, Reich CG, Weaver J, Schuemie MJ, Ryan PB, Callahan A, Shah NH. **Association of Hemoglobin A1c Levels With Use of Sulfonylureas, Dipeptidyl Peptidase 4 Inhibitors, and Thiazolidinediones in Patients With Type 2 Diabetes Treated With Metformin: Analysis From the Observational Health Data Sciences and Informatics Initiative.** JAMA Netw Open. 2018;1(4):e181755. Epub 2019/01/16. doi: 2698083 [pii] 10.1001/jamanetworkopen.2018.1755. PubMed PMID: 30646124; PubMed Central PMCID: PMC6324274.Estimation Study

161. Lai EC, Ryan P, Zhang Y, Schuemie M, Hardy NC, Kamijima Y, Kimura S, Kubota K, Man KK, Cho SY, Park RW, Stang P, Su CC, Wong IC, Kao YY, Setoguchi S. **Applying a common data model to Asian databases for multinational pharmacoepidemiologic studies: opportunities and challenges.** Clin Epidemiol. 2018;10:875-85. Epub 2018/08/14. doi: 10.2147/clep.S149961. PubMed PMID: 30100761; PubMed Central PMCID: PMCPMC6067778. ETL; Data Quality

162. Zhang X, Wang L, Miao S, Xu H, Yin Y, Zhu Y, Dai Z, Shan T, Jing S, Wang J, Zhang X, Huang Z, Wang Z, Guo J, Liu Y. **Analysis of treatment pathways for three chronic diseases using OMOP CDM.** J Med Syst. 2018;42(12):260. Epub 2018/11/14. doi: 10.1007/s10916-018-1076-5 10.1007/s10916-018-1076-5 [pii]. PubMed PMID: 30421323; PubMed Central PMCID: PMC6244882. Characterization Study

163. Dixon BE, Wen C, French T, Williams J, Grannis SJ. **Advanced Visualization and Analysis of Data Quality for Syndromic Surveillance Systems.** Online Journal of Public Health Informatics. 2018;10(1). Data Quality

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

OHDSI Publications

164. Johnston SS, Morton JM, Kalsekar I, Ammann EM, Hsiao C-W, Reps J. **Using machine learning applied to real-world healthcare data for predictive analytics: an applied example in bariatric surgery.** Value in Health. 2019;22(5):580-6. Prediction Study

165. Lima DM, Rodrigues-Jr JF, Traina AJM, Pires FA, Gutierrez MA. **Transforming Two Decades of ePR Data to OMOP CDM for Clinical Research.** Stud Health Technol Inform. 2019;264:233-7. Epub 2019/08/24. doi: SHTI190218 [pii] 10.3233/SHTI190218. PubMed PMID: 31437920. ETL

166. Reps JM, Rijnbeek PR, Ryan PB. **Supplementing claims data analysis using self-reported data to develop a probabilistic phenotype model for current smoking status.** J Biomed Inform. 2019;97:103264. Epub 2019/08/07. doi: S1532-0464(19)30183-2 [pii] 10.1016/j.jbi.2019.103264. PubMed PMID: 31386904. Phenotyping, Prediction Study

167. Haberson A, Rinner C, Gall W. **Standardizing Austrians Claims Data Using the OMOP Common Data Model: A Feasibility Study.** Stud Health Technol Inform. 2019;258:151-2. Epub 2019/04/04. PubMed PMID: 30942734. ETL

168. Belenkaya R, Gurley M, Dymshyts D, Araujo S, Williams A, Chen R, Reich C. **Standardized Observational Cancer Research Using the OMOP CDM Oncology Module.** Stud Health Technol Inform. 2019;264:1831-2. Epub 2019/08/24. doi: SHTI190670 [pii] 10.3233/SHTI190670. PubMed PMID: 31438365. ETL, Vocabulary

169. Glicksberg BS, Oskotsky B, Giangreco N, Thangaraj PM, Rudrapatna V, Datta D, Frazier R, Lee N, Larsen R, Tatonetti NP, Butte AJ. **ROMOP: a light-weight R package for interfacing with OMOP-formatted electronic health record data.** JAMIA open. 2019;2(1):10-4. Epub 2019/10/22. doi: 10.1093/jamiaopen/ooy059. PubMed PMID: 31633087; PubMed Central PMCID: PMC6800657.Methods Other

170. Ross EG, Jung K, Dudley JT, Li L, Leeper NJ, Shah NH. **Predicting future cardiovascular events in patients with peripheral artery disease using electronic health record data.** Circulation: Cardiovascular Quality and Outcomes. 2019;12(3):e004741.Prediction Study

171. Schuemie MJ, Ryan PB, Man KKC, Wong ICK, Suchard MA, Hripcsak G. **A plea to stop using the case-control design in retrospective database studies.** Stat Med. 2019;38(22):4199-208. Epub 2019/08/23. doi: 10.1002/sim.8215. PubMed PMID: 31436848; PubMed Central PMCID: PMCPMC6771795. Estimation Methods, Position

172. Swerdel JN, Hripcsak G, Ryan PB. **PheValuator: Development and evaluation of a phenotype algorithm evaluator.** J Biomed Inform. 2019;97:103258. Epub 2019/08/02. doi: S1532-0464(19)30177-7 [pii] 10.1016/j.jbi.2019.103258. PubMed PMID: 31369862; PubMed Central PMCID: PMC7736922. Phenotyping

173. Glicksberg BS, Oskotsky B, Thangaraj PM, Giangreco N, Badgeley MA, Johnson KW, Datta D, Rudrapatna VA, Rappoport N, Shervey MM, Miotto R, Goldstein TC, Rutenberg E, Frazier R, Lee N, Israni S, Larsen R, Percha B, Li L, Dudley JT, Tatonetti NP, Butte AJ. **PatientExploreR: an extensible application for dynamic visualization of patient clinical history from electronic health records in the OMOP common data model.** Bioinformatics. 2019;35(21):4515-8. Epub 2019/06/20. doi: 5520433 [pii] 10.1093/bioinformatics/btz409. PubMed PMID: 31214700; PubMed Central PMCID: PMC6821222. Phenotyping

174. Kwong M, Gardner HL, Dieterle N, Rentko V. **Optimization of Electronic Medical Records for Data Mining Using a Common Data Model.** Top Companion Anim Med. 2019;37:100364. Epub 2019/12/16. doi: S1938-9736(19)30101-1 [pii] 10.1016/j.tcam.2019.100364. PubMed PMID: 31837755; PubMed Central PMCID: PMC7874511. ETL

175. Burn E, Weaver J, Morales D, Prats-Uribe A, Delmestri A, Strauss VY, He Y, Robinson DE, Pinedo-Villanueva R, Kolovos S, Duarte-Salles T, Sproviero W, Yu D, Van Speybroeck M, Williams R, John LH, Hughes N, Sena AG, Costello R, Birlie B, Culliford D, O'Leary C, Morgan H, Burkard T, Prieto-Alhambra D, Ryan P. **Opioid use, postoperative complications, and implant survival after unicompartmental versus total knee replacement: a population-based network study.** The Lancet Rheumatology. 2019;1(4):e229-e36. doi: https://doi.org/10.1016 S2665-9913(19)30075-X.Estimation Study

176. Kapsner LA, Kampf MO, Seuchter SA, Kamdje-Wabo G, Gradinger T, Ganslandt T, Mate S, Gruendner J, Kraska D, Prokosch H-U. **Moving towards an EHR data quality framework: the MIRACUM approach.** German Medical Data Sciences: Shaping Change–Creative Solutions for Innovative Medicine: IOS Press; 2019. p. 247-53. Data Quality

177. Shang N, Liu C, Rasmussen LV, Ta CN, Caroll RJ, Benoit B, Lingren T, Dikilitas O, Mentch FD, Carrell DS, Wei WQ, Luo Y, Gainer VS, Kullo IJ, Pacheco JA, Hakonarson H, Walunas TL, Denny JC, Wiley K, Murphy SN, Hripcsak G, Weng C. **Making work visible for electronic phenotype implementation: Lessons learned from the eMERGE network.** J Biomed Inform. 2019;99:103293. Epub 2019/09/23. doi: S1532-0464(19)30212-6 [pii]10.1016/j.jbi.2019.103293. PubMed PMID: 31542521; PubMed Central PMCID: PMC6894517. Phenotyping

178. Weeks J, Pardee R. **Learning to Share Health Care Data: A Brief Timeline of Influential Common Data Models and Distributed Health Data Networks in U.S. Health Care Research.** EGEMS (Wash DC). 2019;7(1):4. Epub 2019/04/03. doi: 10.5334/egems.279. PubMed PMID: 30937326; PubMed Central PMCID: PMC6437693. Referring to

179. Gruendner J, Schwachhofer T, Sippl P, Wolf N, Erpenbeck M, Gulden C, Kapsner LA, Zierk J, Mate S, Sturzl M, Croner R, Prokosch HU, Toddenroth D. **KETOS: Clinical decision support and machine learning as a service - A training and deployment platform based on Docker, OMOP-CDM, and FHIR Web Services.** PLoS One. 2019;14(10):e0223010. Epub 2019/10/04. doi: 10.1371/journal.pone.0223010 PONE-D-19-12555 [pii]. PubMed PMID: 31581246; PubMed Central PMCID: PMC6776354.Prediction Methods, FHIR

180. Chandran U, Reps J, Stang PE, Ryan PB. **Inferring disease severity in rheumatoid arthritis using predictive modeling in administrative claims databases.** PLoS One. 2019;14(12):e0226255. Epub 2019/12/19. doi: 10.1371/journal.pone.0226255. PubMed PMID: 31851711; PubMed Central PMCID: PMCPMC6919633. Prediction Study

181. Lynch KE, Deppen SA, DuVall SL, Viernes B, Cao A, Park D, Hanchrow E, Hewa K, Greaves P, Matheny ME. **Incrementally Transforming Electronic Medical Records into the Observational Medical Outcomes Partnership Common Data Model: A Multidimensional Quality Assurance Approach.** Appl Clin Inform. 2019;10(5):794-803. Epub 2019/10/24. doi: 10.1055/s-0039-1697598. PubMed PMID: 31645076; PubMed Central PMCID: PMC6811349. Data Quality

182. Reps JM, Rijnbeek PR, Ryan PB. **Identifying the DEAD: Development and Validation of a Patient-Level Model to Predict Death Status in Population-Level Claims Data.** Drug Saf. 2019;42(11):1377-86. Epub 2019/05/06. doi: 10.1007/s40264-019-00827-0 10.1007/s40264-019-00827-0 [pii]. PubMed PMID: 31054141; PubMed Central PMCID: PMC6834730. Prediction Study

183. Warner JL, Dymshyts D, Reich CG, Gurley MJ, Hochheiser H, Moldwin ZH, Belenkaya R, Williams AE, Yang PC. **HemOnc: A new standard vocabulary for chemotherapy regimen representation in the OMOP common data model.** J Biomed Inform. 2019;96:103239. Epub 2019/06/27. doi: S1532-0464(19)30158-3 [pii] 10.1016/j.jbi.2019.103239. PubMed PMID: 31238109; PubMed Central PMCID: PMC6697579. Vocabulary

184. Shin SJ, You SC, Park YR, Roh J, Kim JH, Haam S, Reich CG, Blacketer C, Son DS, Oh S, Park RW. **Genomic Common Data Model for Seamless Interoperation of Biomedical Data in Clinical Practice: Retrospective Study.** J Med Internet Res. 2019;21(3):e13249. Epub 2019/03/27. doi: v21i3e13249 [pii] 10.2196/13249. PubMed PMID: 30912749; PubMed Central PMCID: PMC6454347. Vocabulary

185. Shin SJ, You SC, Roh J, Park YR, Park RW. **Genomic Common Data Model for Biomedical Data in Clinical Practice.** Stud Health Technol Inform. 2019;264:1843-4. Epub 2019/08/24. doi: SHTI190676 [pii] 10.3233/SHTI190676. PubMed PMID: 31438371. Vocabulary

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

OHDSI Publications

Legend

Orange.....Data Standards
Red.....Methods Research
Green.....Clinical Applications
Black.....Other

186. Banda JM. **Fully connecting the Observational Health Data Science and Informatics (OHDSI) initiative with the world of linked open data.** Genomics Inform. 2019;17(2):e13. Epub 2019/07/16. doi: GI.2019.17.2.e13 [pii] 10.5808/GI.2019.17.2.e13. PubMed PMID: 31307128; PubMed Central PMCID: PMC6808628. Vocabulary

187. Banda JM, Sarraju A, Abbasi F, Parizo J, Pariani M, Ison H, Briskin E, Wand H, Dubois S, Jung K. **Finding missed cases of familial hypercholesterolemia in health systems using machine learning.** NPJ digital medicine. 2019;2(1):1-8. Phenotyping

188. Bartlett VL, Dhruva SS, Shah ND, Ryan P, Ross JS. **Feasibility of Using Real-World Data to Replicate Clinical Trial Evidence.** JAMA Netw Open. 2019;2(10):e1912869. Epub 2019/10/10. doi: 2752575 [pii] 10.1001/jamanetworkopen.2019.12869. PubMed PMID: 31596493; PubMed Central PMCID: PMC6802419. Estimation Methods

189. Haberson A, Rinner C, Schoberl A, Gall W. **Feasibility of Mapping Austrian Health Claims Data to the OMOP Common Data Model.** J Med Syst. 2019;43(10):314. Epub 2019/09/09. doi: 10.1007/s10916-019-1436-9 10.1007/s10916-019-1436-9 [pii]. PubMed PMID: 31494719; PubMed Central PMCID: PMC6732152. ETL

190. Hripcsak G, Shang N, Peissig PL, Rasmussen LV, Liu C, Benoit B, Carroll RJ, Carrell DS, Denny JC, Dikilitas O, Gainer VS, Howell KM, Klann JG, Kullo IJ, Lingren T, Mentch FD, Murphy SN, Natarajan K, Pacheco JA, Wei WQ, Wiley K, Weng C. **Facilitating phenotype transfer using a common data model.** J Biomed Inform. 2019;96:103253. Epub 2019/07/22. doi: S1532-0464(19)30172-8 [pii] 10.1016/j.jbi.2019.103253. PubMed PMID: 31325501; PubMed Central PMCID: PMC6697565. Phenotyping

191. Wang L, Voss EA, Weaver J, Hester L, Yuan Z, DeFalco F, Schuemie MJ, Ryan PB, Sun D, Freedman A, Alba M, Lind J, Meininger G, Berlin JA, Rosenthal N. **Diabetic ketoacidosis in patients with type 2 diabetes treated with sodium glucose co-transporter 2 inhibitors versus other antihyperglycemic agents: An observational study of four US administrative claims databases.** Pharmacoepidemiol Drug Saf. 2019;28(12):1620-8. Epub 2019/08/29. doi: 10.1002/pds.4887. PubMed PMID: 31456304; PubMed Central PMCID: PMCPMC6916409. Estimation Study

192. Sharma H, Mao C, Zhang Y, Vatani H, Yao L, Zhong Y, Rasmussen L, Jiang G, Pathak J, Luo Y. **Developing a portable natural language processing based phenotyping system.** BMC Med Inform Decis Mak. 2019;19(Suppl 3):78. Epub 2019/04/05. doi: 10.1186/s12911-019-0786-z 10.1186/s12911-019-0786-z [pii]. PubMed PMID: 30943974; PubMed Central PMCID: PMC6448187. NLP, Phenotyping

193. Ta CN, Weng C. **Detecting systemic data quality issues in electronic health records.** Studies in health technology and informatics. 2019;264:383. Data Quality

194. Klann JG, Joss MAH, Embree K, Murphy SN. **Data model harmonization for the All Of Us Research Program: Transforming i2b2 data into the OMOP common data model.** PLoS One. 2019;14(2):e0212463. Epub 2019/02/20. doi: 10.1371/journal.pone.0212463 PONE-D-18-34751 [pii]. PubMed PMID: 30779778; PubMed Central PMCID: PMC6380544. ETL

195. Rogers JR, Callahan TJ, Kang T, Bauck A, Khare R, Brown JS, Kahn MG, Weng C. **A Data Element-Function Conceptual Model for Data Quality Checks.** EGEMS (Wash DC). 2019;7(1):17. Epub 2019/05/09. doi: 10.5334/egems.289. PubMed PMID: 31065558; PubMed Central PMCID: PMC6484368. Data Quality

196. Yuan C, Ryan PB, Ta C, Guo Y, Li Z, Hardin J, Makadia R, Jin P, Shang N, Kang T, Weng C. **Criteria2Query: a natural language interface to clinical databases for cohort definition.** J Am Med Inform Assoc. 2019;26(4):294-305. Epub 2019/02/13. doi: 5308980 [pii] 10.1093/jamia/ocy178. PubMed PMID: 30753493; PubMed Central PMCID: PMC6402359.Methods Other

197. Rasmussen LV, Brandt PS, Jiang G, Kiefer RC, Pacheco JA, Adekkanattu P, Ancker JS, Wang F, Xu Z, Pathak J, Luo Y. **Considerations for Improving the Portability of Electronic Health Record-Based Phenotype Algorithms.** AMIA Annu Symp Proc. 2019;2019:755-64. Epub 2020/04/21. PubMed PMID: 32308871; PubMed Central PMCID: PMC7153055. Phenotyping

198. Suchard MA, Schuemie MJ, Krumholz HM, You SC, Chen R, Pratt N, Reich CG, Duke J, Madigan D, Hripcsak G, Ryan PB. **Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis.** Lancet. 2019;394(10211):1816-26. Epub 2019/11/02. doi: 10.1016/s0140-6736(19)32317-7. PubMed PMID: 31668726; PubMed Central PMCID: PMCPMC6924620. Estimation Study

199. Guo GN, Jonnagaddala J, Farshid S, Huser V, Reich C, Liaw ST. **Comparison of the cohort selection performance of Australian Medicines Terminology to Anatomical Therapeutic Chemical mappings.** J Am Med Inform Assoc. 2019;26(11):1237-46. Epub 2019/09/24. doi: 5572827 [pii] 10.1093/jamia/ocz143. PubMed PMID: 31545380; PubMed Central PMCID: PMC7647230. ETL

200. Schuemie MJ, Madigan D, Ryan PB, Reich C, Suchard MA, Berlin JA, Hripcsak G. **Comment on "How pharmacoepidemiology networks can manage distributed analyses to improve replicability and transparency and minimize bias".** Pharmacoepidemiol Drug Saf. 2019;28(7):1032-3. Epub 2019/05/09. doi: 10.1002/pds.4798. PubMed PMID: 31066478. Position

201. Viernes B, Lynch KE, South B, Coronado G, DuVall SL. **Characterizing VA Users with the OMOP Common Data Model.** Stud Health Technol Inform. 2019;264:1614-5. Epub 2019/08/24. doi: SHTI190561 [pii] 10.3233/SHTI190561. PubMed PMID: 31438258. Characterization Study

202. Polubriaginof FCG, Ryan P, Salmasian H, Shapiro AW, Perotte A, Safford MM, Hripcsak G, Smith S, Tatonetti NP, Vawdrey DK. **Challenges with quality of race and ethnicity data in observational databases.** J Am Med Inform Assoc. 2019;26(8-9):730-6. Epub 2019/08/01. doi: 10.1093/jamia/ocz113. PubMed PMID: 31365089; PubMed Central PMCID: PMCPMC6696496. Characterization Study

203. Meystre SM, Heider PM, Kim Y, Aruch DB, Britten CD. **Automatic trial eligibility surveillance based on unstructured clinical data.** Int J Med Inform. 2019;129:13-9. Epub 2019/08/25. doi: S1386-5056(18)31052-9 [pii] 10.1016/j.ijmedinf.2019.05.018. PubMed PMID: 31445247; PubMed Central PMCID: PMC6717538.Trial

204. Jiang G, Yu Y, Kingsbury PR, Shah N. **Augmenting Medical Device Evaluation Using a Reusable Unique Device Identifier Interoperability Solution Based on the OHDSI Common Data Model.** Stud Health Technol Inform. 2019;264:1502-3. Epub 2019/08/24. doi: SHTI190505 [pii] 10.3233/SHTI190505. PubMed PMID: 31438202. Vocabulary

205. Yu Y, Ruddy KJ, Hong N, Tsuji S, Wen A, Shah ND, Jiang G. **ADEpedia-on-OHDSI: A next generation pharmacovigilance signal detection platform using the OHDSI common data model.** J Biomed Inform. 2019;91:103119. Epub 2019/02/11. doi: S1532-0464(19)30037-1 [pii] 10.1016/j.jbi.2019.103119. PubMed PMID: 30738946; PubMed Central PMCID: PMC6432939. ETL

206. Reps JM, Cepeda MS, Ryan PB. Wisdom of the CROUD: **Development and validation of a patient-level prediction model for opioid use disorder using population-level claims data.** PLoS One. 2020;15(2):e0228632. Epub 2020/02/14. doi: 10.1371/journal.pone.0228632. PubMed PMID: 32053653; PubMed Central PMCID: PMCPMC7017997. Prediction Study

207. Chen R, Ryan P, Natarajan K, Falconer T, Crew KD, Reich CG, Vashisht R, Randhawa G, Shah NH, Hripcsak G. **Treatment Patterns for Chronic Comorbid Conditions in Patients With Cancer Using a Large-Scale Observational Data Network.** JCO Clin Cancer Inform. 2020;4:171-83. Epub 2020/03/07. doi: 10.1200/cci.19.00107. PubMed PMID: 32134687; PubMed Central PMCID: PMCPMC7113074. Characterization Study

208. Kern DM, Cepeda MS, Defalco F, Etropolski M. **Treatment patterns and sequences of pharmacotherapy for patients diagnosed with depression in the United States: 2014 through 2019.** BMC psychiatry. 2020;20(1):1-10. Characterization Study

209. Kern DM, Cepeda MS. **Treatment patterns and comorbid burden of patients newly diagnosed with multiple sclerosis in the United States.** BMC neurology. 2020;20(1):1-8. Characterization Study

210. Averitt AJ, Weng C, Ryan P, Perotte A. **Translating evidence into practice: eligibility criteria fail to eliminate clinically significant differences between real-world and study populations.** NPJ Digit Med. 2020;3:67. Epub 2020/05/16. doi: 10.1038/s41746-020-0277-8. PubMed PMID: 32411828; PubMed Central PMCID: PMCPMC7214444. Estimation Methods, Trial

211. Lamer A, Depas N, Doutreligne M, Parrot A, Verloop D, Defebvre MM, Ficheur G, Chazard E, Beuscart JB. **Transforming French Electronic Health Records into the Observational Medical Outcome Partnership's Common Data Model: A Feasibility Study.** Appl Clin Inform. 2020;11(1):13-22. Epub 2020/01/09. doi: 10.1055/s-0039-3402754. PubMed PMID: 31914471; PubMed Central PMCID: PMC6949163. ETL

212. Ryu B, Yoon E, Kim S, Lee S, Baek H, Yi S, Na HY, Kim JW, Baek RM, Hwang H, Yoo S. **Transformation of Pathology Reports Into the Common Data Model With Oncology Module: Use Case for Colon Cancer.** J Med Internet Res. 2020;22(12):e18526. Epub 2020/12/10. doi: v22i12e18526 [pii] 10.2196/18526. PubMed PMID: 33295294; PubMed Central PMCID: PMC7758167. ETL, NLP

213. Brandt PS, Kiefer RC, Pacheco JA, Adekkanattu P, Sholle ET, Ahmad FS, Xu J, Xu Z, Ancker JS, Wang F, Luo Y, Jiang G, Pathak J, Rasmussen LV. **Toward cross-platform electronic health record-driven phenotyping using Clinical Quality Language.** Learn Health Syst. 2020;4(4):e10233. Epub 2020/10/22. doi: 10.1002/lrh2.10233 LRH210233 [pii]. PubMed PMID: 33083538; PubMed Central PMCID: PMC7556419. Phenotyping

214. Syed S, Baghal A, Prior F, Zozus M, Al-Shukri S, Syeda HB, Garza M, Begum S, Gates K, Syed M, Sexton KW. **Toolkit to Compute Time-Based Elixhauser Comorbidity Indices and Extension to Common Data Models.** Healthc Inform Res. 2020;26(3):193-200. Epub 2020/08/21. doi: hir.2020.26.3.193 [pii] 10.4258/hir.2020.26.3.193. PubMed PMID: 32819037; PubMed Central PMCID: PMC7438698. ETL, Phenotyping

215. Gruhl M, Reinecke I, Sedlmayr M. **Specification and Distribution of Vocabularies Among Consortial Partners.** Stud Health Technol Inform. 2020;270:1393-4. Epub 2020/06/24. doi: SHTI200458 [pii] 10.3233/SHTI200458. PubMed PMID: 32570675.Vocabulary

216. Stephens KA, West II, Hallgren KA, Mollis B, Ma K, Donovan DM, Stuvek B, Baldwin L-M. **Service utilization and chronic condition outcomes among primary care patients with substance use disorders and co-occurring chronic conditions.** Journal of substance abuse treatment. 2020;112:49-55. Characterization Study

217. Tong J, Duan R, Li R, Scheuemie MJ, Moore JH, Chen Y. **Robust-ODAL: Learning from heterogeneous health systems without sharing patient-level data.** Pac Symp Biocomput. 2020;25:695-706. Epub 2019/12/05. doi: 9789811215636_0061 [pii]. PubMed PMID: 31797639; PubMed Central PMCID: PMC6905508. Prediction Methods

218. Lee DY, Cho J, You SC, Park RW, Kim CS, Lee EY, Aizenstein H, Andreescu C, Karim H, Hong CH. **Risk of mortality in elderly coronavirus disease 2019 patients with mental health disorders: a nationwide retrospective study in South Korea.** The American Journal of Geriatric Psychiatry. 2020;28(12):1308-16. Characterization Study

219. Lane JCE, Weaver J, Kostka K, Duarte-Salles T, Abrahao MTF, Alghoul H, Alser O, Alshammari TM, Biedermann P, Banda JM, Burn E, Casajust P, Conover MM, Culhane AC, Davydov A, DuVall SL, Dymshyts D, Fernandez-Bertolin S, Fister K, Hardin J, Hester L, Hripcsak G, Kaas-Hansen BS, Kent S, Khosla S, Kolovos S, Lambert CG, van der Lei J, Lynch KE, Makadia R, Margulis AV, Matheny ME, Mehta P, Morales DR, Morgan-Stewart H, Mosseveld M, Newby D, Nyberg F, Ostropolets A, Park RW, Prats-Uribe A, Rao GA, Reich C, Reps J, Rijnbeek P, Sathappan SMK, Schuemie M, Seager S, Sena AG, Shoaibi A, Spotnitz M, Suchard MA, Torre CO, Vizcaya D, Wen H, de Wilde M, Xie J, You SC, Zhang L, Zhuk O, Ryan P, Prieto-Alhambra D, consortium O-C-. **Risk of hydroxychloroquine alone and in combination with azithromycin in the treatment of rheumatoid arthritis: a multinational, retrospective study.** Lancet Rheumatol. 2020;2(11):e698-e711. Epub 2020/08/31. doi: 10.1016/S2665-9913(20)30276-9 S2665-9913(20)30276-9 [pii]. PubMed PMID: 32864627; PubMed Central PMCID: PMC7442425. Estimation Study

220. Lane JCE, Weaver J, Kostka K, Duarte-Salles T, Abrahao MTF, Alghoul H, Alser O, Alshammari TM, Areia C, Biedermann P, Banda JM, Burn E, Casajust P, Fister K, Hardin J, Hester L, Hripcsak G, Kaas-Hansen BS, Khosla S, Kolovos S, Lynch KE, Makadia R, Mehta PP, Morales DR, Morgan-Stewart H, Mosseveld M, Newby D, Nyberg F, Ostropolets A, Woong Park R, Prats-Uribe A, Rao GA, Reich C, Rijnbeek P, Sena AG, Shoaibi A, Spotnitz M, Vignesh S, Suchard MA, Vizcaya D, Wen H, de Wilde M, Xie J, You SC, Zhang L, Lovestone S, Ryan P, Prieto-Alhambra D, consortium O-C-. **Risk of depression, suicide and psychosis with hydroxychloroquine treatment for rheumatoid arthritis: a multinational network cohort study.** Rheumatology (Oxford). 2020. Epub 2020/12/29. doi: 6048420 [pii] 10.1093/rheumatology/keaa771. PubMed PMID: 33367863; PubMed Central PMCID: PMC7798671. Estimation Study

221. Callahan A, Shah NH, Chen JH. **Research and Reporting Considerations for Observational Studies Using Electronic Health Record Data.** Ann Intern Med. 2020;172(11 Suppl):S79-s84. Epub 2020/06/02. doi: 10.7326/m19-0873. PubMed PMID: 32479175; PubMed Central PMCID: PMCPMC7413106. Phenotyping

222. Spotnitz ME, Natarajan K, Ryan PB, Westhoff CL. **Relative Risk of Cervical Neoplasms Among Copper and Levonorgestrel-Releasing Intrauterine System Users.** Obstet Gynecol. 2020;135(2):319-27. Epub 2020/01/11. doi: 10.1097/aog.0000000000003656. PubMed PMID: 31923062; PubMed Central PMCID: PMCPMC7012337. Estimation Study

223. Alnofal FA, Alrwisan AA, Alshammari TM. **Real-world data in Saudi Arabia: Current situation and challenges for regulatory decision-making.** Pharmacoepidemiology and Drug Safety. 2020;29(10):1303-6. ETL

224. Kim HI, Yoon JY, Kwak MS, Cha JM. **Real-World Use of Colonoscopy in an Older Population: A Nationwide Standard Cohort Study Using a Common Data Model.** Digestive Diseases and Sciences. 2020:1-8. Characterization Study

225. Alnofal FA, Alrwisan AA, Alshammari TM. **Real-world data in Saudi Arabia: Current situation and challenges for regulatory decision-making.** Pharmacoepidemiol Drug Saf. 2020;29(10):1303-6. Epub 2020/05/28. doi: 10.1002/pds.5025. PubMed PMID: 32458499. ETL

226. Jeon S, Seo J, Kim S, Lee J, Kim JH, Sohn JW, Moon J, Joo HJ. **Proposal and Assessment of a De-Identification Strategy to Enhance Anonymity of the Observational Medical Outcomes Partnership Common Data Model (OMOP-CDM) in a Public Cloud-Computing Environment: Anonymization of Medical Data Using Privacy Models.** J Med Internet Res. 2020;22(11):e19597. Epub 2020/11/13. doi: v22i11e19597 [pii] 10.2196/19597. PubMed PMID: 33177037; PubMed Central PMCID: PMC7728527. Methods Other

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

OHDSI Publications

227. Schuemie MJ, Ryan PB, Pratt N, Chen R, You SC, Krumholz HM, Madigan D, Hripcsak G, Suchard MA. **Principles of Large-scale Evidence Generation and Evaluation across a Network of Databases (LEGEND).** J Am Med Inform Assoc. 2020;27(8):1331-7. Epub 2020/09/11. doi: 10.1093/jamia/ocaa103. PubMed PMID: 32909033; PubMed Central PMCID: PMCPMC7481029. Estimation Methods

228. Jin S, Kostka K, Posada JD, Kim Y, Seo SI, Lee DY, Shah NH, Roh S, Lim Y-H, Chae SG. **Prediction of major depressive disorder following beta-blocker therapy in patients with cardiovascular diseases.** Journal of personalized medicine. 2020;10(4):288. Prediction Study

229. Zuo X, Li J, Zhao B, Zhou Y, Dong X, Duke J, Natarajan K, Hripcsak G, Shah N, Banda JM, Reeves R, Miller T, Xu H. **Normalizing Clinical Document Titles to LOINC Document Ontology: an Initial Study.** AMIA Annu Symp Proc. 2020;2020:1441-50. Epub 2021/05/04. doi: 181_3416722 [pii]. PubMed PMID: 33936520; PubMed Central PMCID: PMC8075502. Referring to, NLP

230. Chandler RE. **Nintedanib and ischemic colitis: Signal assessment with the integrated use of two types of real-world evidence, spontaneous reports of suspected adverse drug reactions, and observational data from large health-care databases.** Pharmacoepidemiol Drug Saf. 2020;29(8):951-7. Epub 2020/05/14. doi: 10.1002/pds.5022. PubMed PMID: 32399991; PubMed Central PMCID: PMC7496543. Estimation Study

231. Michael CL, Sholle ET, Wulff RT, Roboz GJ, Campion TR, Jr. **Mapping Local Biospecimen Records to the OMOP Common Data Model.** AMIA Jt Summits Transl Sci Proc. 2020;2020:422-9. Epub 2020/06/02. PubMed PMID: 32477663; PubMed Central PMCID: PMC7233045. ETL, Vocabulary

232. Fishbein HA, Birch RJ, Mathew SM, Sawyer HL, Pulver G, Poling J, Kaelber D, Mardon R, Johnson MC, Pace W, editors. **The Longitudinal Epidemiologic Assessment of Diabetes Risk (LEADR): Unique 1.4 M patient Electronic Health Record cohort.** Healthcare; 2020: Elsevier. ETL

233. Duan R, Luo C, Schuemie MJ, Tong J, Liang CJ, Chang HH, Boland MR, Bian J, Xu H, Holmes JH. **Learning from local to global: An efficient distributed algorithm for modeling time-to-event data.** Journal of the American Medical Informatics Association. 2020;27(7):1028-36. Estimation Methods

234. Schuemie MJ, Ryan PB, Pratt N, Chen R, You SC, Krumholz HM, Madigan D, Hripcsak G, Suchard MA. **Large-scale evidence generation and evaluation across a network of databases (LEGEND): assessing validity using hypertension as a case study.** J Am Med Inform Assoc. 2020;27(8):1268-77. Epub 2020/08/23. doi: 10.1093/jamia/ocaa124. PubMed PMID: 32827027; PubMed Central PMCID: PMCPMC7481033. Estimation Study

235. Brat GA, Weber GM, Gehlenborg N, Avillach P, Palmer NP, Chiovato L, Cimino J, Waitman LR, Omenn GS, Malovini A, Moore JH, Beaulieu-Jones BK, Tibollo V, Murphy SN, Yi SL, Keller MS, Bellazzi R, Hanauer DA, Serret-Larmande A, Gutierrez-Sacristan A, Holmes JJ, Bell DS, Mandl KD, Follett RW, Klann JG, Murad DA, Scudeller L, Bucalo M, Kirchoff K, Craig J, Obeid J, Jouhet V, Griffier R, Cossin S, Moal B, Patel LP, Bellasi A, Prokosch HU, Kraska D, Sliz P, Tan ALM, Ngiam KY, Zambelli A, Mowery DL, Schiver E, Devkota B, Bradford RL, Daniar M, Daniel C, Benoit V, Bey R, Paris N, Serre P, Orlova N, Dubiel J, Hilka M, Jannot AS, Breant S, Leblanc J, Griffon N, Burgun A, Bernaux M, Sandrin A, Salamanca E, Cormont S, Ganslandt T, Gradinger T, Champ J, Boeker M, Martel P, Esteve L, Gramfort A, Grisel O, Leprovost D, Moreau T, Varoquaux G, Vie JJ, Wassermann D, Mensch A, Caucheteux C, Haverkamp C, Lemaitre G, Bosari S, Krantz ID, South A, Cai T, Kohane IS. **International electronic health record-derived COVID-19 clinical course profiles: the 4CE consortium.** NPJ Digit Med. 2020;3:109. Epub 2020/08/31. doi: 10.1038/s41746-020-00308-0 308 [pii]. PubMed PMID: 32864472; PubMed Central PMCID: PMC7438496. Characterization Study, Trial

236. Burn E, You SC, Sena AG, Kostka K, Abedtash H, Abrahao MTF, Alberga A, Alghoul H, Alser O, Alshammari TM, Areia C, Banda JM, Cho J, Culhane AC, Davydov A, DeFalco FJ, Duarte-Salles T, DuVall S, Falconer T, Gao W, Golozar A, Hardin J, Hripcsak G, Huser V, Jeon H, Jing Y, Jung CY, Kaas-Hansen BS, Kaduk D, Kent S, Kim Y, Kolovos S, Lane JCE, Lee H, Lynch KE, Makadia R, Matheny ME, Mehta P, Morales DR, Natarajan K, Nyberg F, Ostropolets A, Park RW, Park J, Posada JD, Prats-Uribe A, Rao G, Reich C, Rho Y, Rijnbeek P, Sathappan SMK, Schilling LM, Schuemie M, Shah NH, Shoaibi A, Song S, Spotnitz M, Suchard MA, Swerdel JN, Vizcaya D, Volpe S, Wen H, Williams AE, Yimer BB, Zhang L, Zhuk O, Prieto-Alhambra D, Ryan P. **An international characterisation of patients hospitalised with COVID-19 and a comparison with those previously hospitalised with influenza.** medRxiv. 2020. Epub 2020/06/09. doi: 10.1101/2020.04.22.20074336. PubMed PMID: 32511443; PubMed Central PMCID: PMC7239064. Characterization Study

237. Yu Y, Ruddy KJ, Wen A, Zong N, Tsuji S, Chen J, Shah ND, Jiang G. **Integrating Electronic Health Record Data into the ADEpedia-on-OHDSI Platform for Improved Signal Detection: A Case Study of Immune-related Adverse Events.** AMIA Jt Summits Transl Sci Proc. 2020;2020:710-9. Epub 2020/06/02. PubMed PMID: 32477694; PubMed Central PMCID: PMC7233056. ETL

238. Su CC, Chia-Cheng Lai E, Kao Yang YH, Man KKC, Kubota K, Stang P, Schuemie M, Ryan P, Hardy C, Zhang Y, Kimura S, Kamijima Y, Wong ICK, Setoguchi S. **Incidence, prevalence and prescription patterns of antipsychotic medications use in Asia and US: A cross-nation comparison with common data model.** J Psychiatr Res. 2020;131:77-84. Epub 2020/09/19. doi: 10.1016/j.jpsychires.2020.08.025. PubMed PMID: 32947205. Characterization Study

239. Kumar P, Nestsiarovich A, Nelson SJ, Kerner B, Perkins DJ, Lambert CG. **Imputation and characterization of uncoded self-harm in major mental illness using machine learning.** J Am Med Inform Assoc. 2020;27(1):136-46. Epub 2019/10/28. doi: 10.1093/jamia/ocz173. PubMed PMID: 31651956; PubMed Central PMCID: PMCPMC7647246. Characterization Study

240. Liu S, Wang Y, Wen A, Wang L, Hong N, Shen F, Bedrick S, Hersh W, Liu H. **Implementation of a cohort retrieval system for clinical data repositories using the observational medical outcomes partnership common data model: Proof-of-concept system validation.** JMIR medical informatics. 2020;8(10):e17376. NLP

241. Schuemie MJ, Cepeda MS, Suchard MA, Yang J, Tian Y, Schuler A, Ryan PB, Madigan D, Hripcsak G. **How Confident Are We about Observational Findings in Healthcare: A Benchmark Study.** Harv Data Sci Rev. 2020;2(1). Epub 2020/12/29. doi: 10.1162/99608f92.147cc28e. PubMed PMID: 33367288; PubMed Central PMCID: PMC7755157. Estimation Methods

242. Reps JM, Williams RD, You SC, Falconer T, Minty E, Callahan A, Ryan PB, Park RW, Lim HS, Rijnbeek P. **Feasibility and evaluation of a large-scale external validation approach for patient-level prediction in an international data network: validation of models predicting stroke in female patients newly diagnosed with atrial fibrillation.** BMC Med Res Methodol. 2020;20(1):102. Epub 2020/05/08. doi: 10.1186/s12874-020-00991-3 10.1186/s12874-020-00991-3 [pii]. PubMed PMID: 32375693; PubMed Central PMCID: PMC7201646. Prediction Study

243. Dixon BE, Wen C, French T, Williams JL, Duke JD, Grannis SJ. **Extending an open-source tool to measure data quality: case report on Observational Health Data Science and Informatics (OHDSI).** BMJ Health Care Inform. 2020;27(1). Epub 2020/04/02. doi: bmjhci-2019-100054 [pii] 10.1136/bmjhci-2019-100054. PubMed PMID: 32229499; PubMed Central PMCID: PMC7254131. Methods Other, Data Quality

Legend

- Orange ..... Data Standards
- Red ..... Methods Research
- Green ..... Clinical Applications
- Black ..... Other

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

OHDSI Publications

244. Lovestone S. **The European medical information framework: A novel ecosystem for sharing healthcare data across Europe.** Learn Health Syst. 2020;4(2):e10214. Epub 2020/04/22. doi: 10.1002/lrh2.10214. PubMed PMID: 32313838; PubMed Central PMCID: PMCPMC7156868. Referring to

245. Tian Y, Chen W, Zhou T, Li J, Ding K, Li J. **Establishment and evaluation of a multicenter collaborative prediction model construction framework supporting model generalization and continuous improvement: A pilot study.** Int J Med Inform. 2020;141:104173. Epub 2020/06/13. doi: S1386-5056(20)30136-2 [pii] 10.1016/j.ijmedinf.2020.104173. PubMed PMID: 32531725. Prediction Methods

246. Thurin NH, Lassalle R, Schuemie M, Penichon M, Gagne JJ, Rassen JA, Benichou J, Weill A, Blin P, Moore N, Droz-Perroteau C. **Empirical assessment of case-based methods for identification of drugs associated with upper gastrointestinal bleeding in the French National Healthcare System database (SNDS).** Pharmacoepidemiol Drug Saf. 2020;29(8):890-903. Epub 2020/06/12. doi: 10.1002/pds.5038. PubMed PMID: 32524701. Estimation Study

247. Thurin NH, Lassalle R, Schuemie M, Penichon M, Gagne JJ, Rassen JA, Benichou J, Weill A, Blin P, Moore N, Droz-Perroteau C. **Empirical assessment of case-based methods for drug safety alert identification in the French National Healthcare System database (SNDS): Methodology of the ALCAPONE project.** Pharmacoepidemiol Drug Saf. 2020;29(9):993-1000. Epub 2020/03/07. doi: 10.1002/pds.4983. PubMed PMID: 32133717. Estimation Methods

248. Unberath P, Prokosch HU, Grundner J, Erpenbeck M, Maier C, Christoph J. **EHR-Independent Predictive Decision Support Architecture Based on OMOP.** Appl Clin Inform. 2020;11(3):399-404. Epub 2020/06/04. doi: 10.1055/s-0040-1710393. PubMed PMID: 32492716; PubMed Central PMCID: PMC7269719. Prediction Methods

249. Choi YI, Kim YJ, Chung JW, Kim KO, Kim H, Park RW, Park DK. **Effect of Age on the Initiation of Biologic Agent Therapy in Patients With Inflammatory Bowel Disease: Korean Common Data Model Cohort Study.** JMIR Med Inform. 2020;8(4):e15124. Epub 2020/04/16. doi: v8i4e15124 [pii] 10.2196/15124. PubMed PMID: 32293578; PubMed Central PMCID: PMC7191339. Referring to, Estimation Study

250. Kashyap M, Seneviratne M, Banda JM, Falconer T, Ryu B, Yoo S, Hripcsak G, Shah NH. **Development and validation of phenotype classifiers across multiple sites in the observational health data sciences and informatics network.** J Am Med Inform Assoc. 2020;27(6):877-83. Epub 2020/05/07. doi: 5831103 [pii] 10.1093/jamia/ocaa032. PubMed PMID: 32374408; PubMed Central PMCID: PMC7309227. Phenotyping

251. Wang Q, Reps JM, Kostka KF, Ryan PB, Zou Y, Voss EA, Rijnbeek PR, Chen R, Rao GA, Morgan Stewart H, Williams AE, Williams RD, Van Zandt M, Falconer T, Fernandez-Chas M, Vashisht R, Pfohl SR, Shah NH, Kasthurirathne SN, You SC, Jiang Q, Reich C, Zhou Y. **Development and validation of a prognostic model predicting symptomatic hemorrhagic transformation in acute ischemic stroke at scale in the OHDSI network.** PLoS One. 2020;15(1):e0226718. Epub 2020/01/08. doi: 10.1371/journal.pone.0226718 PONE-D-19-25005 [pii]. PubMed PMID: 31910437; PubMed Central PMCID: PMC6946584 Development. This does not alter our adherence to PLOS ONE policies on sharing data and materials. Prediction Study

252. Swerdel JN, Reps JM, Fife D, Ryan PB. **Developing Predictive Models to Determine Patients in End-of-Life Care in Administrative Datasets.** Drug Saf. 2020;43(5):447-55. Epub 2020/01/16. doi: 10.1007/s40264-020-00906-7 10.1007/s40264-020-00906-7 [pii]. PubMed PMID: 31939079; PubMed Central PMCID: PMC7165142. Prediction Study

253. Yu Y, Ruddy K, Mansfield A, Zong N, Wen A, Tsuji S, Huang M, Liu H, Shah N, Jiang G. **Detecting and Filtering Immune-Related Adverse Events Signal Based on Text Mining and Observational Health Data Sciences and Informatics Common Data Model: Framework Development Study.** JMIR Med Inform. 2020;8(6):e17353. Epub 2020/06/13. doi: v8i6e17353 [pii] 10.2196/17353. PubMed PMID: 32530430; PubMed Central PMCID: PMC7320306. NLP

254. Reinecke I, Gulden C, Kummel M, Nassirian A, Blasini R, Sedlmayr M. **Design for a Modular Clinical Trial Recruitment Support System Based on FHIR and OMOP.** Stud Health Technol Inform. 2020;270:158-62. Epub 2020/06/24. doi: SHTI200142 [pii] 10.3233/SHTI200142. PubMed PMID: 32570366. Phenotyping, FHIR, Trial

255. Weng C, Shah NH, Hripcsak G. **Deep phenotyping: Embracing complexity and temporality-Towards scalability, portability, and interoperability.** J Biomed Inform. 2020;105:103433. Epub 2020/04/27. doi: 10.1016/j.jbi.2020.103433. PubMed PMID: 32335224; PubMed Central PMCID: PMCPMC7179504. Phenotyping

256. Burn E, You SC, Sena AG, Kostka K, Abedtash H, Abrahão MTF, Alberga A, Alghoul H, Alser O, Alshammari TM, Aragon M, Areia C, Banda JM, Cho J, Culhane AC, Davydov A, DeFalco FJ, Duarte-Salles T, DuVall S, Falconer T, Fernandez-Bertolin S, Gao W, Golozar A, Hardin J, Hripcsak G, Huser V, Jeon H, Jing Y, Jung CY, Kaas-Hansen BS, Kaduk D, Kent S, Kim Y, Kolovos S, Lane JCE, Lee H, Lynch KE, Makadia R, Matheny ME, Mehta PP, Morales DR, Natarajan K, Nyberg F, Ostropolets A, Park RW, Park J, Posada JD, Prats-Uribe A, Rao G, Reich C, Rho Y, Rijnbeek P, Schilling LM, Schuemie M, Shah NH, Shoaibi A, Song S, Spotnitz M, Suchard MA, Swerdel JN, Vizcaya D, Volpe S, Wen H, Williams AE, Yimer BB, Zhang L, Zhuk O, Prieto-Alhambra D, Ryan P. **Deep phenotyping of 34,128 adult patients hospitalised with COVID-19 in an international network study.** Nat Commun. 2020;11(1):5009. Epub 2020/10/08. doi: 10.1038/s41467-020-18849-z. PubMed PMID: 33024121; PubMed Central PMCID: PMCPMC7538555. Characterization Study

257. Bompelli A, Li J, Xu Y, Wang N, Wang Y, Adam T, He Z, Zhang R. **Deep Learning Approach to Parse Eligibility Criteria in Dietary Supplements Clinical Trials Following OMOP Common Data Model.** AMIA Annu Symp Proc. 2020;2020:243-52. Epub 2021/05/04. doi: 056_3415023 [pii]. PubMed PMID: 33936396; PubMed Central PMCID: PMC8075443.NLP

258. Fischer P, Stohr MR, Gall H, Michel-Backofen A, Majeed RW. **Data Integration into OMOP CDM for Heterogeneous Clinical Data Collections via HL7 FHIR Bundles and XSLT.** Stud Health Technol Inform. 2020;270:138-42. Epub 2020/06/24. doi: SHTI200138 [pii] 10.3233/SHTI200138. PubMed PMID: 32570362. ETL, FHIR

259. Averitt AJ, Vanitchanant N, Ranganath R, Perotte AJ. **The Counterfactual χ-GAN: Finding comparable cohorts in observational health data.** J Biomed Inform. 2020;109:103515. Epub 2020/08/11. doi: 10.1016/j.jbi.2020.103515. PubMed PMID: 32771540. Estimation Methods

260. Ji H, Kim S, Yi S, Hwang H, Kim JW, Yoo S. **Converting clinical document architecture documents to the common data model for incorporating health information exchange data in observational health studies: CDA to CDM.** J Biomed Inform. 2020;107:103459. Epub 2020/05/30. doi: S1532-0464(20)30087-3 [pii] 10.1016/j.jbi.2020.103459. PubMed PMID: 32470694. ETL, NLP

261. Cho S, Sin M, Tsapepas D, Dale LA, Husain SA, Mohan S, Natarajan K. **Content Coverage Evaluation of the OMOP Vocabulary on the Transplant Domain Focusing on Concepts Relevant for Kidney Transplant Outcomes Analysis.** Appl Clin Inform. 2020;11(4):650-8. Epub 2020/10/08. doi: 10.1055/s-0040-1716528. PubMed PMID: 33027834; PubMed Central PMCID: PMC7557323 Kidney International Reports, outside the submitted work. ETL, Vocabulary

262. You SC, Jung S, Swerdel JN, Ryan PB, Schuemie MJ, Suchard MA, Lee S, Cho J, Hripcsak G, Park RW, Park S. **Comparison of First-Line Dual Combination Treatments in Hypertension: Real-World Evidence from Multinational Heterogeneous Cohorts.** Korean Circ J. 2020;50(1):52-68. Epub 2019/10/24. doi: 10.4070/kcj.2019.0173. PubMed PMID: 31642211; PubMed Central PMCID: PMCPMC6923236. Estimation Study

263. Hripcsak G, Suchard MA, Shea S, Chen R, You SC, Pratt N, Madigan D, Krumholz HM, Ryan PB, Schuemie MJ. **Comparison of Cardiovascular and Safety Outcomes of Chlorthalidone vs Hydrochlorothiazide to Treat Hypertension.** JAMA Intern Med. 2020;180(4):542-51. Epub 2020/02/18. doi: 10.1001/jamainternmed.2019.7454. PubMed PMID: 32065600; PubMed Central PMCID: PMCPMC7042845. Estimation Study

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

OHDSI Publications

Legend

Orange.....Data Standards
Red.....Methods Research
Green.....Clinical Applications
Black.....Other

264. Kim Y, Tian Y, Yang J, Huser V, Jin P, Lambert CG, Park H, You SC, Park RW, Rijnbeek PR, Van Zandt M, Reich C, Vashisht R, Wu Y, Duke J, Hripcsak G, Madigan D, Shah NH, Ryan PB, Schuemie MJ, Suchard MA. **Comparative safety and effectiveness of alendronate versus raloxifene in women with osteoporosis.** Sci Rep. 2020;10(1):11115. Epub 2020/07/08. doi: 10.1038/s41598-020-68037-8 10.1038/s41598-020-68037-8 [pii]. PubMed PMID: 32632237; PubMed Central PMCID: PMC7338498.Estimation Study

265. Seo SI, You SC, Park CH, Kim TJ, Ko YS, Kim Y, Yoo JJ, Kim J, Shin WG. **Comparative risk of Clostridium difficile infection between proton pump inhibitors and histamine-2 receptor antagonists: A 15-year hospital cohort study using a common data model.** Journal of gastroenterology and hepatology. 2020;35(8):1325-30.Estimation Study

266. Davidson L, Boland MR. **Comparative Analysis and Evaluation of State-of-the-Art Medication Mapping Tools to Transform a Local Medication Terminology to RxNorm.** AMIA Jt Summits Transl Sci Proc. 2020;2020:126-35. Epub 2020/06/02. PubMed PMID: 32477631; PubMed Central PMCID: PMCPMC7233099. ETL

267. Ostropolets A, Reich C, Ryan P, Weng C, Molinaro A, DeFalco F, Jonnagaddala J, Liaw ST, Jeon H, Park RW, Spotnitz ME, Natarajan K, Argyriou G, Kostka K, Miller R, Williams A, Minty E, Posada J, Hripcsak G. **Characterizing database granularity using SNOMED-CT hierarchy.** AMIA Annu Symp Proc. 2020;2020:983-92. Epub 2021/05/04. doi: 134_3416797 [pii]. PubMed PMID: 33936474; PubMed Central PMCID: PMC8075504. Characterization Methods, Vocabulary

268. Kim H, Yoo S, Jeon Y, Yi S, Kim S, Choi SA, Hwang H, Kim KJ. **Characterization of Anti-seizure Medication Treatment Pathways in Pediatric Epilepsy Using the Electronic Health Record-Based Common Data Model.** Front Neurol. 2020;11:409. Epub 2020/06/02. doi: 10.3389/fneur.2020.00409. PubMed PMID: 32477256; PubMed Central PMCID: PMC7235379. Characterization Study

269. Weinstein RB, Ryan PB, Berlin JA, Schuemie MJ, Swerdel J, Fife D. **Channeling Bias in the Analysis of Risk of Myocardial Infarction, Stroke, Gastrointestinal Bleeding, and Acute Renal Failure with the Use of Paracetamol Compared with Ibuprofen.** Drug Saf. 2020;43(9):927-42. Epub 2020/06/06. doi: 10.1007/s40264-020-00950-3. PubMed PMID: 32500272; PubMed Central PMCID: PMCPMC7434801. Estimation Study

270. Candore G, Hedenmalm K, Slattery J, Cave A, Kurz X, Arlett P. **Can We Rely on Results From IQVIA Medical Research Data UK Converted to the Observational Medical Outcome Partnership Common Data Model?: A Validation Study Based on Prescribing Codeine in Children.** Clin Pharmacol Ther. 2020;107(4):915-25. Epub 2020/01/21. doi: 10.1002/cpt.1785. PubMed PMID: 31956997; PubMed Central PMCID: PMC7158210. Characterization Study, Position

271. You SC, Rho Y, Bikdeli B, Kim J, Siapos A, Weaver J, Londhe A, Cho J, Park J, Schuemie M, Suchard MA, Madigan D, Hripcsak G, Gupta A, Reich CG, Ryan PB, Park RW, Krumholz HM. **Association of Ticagrelor vs Clopidogrel With Net Adverse Clinical Events in Patients With Acute Coronary Syndrome Undergoing Percutaneous Coronary Intervention.** Jama. 2020;324(16):1640-50. Epub 2020/10/28. doi: 10.1001/jama.2020.16167. PubMed PMID: 33107944; PubMed Central PMCID: PMCPMC7592033. Estimation Study

272. Lee SM, Kim K, Yoon J, Park SK, Moon S, Lee SE, Oh J, Yoo S, Kim K-I, Yoon H-J. **Association between use of hydrochlorothiazide and nonmelanoma skin cancer: Common data model cohort study in Asian population.** Journal of clinical medicine. 2020;9(9):2910. Estimation Study

273. Cho J, You SC, Lee S, Park D, Park B, Hripcsak G, Park RW. **Application of Epidemiological Geographic Information System: An Open-Source Spatial Analysis Tool Based on the OMOP Common Data Model.** Int J Environ Res Public Health. 2020;17(21). Epub 2020/10/30. doi: ijerph17217824 [pii] 10.3390/ijerph17217824. PubMed PMID: 33114631; PubMed Central PMCID: PMC7663469. Methods Other

274. Brauer R, Wong ICK, Man KK, Pratt NL, Park RW, Cho SY, Li YJ, Iqbal U, Nguyen PA, Schuemie M. **Application of a Common Data Model (CDM) to rank the paediatric user and prescription prevalence of 15 different drug classes in South Korea, Hong Kong, Taiwan, Japan and Australia: an observational, descriptive study.** BMJ open. 2020;10(1):e032426. Epub 2020/01/16. doi: 10.1136/bmjopen-2019-032426. PubMed PMID: 31937652; PubMed Central PMCID: PMCPMC7044847. Characterization Study

275. Choi SA, Kim H, Kim S, Yoo S, Yi S, Jeon Y, Hwang H, Kim KJ. **Analysis of antiseizure drug-related adverse reactions from the electronic health record using the common data model.** Epilepsia. 2020;61(4):610-6. Characterization Study

276. Ostropolets A, Reich C, Ryan P, Shang N, Hripcsak G, Weng C. **Adapting electronic health records-derived phenotypes to claims data: Lessons learned in using limited clinical data for phenotyping.** J Biomed Inform. 2020;102:103363. Epub 2019/12/24. doi: S1532-0464(19)30283-7 [pii] 10.1016/j.jbi.2019.103363. PubMed PMID: 31866433; PubMed Central PMCID: PMC7390483. Phenotyping

277. Yuan Z, DeFalco F, Wang L, Hester L, Weaver J, Swerdel JN, Freedman A, Ryan P, Schuemie M, Qiu R, Yee J, Meininger G, Berlin JA, Rosenthal N. **Acute pancreatitis risk in type 2 diabetes patients treated with canagliflozin versus other antihyperglycemic agents: an observational claims database study.** Curr Med Res Opin. 2020;36(7):1117-24. Epub 2020/04/28. doi: 10.1080/03007995.2020.1761312. PubMed PMID: 32338068. Estimation Study

278. Bhuyan P, Medin J, da Silva HG, Yadavalli M, Shankar NK, Mullerova H, Arnold M, Nord M. **Very rare thrombosis with thrombocytopenia after second AZD1222 dose: a global safety database analysis.** The Lancet. 2021. doi: 10.1016/S0140-6736(21)01693-7. Active Surveillance, Characterization Study

279. Prats-Uribe A, Sena AG, Lai LYH, Ahmed WU, Alghoul H, Alser O, Alshammari TM, Areia C, Carter W, Casajust P, Dawoud D, Golozar A, Jonnagaddala J, Mehta PP, Gong M, Morales DR, Nyberg F, Posada JD, Recalde M, Roel E, Shah K, Shah NH, Schilling LM, Subbian V, Vizcaya D, Zhang L, Zhang Y, Zhu H, Liu L, Cho J, Lynch KE, Matheny ME, You SC, Rijnbeek PR, Hripcsak G, Lane JC, Burn E, Reich C, Suchard MA, Duarte-Salles T, Kostka K, Ryan PB, Prieto-Alhambra D. **Use of repurposed and adjuvant drugs in hospital patients with covid-19: multinational network cohort study.** bmj. 2021;373:n1038. Epub 2021/05/13. doi: 10.1136/bmj.n1038. PubMed PMID: 33975825; PubMed Central PMCID: PMC8111167. Characterization Study

280. Wang Y, Makadia R, Knoll C, Hardin J, Voss EA, Fife D, Davis K, Sloan S. **Understanding patient journey in ulcerative colitis prior to biologic initiation: a 5-year exploration.** BMC gastroenterology. 2021;21(1):1-8. Characterization Study

281. Almeida JR, Silva JF, Matos S, Oliveira JL. **A two-stage workflow to extract and harmonize drug mentions from clinical notes into observational databases.** J Biomed Inform. 2021;120:103849. Epub 2021/07/03. doi: 10.1016/j.jbi.2021.103849. PubMed PMID: 34214696. ETL, NLP

282. Kim JW, Kim S, Ryu B, Song W, Lee HY, Yoo S. **Transforming electronic health record polysomnographic data into the Observational Medical Outcome Partnership's Common Data Model: a pilot feasibility study.** Sci Rep. 2021;11(1):7013. Epub 2021/03/31. doi: 10.1038/s41598-021-86564-w 10.1038/s41598-021-86564-w [pii]. PubMed PMID: 33782494; PubMed Central PMCID: PMC8007756. ETL, Vocabulary

283. Kim JH, Ta CN, Liu C, Sung C, Butler AM, Stewart LA, Ena L, Rogers JR, Lee J, Ostropolets A, Ryan PB, Liu H, Lee SM, Elkind MSV, Weng C. **Towards clinical data-driven eligibility criteria optimization for interventional COVID-19 clinical trials.** J Am Med Inform Assoc. 2021;28(1):14-22. Epub 2020/12/02. doi: 10.1093/jamia/ocaa276. PubMed PMID: 33260201; PubMed Central PMCID: PMCPMC7798960. Trial Feasibility

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

284. Fife D, Blacketer C, Knight RK, Weaver J. **Stroke risk among elderly users of haloperidol and typical antipsychotics versus atypical antipsychotics: a real-world study from a US health insurance claims database.** The American Journal of Geriatric Psychiatry. 2021;29(5):499-510. Estimation Study

285. Yoon JY, Cha JM, Kim HI, Kwak MS. **Seasonal variation of peptic ulcer disease, peptic ulcer bleeding, and acute pancreatitis: A nationwide population-based study using a common data model.** Medicine. 2021;100(21). Characterization Study

286. Kim GL, Yi YH, Hwang HR, Kim J, Park Y, Kim YJ, Lee JG, Tak YJ, Lee SH, Lee SY, Cho YH, Park EJ, Lee Y. **The Risk of Osteoporosis and Osteoporotic Fracture Following the Use of Irritable Bowel Syndrome Medical Treatment: An Analysis Using the OMOP CDM Database.** J Clin Med. 2021;10(9). Epub 2021/06/03. doi: jcm10092044 [pii] 10.3390/jcm10092044. PubMed PMID: 34068814; PubMed Central PMCID: PMC8126251. Estimation Study, Characterization Study

287. Lane JC, Weaver J, Kostka K, Duarte-Salles T, Abrahao MTF, Alghoul H, Alser O, Alshammari TM, Areia C, Biedermann P. **Risk of depression, suicide and psychosis with hydroxychloroquine treatment for rheumatoid arthritis: a multinational network cohort study.** Rheumatology. 2021;60(7):3222-34. Characterization Study

288. Morales DR, Conover MM, You SC, Pratt N, Kostka K, Duarte-Salles T, Fernandez-Bertolin S, Aragon M, DuVall SL, Lynch K, Falconer T, van Bochove K, Sung C, Matheny ME, Lambert CG, Nyberg F, Alshammari TM, Williams AE, Park RW, Weaver J, Sena AG, Schuemie MJ, Rijnbeek PR, Williams RD, Lane JCE, Prats-Uribe A, Zhang L, Areia C, Krumholz HM, Prieto-Alhambra D, Ryan PB, Hripcsak G, Suchard MA. **Renin-angiotensin system blockers and susceptibility to COVID-19: an international, open science, cohort analysis.** Lancet Digit Health. 2021;3(2):e98-e114. Epub 2020/12/22. doi: S2589-7500(20)30289-2 [pii] 10.1016/S2589-7500(20)30289-2. PubMed PMID: 33342753; PubMed Central PMCID: PMC7834915. Estimation Study

289. Schuemie MJ, Weinstein R, Ryan PB, Berlin JA. **Quantifying bias in epidemiologic studies evaluating the association between acetaminophen use and cancer.** Regul Toxicol Pharmacol. 2021;120:104866. Epub 2021/01/18. doi: 10.1016/j.yrtph.2021.104866. PubMed PMID: 33454352. Estimation Methods

290. Vogelsang RP, Bojesen RD, Hoelmich ER, Orhan A, Buzquurz F, Cai L, Grube C, Zahid JA, Allakhverdiiev E, Raskov HH, Drakos I, Derian N, Ryan PB, Rijnbeek PR, Gögenur I. **Prediction of 90-day mortality after surgery for colorectal cancer using standardized nationwide quality-assurance data.** BJS Open. 2021;5(3). Epub 2021/05/09. doi: 10.1093/bjsopen/zrab023. PubMed PMID: 33963368; PubMed Central PMCID: PMCPMC8105588 .Prediction Study

291. Chen Z, Liu H, Butler A, Ostropolets A, Weng C. **Potential Role of Clinical Trial Eligibility Criteria in Electronic Phenotyping.** Stud Health Technol Inform. 2021;281:148-52. Epub 2021/05/28. doi: SHTI210138 [pii] 10.3233/SHTI210138. PubMed PMID: 34042723. Methods Other, Trial

292. Dimitriadis VK, Gavriilidis GI, Natsiavas P. **Pharmacovigilance and Clinical Environment: Utilizing OMOP-CDM and OHDSI Software Stack to Integrate EHR Data.** Stud Health Technol Inform. 2021;281:555-9. Epub 2021/05/28. doi: SHTI210232 [pii] 10.3233/SHTI210232. PubMed PMID: 34042637. ETL, Active Surveillance

293. Maier C, Kapsner LA, Mate S, Prokosch HU, Kraus S. **Patient Cohort Identification on Time Series Data Using the OMOP Common Data Model.** Appl Clin Inform. 2021;12(1):57-64. Epub 2021/01/29. doi: 10.1055/s-0040-1721481. PubMed PMID: 33506478; PubMed Central PMCID: PMC7840432.Methods Other

294. Ge J, Pletcher MJ, Lai JC. **Outcomes of SARS-CoV-2 Infection in Patients with Chronic Liver Disease and Cirrhosis: a N3C Study.** Gastroenterology. 2021. Epub 2021/07/21. doi: 10.1053/j.gastro.2021.07.010. PubMed PMID: 34284037; PubMed Central PMCID: PMCPMC8286237. Characterization Study

295. Sharafeldin N, Bates B, Song Q, Madhira V, Yan Y, Dong S, Lee E, Kuhrt N, Shao YR, Liu F, Bergquist T, Guinney J, Su J, Topaloglu U. **Outcomes of COVID-19 in Patients With Cancer: Report From the National COVID Cohort Collaborative (N3C).** J Clin Oncol. 2021;39(20):2232-46. Epub 2021/06/05. doi: 10.1200/jco.21.01074. PubMed PMID: 34085538; PubMed Central PMCID: PMCPMC8260918. Characterization Study

296. Shin H, Lee S. **An OMOP-CDM based pharmacovigilance data-processing pipeline (PDP) providing active surveillance for ADR signal detection from real-world data sources.** BMC Med Inform Decis Mak. 2021;21(1):159. Epub 2021/05/19. doi: 10.1186/s12911-021-01520-y 10.1186/s12911-021-01520-y [pii]. PubMed PMID: 34001114; PubMed Central PMCID: PMC8130307.Active Surveillance, Methods Other

297. Burn E, Tebé C, Fernandez-Bertolin S, Aragon M, Recalde M, Roel E, Prats-Uribe A, Prieto-Alhambra D, Duarte-Salles T. **The natural history of symptomatic COVID-19 during the first wave in Catalonia.** Nature communications. 2021;12(1):1-12. Characterization Study

298. Gaudet-Blavignac C, Raisaro JL, Touré V, Österle S, Crameri K, Lovis C. **A National, Semantic-Driven, Three-Pillar Strategy to Enable Health Data Secondary Usage Interoperability for Research Within the Swiss Personalized Health Network: Methodological Study.** JMIR Med Inform. 2021;9(6):e27591. Epub 2021/06/30. doi: 10.2196/27591. PubMed PMID: 34185008; PubMed Central PMCID: PMCPMC8277320. Characterization Methods, Methods Other

299. Haendel MA, Chute CG, Bennett TD, Eichmann DA, Guinney J, Kibbe WA, Payne PR, Pfaff ER, Robinson PN, Saltz JH. **The National COVID Cohort Collaborative (N3C): rationale, design, infrastructure, and deployment.** Journal of the American Medical Informatics Association. 2021;28(3):427-43. ETL

300. Kim C, You SC, Reps JM, Cheong JY, Park RW. **Machine-learning model to predict the cause of death using a stacking ensemble method for observational data.** Journal of the American Medical Informatics Association. 2021;28(6):1098-107. Prediction Methods

301. Liu H, Chi Y, Butler A, Sun Y, Weng C. **A knowledge base of clinical trial eligibility criteria.** J Biomed Inform. 2021;117:103771. Epub 2021/04/05. doi: S1532-0464(21)00100-3 [pii] 10.1016/j.jbi.2021.103771. PubMed PMID: 33813032. Methods Other, Trial

302. Blacketer C, Defalco FJ, Ryan PB, Rijnbeek PR. **Increasing trust in real-world evidence through evaluation of observational data quality.** J Am Med Inform Assoc. 2021. Epub 2021/07/28. doi: 10.1093/jamia/ocab132. PubMed PMID: 34313749.Data Quality

303. Seong Y, You SC, Ostropolets A, Rho Y, Park J, Cho J, Dymshyts D, Reich CG, Heo Y, Park RW. **Incorporation of Korean Electronic Data Interchange Vocabulary into Observational Medical Outcomes Partnership Vocabulary.** Healthc Inform Res. 2021;27(1):29-38. Epub 2021/02/22. doi: hir.2021.27.1.29 [pii] 10.4258/hir.2021.27.1.29. PubMed PMID: 33611874; PubMed Central PMCID: PMC7921574. Vocabulary, ETL

304. Wang L, Swerdel JN, Weaver J, Weiss B, Pan G, Yuan Z, DiBattiste PM. **Incidence rate of hospitalization and mortality in the first year following initial diagnosis of cardiac amyloidosis in the US claims databases.** Current Medical Research and Opinion. 2021:1-11. Characterization Study

305. Reps JM, Kim C, Williams RD, Markus AF, Yang C, Duarte-Salles T, Falconer T, Jonnagaddala J, Williams A, Fernández-Bertolín S, DuVall SL, Kostka K, Rao G, Shoaibi A, Ostropolets A, Spotnitz ME, Zhang L, Casajust P, Steyerberg EW, Nyberg F, Kaas-Hansen BS, Choi YH, Morales D, Liaw ST, Abrahão MTF, Areia C, Matheny ME, Lynch KE, Aragón M, Park RW, Hripcsak G, Reich CG, Suchard MA, You SC, Ryan PB, Prieto-Alhambra D, Rijnbeek PR. **Implementation of the COVID-19 Vulnerability Index Across an International Network of Health Care Data Sets: Collaborative External Validation Study.** JMIR Med Inform. 2021;9(4):e21547. Epub 2021/03/05. doi: 10.2196/21547. PubMed PMID: 33661754; PubMed Central PMCID: PMCPMC8023380.Prediction Study

306. Majeed RW, Stohr MR, Gunther A. HIStream-Import: A Generic. **ETL Framework for Processing Arbitrary Patient Data Collections or Hospital Information Systems into HL7 FHIR Bundles.** Stud Health Technol Inform. 2021;278:75-9. Epub 2021/05/28. doi: SHTI210053 [pii] 10.3233/SHTI210053. PubMed PMID: 34042878. ETL, FHIR

2012	2013	2014	2015	2016	2017	2018	2019	2020	Thru July '21
18	35	14	16	22	28	36	42	72	67

OHDSI Publications

Legend

Orange.....Data Standards
Red.....Methods Research
Green.....Clinical Applications
Black.....Other

307. Kim HI, Yoon JY, Kwak MS, Cha JM. **Gastrointestinal and Nongastrointestinal Complications of Esophagogastroduodenoscopy and Colonoscopy in the Real World: A Nationwide Standard Cohort Using the Common Data Model Database.** Gut and Liver. 2021. Characterization Study

308. Rinaldi E, Thun S. From OpenEHR to FHIR and OMOP Data Model for Microbiology Findings. Stud Health Technol Inform. 2021;281:402-6. Epub 2021/05/28. doi: SHTI210189 [pii] 10.3233/SHTI210189. PubMed PMID: 34042774. ETL, FHIR

309. Park J, You SC, Jeong E, Weng C, Park D, Roh J, Lee DY, Cheong JY, Choi JW, Kang M, Park RW. **A Framework (SOCRATex) for Hierarchical Annotation of Unstructured Electronic Health Records and Integration Into a Standardized Medical Database: Development and Usability Study.** JMIR Med Inform. 2021;9(3):e23983. Epub 2021/03/31. doi: v9i3e23983 [pii] 10.2196/23983. PubMed PMID: 33783361; PubMed Central PMCID: PMC8044740. ETL, NLP

310. Belenkaya R, Gurley MJ, Golozar A, Dymshyts D, Miller RT, Williams AE, Ratwani S, Siapos A, Korsik V, Warner J, Campbell WS, Rivera D, Banokina T, Modina E, Bethusamy S, Stewart HM, Patel M, Chen R, Falconer T, Park RW, You SC, Jeon H, Shin SJ, Reich C. **Extending the OMOP Common Data Model and Standardized Vocabularies to Support Observational Cancer Research.** JCO Clin Cancer Inform. 2021;5:12-20. Epub 2021/01/08. doi: 10.1200/CCI.20.00079. PubMed PMID: 33411620; PubMed Central PMCID: PMC8140810. ETL, Vocabulary

311. Schüttler C, Prokosch HU, Sedlmayr M, Sedlmayr B. **Evaluation of Three Feasibility Tools for Identifying Patient Data and Biospecimen Availability: Comparative Usability Study.** JMIR Med Inform. 2021;9(7):e25531. Epub 2021/07/22. doi: 10.2196/25531. PubMed PMID: 34287211. Characterization Methods

312. Oh S, Sung M, Rhee Y, Hong N, Park YR. **Evaluation of the Privacy Risks of Personal Health Identifiers and Quasi-Identifiers in a Distributed Research Network: Development and Validation Study.** JMIR Med Inform. 2021;9(5):e24940. Epub 2021/06/01. doi: v9i5e24940 [pii] 10.2196/24940. PubMed PMID: 34057426. Methods Other

313. Giangreco NP, Tatonetti NP. **Evaluating risk detection methods to uncover ontogenic-mediated adverse drug effect mechanisms in children.** BioData Min. 2021;14(1):34. Epub 2021/07/24. doi: 10.1186/s13040-021-00264-9. PubMed PMID: 34294093. Methods Other

314. Thurin NH, Lassalle R, Schuemie M, Penichon M, Gagne JJ, Rassen JA, Benichou J, Weill A, Blin P, Moore N, Droz-Perroteau C. **Empirical assessment of case-based methods for identification of drugs associated with acute liver injury in the French National Healthcare System database (SNDS).** Pharmacoepidemiol Drug Saf. 2021;30(3):320-33. Epub 2020/10/26. doi: 10.1002/pds.5161. PubMed PMID: 33099844. Referring to, Estimation Study

315. Reps JM, Rijnbeek P, Cuthbert A, Ryan PB, Pratt N, Schuemie M. **An empirical analysis of dealing with patients who are lost to follow-up when developing prognostic models using a cohort design.** BMC Med Inform Decis Mak. 2021;21(1):43. Epub 2021/02/08. doi: 10.1186/s12911-021-01408-x. PubMed PMID: 33549087; PubMed Central PMCID: PMCPMC7866757. Prediction Methods, Prediction Study

316. Hripcsak G, Schuemie MJ, Madigan D, Ryan PB, Suchard MA. **Drawing Reproducible Conclusions from Observational Clinical Data with OHDSI.** Yearb Med Inform. 2021. Epub 2021/04/22. doi: 10.1055/s-0041-1726481. PubMed PMID: 33882595. Position

317. Mamidi TKK, Tran-Nguyen TK, Melvin RL, Worthey EA. **Development of An Individualized Risk Prediction Model for COVID-19 Using Electronic Health Record Data.** Front Big Data. 2021;4:675882. Epub 2021/06/22. doi: 10.3389/fdata.2021.675882. PubMed PMID: 34151259; PubMed Central PMCID: PMCPMC8211871. Prediction Study

318. Sivesind TE, Runion T, Branda M, Schilling LM, Dellavalle RP. **Dermatologic Research Potential of the Observational Health Data Sciences and Informatics (OHDSI) Network.** Dermatology. 2021:1-9. Epub 2021/03/19. doi: 000514536 [pii] 10.1159/000514536. PubMed PMID: 33735862. Position

319. Kang B, Yoon J, Kim HY, Jo SJ, Lee Y, Kam HJ. **Deep-learning-based automated terminology mapping in OMOP-CDM.** J Am Med Inform Assoc. 2021. Epub 2021/05/15. doi: 6275415 [pii] 10.1093/jamia/ocab030. PubMed PMID: 33987667. ETL, Vocabulary

320. Ostropolets A, Zachariah P, Ryan P, Chen R, Hripcsak G. **Data Consult Service: Can we use observational data to address immediate clinical needs?** Journal of the American Medical Informatics Association. 2021. doi: 10.1093/jamia/ocab122.Estimation Study; Characterization Study

321. Wang J, Abu-El-Rub N, Gray J, Pham HA, Zhou Y, Manion FJ, Liu M, Song X, Xu H, Rouhizadeh M, Zhang Y. **COVID-19 SignSym: a fast adaptation of a general clinical NLP tool to identify and normalize COVID-19 signs and symptoms to OMOP common data model.** J Am Med Inform Assoc. 2021. Epub 2021/03/07. doi: 6155732 [pii] 10.1093/jamia/ocab015. PubMed PMID: 33674830; PubMed Central PMCID: PMC7989301. ETL, NLP

322. Tan EH, Sena AG, Prats-Uribe A, You SC, Ahmed WU, Kostka K, Reich C, Duvall SL, Lynch KE, Matheny ME, Duarte-Salles T, Bertolin SF, Hripcsak G, Natarajan K, Falconer T, Spotnitz M, Ostropolets A, Blacketer C, Alshammari TM, Alghoul H, Alser O, Lane JCE, Dawoud DM, Shah K, Yang Y, Zhang L, Areia C, Golozar A, Recalde M, Casajust P, Jonnagaddala J, Subbian V, Vizcaya D, Lai LYH, Nyberg F, Morales DR, Posada JD, Shah NH, Gong M, Vivekanantham A, Abend A, Minty EP, Suchard M, Rijnbeek P, Ryan PB, Prieto-Alhambra D. **COVID-19 in patients with autoimmune diseases: characteristics and outcomes in a multinational network of cohorts across three countries.** Rheumatology (Oxford). 2021. Epub 2021/03/17. doi: 6174122 [pii] 10.1093/rheumatology/keab250. PubMed PMID: 33725121; PubMed Central PMCID: PMC7989171. Characterization Study

323. Chan You S, Krumholz HM, Suchard MA, Schuemie MJ, Hripcsak G, Chen R, Shea S, Duke J, Pratt N, Reich CG, Madigan D, Ryan PB, Woong Park R, Park S. **Comprehensive Comparative Effectiveness and Safety of First-Line β-Blocker Monotherapy in Hypertensive Patients: A Large-Scale Multicenter Observational Study.** Hypertension. 2021;77(5):1528-38. Epub 2021/03/30. doi: 10.1161/hypertensionaha.120.16402. PubMed PMID: 33775125; PubMed Central PMCID: PMCPMC8035236. Estimation Study

324. Jensen ET, Dabelea DA, Praveen PA, Amutha A, Hockett CW, Isom SP, Ong TC, Mohan V, D'Agostino R, Jr., Kahn MG, Hamman RF, Wadwa P, Dolan L, Lawrence JM, Madhu SV, Chhokar R, Goel K, Tandon N, Mayer-Davis E. **Comparison of the incidence of diabetes in United States and Indian youth: An international harmonization of youth diabetes registries.** Pediatr Diabetes. 2021;22(1):8-14. Epub 2020/03/21. doi: 10.1111/pedi.13009. PubMed PMID: 32196874; PubMed Central PMCID: PMC7748376. Characterization Study

325. Cronin RM, Halvorson AE, Springer C, Feng X, Sulieman L, Loperena-Cortes R, Mayo K, Carroll RJ, Chen Q, Ahmedani BK. **Comparison of family health history in surveys vs electronic health record data mapped to the observational medical outcomes partnership data model in the All of Us Research Program.** Journal of the American Medical Informatics Association. 2021;28(4):695-703. ETL

326. Rogers JR, Liu C, Hripcsak G, Cheung YK, Weng C. **Comparison of Clinical Characteristics Between Clinical Trial Participants and Nonparticipants Using Electronic Health Record Data.** JAMA network open. 2021;4(4):e214732-e.Trial Feasibility

327. Chen R, Suchard MA, Krumholz HM, Schuemie MJ, Shea S, Duke J, Pratt N, Reich CG, Madigan D, You SC, Ryan PB, Hripcsak G. **Comparative First-Line Effectiveness and Safety of ACE (Angiotensin-Converting Enzyme) Inhibitors and Angiotensin Receptor Blockers: A Multinational Cohort Study.** Hypertension. 2021:HYPERTENSIONAHA12016667. Epub 2021/07/27. doi: 10.1161/HYPERTENSIONAHA.120.16667. PubMed PMID: 34304580. Active Surveillance

328. Lee J, Liu C, Kim JH, Butler A, Shang N, Pang C, Natarajan K, Ryan P, Ta C, Weng C. **Comparative effectiveness of medical concept embedding for feature engineering in phenotyping.** JAMIA open. 2021;4(2):ooab028. Epub 2021/06/19. doi: 10.1093/jamiaopen/ooab028. PubMed PMID: 34142015; PubMed Central PMCID: PMCPMC8206403. Phenotyping

329. Shoaibi A, Fortin SP, Weinstein R, Berlin JA, Ryan P. **Comparative Effectiveness of Famotidine in Hospitalized COVID-19 Patients.** Am J Gastroenterol. 2021;116(4):692-9. Epub 2021/05/14. doi: 10.14309/ajg.0000000000001153. PubMed PMID: 33982938. Estimation Study

330. Kent S, Burn E, Dawoud D, Jonsson P, Ostby JT, Hughes N, Rijnbeek P, Bouvy JC. **Common Problems, Common Data Model Solutions: Evidence Generation for Health Technology Assessment.** PharmacoEconomics. 2021;39(3):275-85. Epub 2020/12/19. doi: 10.1007/s40273-020-00981-9 10.1007/s40273-020-00981-9 [pii]. PubMed PMID: 33336320; PubMed Central PMCID: PMC7746423. Position

331. Bennett TD, Moffitt RA, Hajagos JG, Amor B, Anand A, Bissell MM, Bradwell KR, Bremer C, Byrd JB, Denham A, DeWitt PE, Gabriel D, Garibaldi BT, Girvin AT, Guinney J, Hill EL, Hong SS, Jimenez H, Kavuluru R, Kostka K, Lehmann HP, Levitt E, Mallipattu SK, Manna A, McMurry JA, Morris M, Muschelli J, Neumann AJ, Palchuk MB, Pfaff ER, Qian Z, Qureshi N, Russell S, Spratt H, Walden A, Williams AE, Wooldridge JT, Yoo YJ, Zhang XT, Zhu RL, Austin CP, Saltz JH, Gersing KR, Haendel MA, Chute CG, National CCCC. **Clinical Characterization and Prediction of Clinical Severity of SARS-CoV-2 Infection Among US Adults Using Data From the US National COVID Cohort Collaborative.** JAMA Netw Open. 2021;4(7):e2116901. Epub 2021/07/14. doi: 10.1001/jamanetworkopen.2021.16901. PubMed PMID: 34255046; PubMed Central PMCID: PMCPMC8278272. Characterization Study, Prediction Study

332. Jeon H, You SC, Kang SY, Seo SI, Warner JL, Belenkaya R, Park RW. **Characterizing the Anticancer Treatment Trajectory and Pattern in Patients Receiving Chemotherapy for Cancer Using Harmonized Observational Databases: Retrospective Study.** JMIR Med Inform. 2021;9(4):e25035. Epub 2021/03/16. doi: v9i4e25035 [pii] 10.2196/25035. PubMed PMID: 33720842; PubMed Central PMCID: PMC8058693. Methods Other

333. Roel E, Pistillo A, Recalde M, Sena AG, Fernandez-Bertolin S, Aragon M, Puente D, Ahmed WU, Alghoul H, Alser O, Alshammari TM, Areia C, Blacketer C, Carter W, Casajust P, Culhane AC, Dawoud D, DeFalco F, DuVall SL, Falconer T, Golozar A, Gong M, Hester L, Hripcsak G, Tan EH, Jeon H, Jonnagaddala J, Lai LY, Lynch KE, Matheny ME, Morales DR, Natarajan K, Nyberg F, Ostropolets A, Posada JD, Prats-Uribe A, Reich CG, Rivera DR, Schilling LM, Soerjomataram I, Shah K, Shah NH, Shen Y, Spotnitz M, Subbian V, Suchard MA, Trama A, Zhang L, Zhang Y, Ryan PB, Prieto-Alhambra D, Kostka K, Duarte-Salles T. **Characteristics and outcomes of over 300,000 COVID-19 individuals with history of cancer in the United States and Spain.** Cancer Epidemiol Biomarkers Prev. 2021. Epub 2021/07/18. doi: 10.1158/1055-9965.EPI-21-0266. PubMed PMID: 34272262. Characterization Study

334. Recalde M, Roel E, Pistillo A, Sena AG, Prats-Uribe A, Ahmed WU, Alghoul H, Alshammari TM, Alser O, Areia C, Burn E, Casajust P, Dawoud D, DuVall SL, Falconer T, Fernández-Bertolín S, Golozar A, Gong M, Lai LYH, Lane JCE, Lynch KE, Matheny ME, Mehta PP, Morales DR, Natarjan K, Nyberg F, Posada JD, Reich CG, Rijnbeek PR, Schilling LM, Shah K, Shah NH, Subbian V, Zhang L, Zhu H, Ryan P, Prieto-Alhambra D, Kostka K, Duarte-Salles T. **Characteristics and outcomes of 627 044 COVID-19 patients living with and without obesity in the United States, Spain, and the United Kingdom.** Int J Obes (Lond). 2021:1-11. Epub 2021/07/17. doi: 10.1038/s41366-021-00893-4. PubMed PMID: 34267326; PubMed Central PMCID: PMCPMC8281807. Characterization Study

335. Li X, Ostropolets A, Makadia R, Shoaibi A, Rao G, Sena AG, Martinez-Hernandez E, Delmestri A, Verhamme K, Rijnbeek PR, Duarte-Salles T, Suchard MA, Ryan PB, Hripcsak G, Prieto-Alhambra D. **Characterising the background incidence rates of adverse events of special interest for covid-19 vaccines in eight countries: multinational network cohort study.** bmj. 2021;373:n1435. doi: 10.1136/bmj.n1435. Characterization Study, Estimation Study

336. Reps JM, Kim C, Williams RD, Markus AF, Yang C, Salles TD, Falconer T, Jonnagaddala J, Williams A, Fernández-Bertolín S. **Can we trust the prediction model? Illustrating the importance of external validation by implementing the COVID-19 Vulnerability (C-19) Index across an international network of observational healthcare datasets.** JMIR medical informatics. 2021;9(4). Prediction Methods

337. Sun Y, Butler A, Stewart LA, Liu H, Yuan C, Southard CT, Kim JH, Weng C. **Building an OMOP common data model-compliant annotated corpus for COVID-19 clinical trials.** J Biomed Inform. 2021;118:103790. Epub 2021/04/23. doi: S1532-0464(21)00119-2 [pii] 10.1016/j.jbi.2021.103790. PubMed PMID: 33887457; PubMed Central PMCID: PMC8079156. Methods Other, Trial

338. Lenert LA, Ilatovskiy AV, Agnew J, Rudsill P, Jacobs J, Weatherston D, Deans K. **Automated Production of Research Data Marts from a Canonical Fast Healthcare Interoperability Resource (FHIR) Data Repository: Applications to COVID-19 Research.** J Am Med Inform Assoc. 2021. Epub 2021/05/17. doi: 6276433 [pii] 10.1093/jamia/ocab108. PubMed PMID: 33993254. ETL, FHIR

339. Boudis F, Clement G, Bruandet A, Lamer A. **Automated Generation of Individual and Population Clinical Pathways with the OMOP Common Data Model.** Stud Health Technol Inform. 2021;281:218-22. Epub 2021/05/28. doi: SHTI210152 [pii] 10.3233/SHTI210152. PubMed PMID: 34042737. Characterization Methods

340. Cepeda MS, Kern DM, Canuso CM. **At baseline patients treated with esketamine have higher burden of disease than other patients with treatment resistant depression: Learnings from a population based study.** Depression and anxiety. 2021;38(5):521-7. Characterization Study

341. Fortin SP, Johnston SS, Schuemie MJ. **Applied comparison of large-scale propensity score matching and cardinality matching for causal inference in observational research.** BMC medical research methodology. 2021;21(1):1-11. Estimation Methods

342. Zoch M, Gierschner C, Peng Y, Gruhl M, Leutner LA, Sedlmayr M, Bathelt F. **Adaption of the OMOP CDM for Rare Diseases.** Stud Health Technol Inform. 2021;281:138-42. Epub 2021/05/28. doi: SHTI210136 [pii] 10.3233/SHTI210136. PubMed PMID: 34042721. ETL, Vocabulary

343. Majeed RW, Fischer P, Gunther A. **Accessing OMOP Common Data Model Repositories with the i2b2 Webclient - Algorithm for Automatic Query Translation.** Stud Health Technol Inform. 2021;278:251-9. Epub 2021/05/28. doi: SHTI210077 [pii] 10.3233/SHTI210077. PubMed PMID: 34042902.Methods Other

344. Duarte-Salles T, Vizcaya D, Pistillo A, Casajust P, Sena AG, Lai LYH, Prats-Uribe A, Ahmed WU, Alshammari TM, Alghoul H, Alser O, Burn E, You SC, Areia C, Blacketer C, DuVall S, Falconer T, Fernandez-Bertolin S, Fortin S, Golozar A, Gong M, Tan EH, Huser V, Iveli P, Morales DR, Nyberg F, Posada JD, Recalde M, Roel E, Schilling LM, Shah NH, Shah K, Suchard MA, Zhang L, Zhang Y, Williams AE, Reich CG, Hripcsak G, Rijnbeek P, Ryan P, Kostka K, Prieto-Alhambra D. **30-Day Outcomes of Children and Adolescents With COVID-19: An International Experience.** Pediatrics. 2021. Epub 2021/05/30. doi: 10.1542/peds.2020-042929. PubMed PMID: 34049958. Characterization Study

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-18849-z OPEN

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

Edward Burn et al.#

Comparative Effectiveness of Famotidine in Hospitalized COVID-19 Patients

International Journal of Obesity

ARTICLE

Epidemiology and Population Health

Characteristics and outcomes of 627 044 COVID-19 patients living with and without obesity in the United States, Spain, and the United Kingdom

Renin–angiotensin system blockers and susceptibility to COVID-19: an international, open science, cohort analysis

Daniel R Morales, Mitchell M Conover, Seng Chan You, Nicole Pratt, Kristin Kostka, Talita Duarte-Salles, Sergio Fernández-Bertolín, Maria Aragón, Scott L DuVall, Kristine Lynch, Thomas Falconer, Kees van Bochove, Cynthia Sung, Michael E Matheny, Christophe G Lambert, Fredrik Nyberg, Thamir M Alshammari, Andrew E Williams, Rae Woong Park, James Weaver, Anthony G Sena, Martijn J Schuemie, Peter R Rijnbeek, Ross D Williams, Jennifer C E Lane, Albert Prats-Uribe, Lin Zhang, Carlos Areia, Harlan M Krumholz, Daniel Prieto-Alhambra, Patrick B Ryan, George Hripcsak, Marc A Suchard

Summary

Background Angiotensin-converting enzyme inhibitors (ACEIs) and angiotensin receptor blockers (ARBs) have been postulated to affect susceptibility to COVID-19. Observational studies so far have lacked rigorous ascertainment adjustment and international generalisability. We aimed to determine whether use of ACEIs or ARBs is associated with an increased susceptibility to COVID-19 in patients with hypertension.

Lancet Digit Health 2021; 3: e98–134
Published Online December 17, 2020 https://doi.org/10.1016/

Risk of hydroxychloroquine alone and in combination with azithromycin in the treatment of rheumatoid arthritis: a multinational, retrospective study

Jennifer C E Lane*, James Weaver*, Kristin Kostka, Talita Duarte-Salles, Maria Tereza F Abrahao, Heba Alghoul, Osaid Alser, Thamir M Alshammari, Patricia Biedermann, Juan M Banda, Edward Burn, Paula Casajust, Mitchell M Conover, Aedin C Culhane, Alexander Davydov, Scott L DuVall, Dmitry Dymshyts, Sergio Fernandez-Bertolin, Kristina Fitter, Jill Hardin, Laura Hester, George Hripcsak, Benjamin Skov Kaas-Hansen, Seamus Kent, Sejan Khosla, Spyros Kolovos, Christophe G Lambert, Johan van der Lei, Kristine E Lynch, Rupa Makadia, Andrea V Margulis, Michael E Matheny, Paras Mehta, Daniel R Morales, Henry Morgan-Stewart, Mees Mosseveld, Danielle Newby, Fredrik Nyberg, Anna Ostropolets, Rae Woong Park, Albert Prats-Uribe, Gowtham A Rao, Christian Reich, Jenna Reps, Peter Rijnbeek, Selva Muthu Kumaran Sathappan, Martijn Schuemie, Sarah Seager, Anthony G Sena, Azza Shoaibi, Matthew Spotnitz, Marc A Suchard, Carmen O Torre, David Vizcaya, Haini Wen, Marcel de Wilde, Junqing Xie, Seng Chan You, Lin Zhang, Oleg Zhuk, Patrick Ryan, Daniel Prieto-Alhambra, on behalf of the OHDSI-COVID-19 consortium

Summary

Background Hydroxychloroquine, a drug commonly used in the treatment of rheumatoid arthritis, has received much negative publicity for adverse events associated with its authorisation for emergency use to treat patients with COVID-19 pneumonia. We studied the safety of hydroxychloroquine, alone and in combination with azithromycin, to determine the risk associated with its use in routine care in the patients with rheumatoid arthritis.

The Lancet Rheumatol 2020
Published Online August 21, 2020 https://doi.org/10.1016/S2665-9913(20)30276-9

RHEUMATOLOGY

Original article

COVID-19 in patients with autoimmune diseases: characteristics and outcomes in a multinational network of cohorts across three countries

Eng Hooi Tan[1], Anthony G. Sena[2,3], Albert Prats-Uribe[4], Seng Chan You[4], Waheed-Ul-Rahman Ahmed[5,6], Kristin Kostka[7], Christian Reich[7], Scott L. DuVall[5,8], Kristine E. Lynch[5,9], Michael E. Matheny[10,11], Talita Duarte-Salles[12], Sergio Fernandez Bertolin[12], George Hripcsak[13,14], Karthik Natarajan[13,14], Thomas Falconer[13], Matthew Spotnitz[13], Anna Ostropolets[13], Clair Blacketer[2,3], Thamir M Alshammari[15], Heba Alghoul[16], Osaid Alser[17], Jennifer C. E. Lane[1], Dalia M. Dawoud[18], Karishma Shah[5], Yue Yang[19], Lin Zhang[20,21], Carlos Areia[22], Asieh Golozar[23,24], Martina Recalde[12,25], Paula Casajust[26], Jitendra Jonnagaddala[27], Vignesh Subbian[28], David Vizcaya[29], Lana Y. H. Lai[30], Fredrik Nyberg[31], Daniel R. Morales[32], Jose D. Posada[33], Nigam H. Shah[34], Mengchun Gong[35], Arani Vivekanantham[4], Aaron Abend[36], Evan P Minty[36], Marc Suchard[37], Peter Rijnbeek[3], Patrick B. Ryan[2,13] and Daniel Prieto-Alhambra[1]

Prepublication Release

PEDIATRICS

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

30-Day Outcomes of Children and Adolescents With COVID-19: An International Experience

Talita Duarte-Salles, David Vizcaya, Andrea Pistillo, Paula Casajust, Anthony G. Sena, Lana Yin Hui Lai, Albert Prats-Uribe, Waheed-Ul-Rahman Ahmed, Thamir M Alshammari, Heba Alghoul, Osaid Alser, Edward Burn, Seng Chan You, Carlos Areia, Clair Blacketer, Scott DuVall, Thomas Falconer, Sergio Fernandez-Bertolin, Stephen Fortin, Asieh Golozar, Mengchun Gong, Eng Hooi Tan, Vojtech Huser, Pablo Ivell, Daniel R. Morales, Fredrik Nyberg, Jose D. Posada, Martina Recalde, Elena Roel, Lisa M. Schilling, Nigam H. Shah, Karishma Shah, Marc A. Suchard, Lin Zhang, Ying Zhang, Andrew E. Williams, Christian G. Reich, George Hripcsak, Peter Rijnbeek, Patrick Ryan, Kristin Kostka, Daniel Prieto-Alhambra

CANCER EPIDEMIOLOGY, BIOMARKERS & PREVENTION

Characteristics and outcomes of over 300,000 COVID-19 individuals with history of cancer in the United States and Spain

Elena Roel, Andrea Pistillo, Martina Recalde, Anthony G Sena, Sergio Fernández-Bertolín, Maria Aragon, Dalia Dawoud, Waheed-Ul-Rahman Ahmed, Heba Alghoul, Osaid Alser, Thamir M. Alshammari, Carlos Areia, Clair Blacketer, William Carter, Paula Casajust, Aedin C. Culhane, Dalia Dawood, Frank DeFalco, Scott L DuVall, Thomas Falconer, Asieh Golozar, Mengchun Gong, Laura Heras, George Hripcsak, Eng Hooi Tan, Hokyun Jeon, Jitendra Jonnagaddala, Lana YH Lai, Kristine E Lynch, Michael E Matheny, Daniel R Morales, Karthik Natarajan, Fredrik Nyberg, Anna Ostropolets, Jose D Posada, Albert Prats-Uribe, Christian G Reich, Doris N Rivera, Lisa M Schilling, Isabella Scorpinaniam, Karishma Shah, Nigam H. Shah, Yong Shan, Matthew Spotnitz, Vignesh Subbian, Marc A Suchard, Annalisa Towne, Lin Zhang, Ying Zhang, Patrick B Ryan, Daniel Prieto-Alhambra, Kristin Kostka and Talita Duarte-Salles

thebmj

RESEARCH

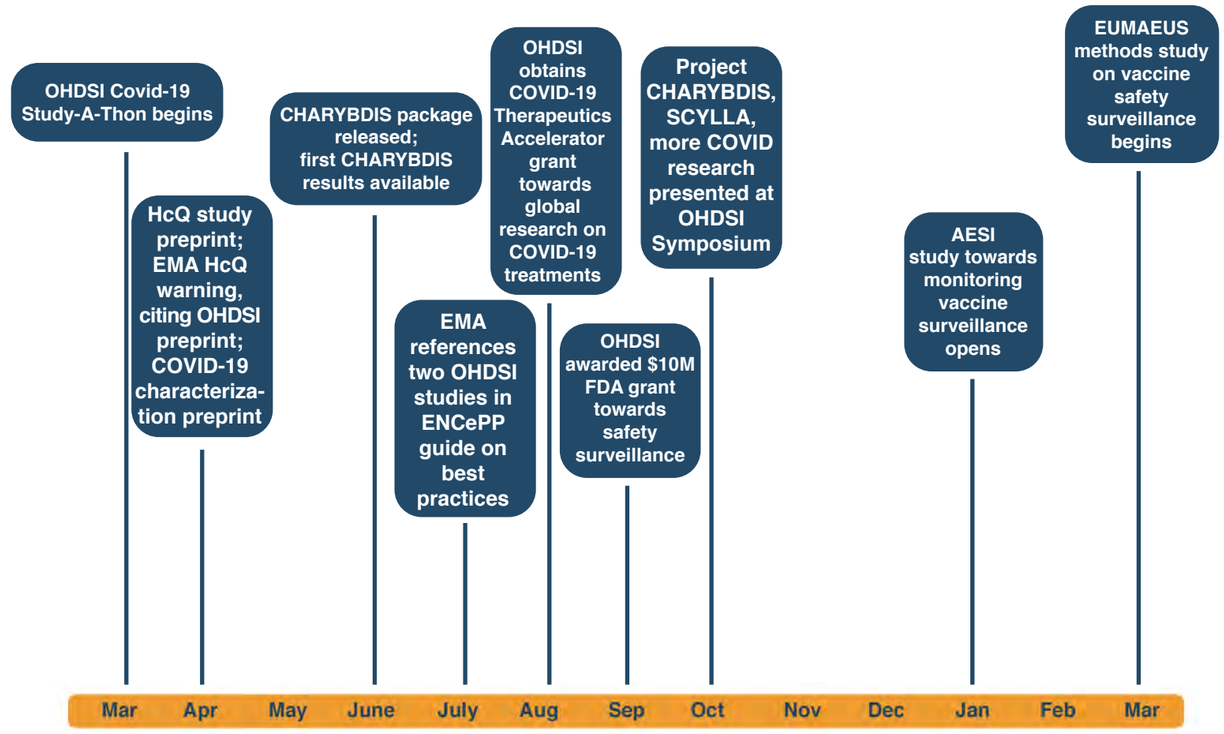
Use of repurposed and adjuvant drugs in hospital patients with covid-19: multinational network cohort study

Albert Prats-Uribe,[1] Anthony G Sena,[2,3] Lana Yin Hui Lai,[4] Waheed-Ul-Rahman Ahmed,[5,6] Heba Alghoul,[7] Osaid Alser,[8] Thamir M Alshammari,[9] Carlos Areia,[10] William Carter,[11] Paula Casajust,[12] Dalia Dawoud,[13,14] Asieh Golozar,[15,16] Jitendra Jonnagaddala,[17] Paras P Mehta,[18] Mengchun Gong,[19] Daniel R Morales,[20,21] Fredrik Nyberg,[22] Jose D Posada,[23] Martina Recalde,[24,25] Elena Roel,[24,25] Karishma Shah,[5] Nigam H Shah,[23] Lisa M Schilling,[11] Vignesh Subbian,[26] David Vizcaya,[27] Lin Zhang,[28,29] Ying Zhang,[19] Hong Zhu,[30] Li Liu,[30] Jaehyeong Cho,[31] Kristine E Lynch,[32] Michael E Matheny,[31,26] Seng Chan You,[33] Peter R Rijnbeek,[3] George Hripcsak,[34] Jennifer CE Lane,[3] Edward Burn,[1,24] Christian Reich,[37] Marc A Suchard,[35] Talita Duarte-Salles,[24] Kristin Kostka,[37,38] Patrick B Ryan,[2,40] Daniel Prieto-Alhambra[1]

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

Xintong Li,[1] Anna Ostropolets,[2] Rupa Makadia,[3] Azza Shoaibi,[3] Gowtham Rao,[3] Anthony G Sena,[3,6] Eugenia Martinez-Hernandez,[4] Antonella Delmestri,[1] Katia Verhamme,[6,7] Peter R Rijnbeek,[6] Talita Duarte-Salles,[5] Marc A Suchard,[8,9] Patrick B Ryan,[2,3] George Hripcsak,[2] Daniel Prieto-Alhambra[1,6]

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.

88 Hours: OHDSI's Signature Moment

OHDSI's COVID-19 work began with the ultimate show of collaboration & community.

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 late March of 2020. The symposium 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 an observational data science network), the needs were immediate, and everybody was staying home.

Those factors would be a hard stop to most, but the virtual OHDSI community has thrived on overcoming obstacles, and there was never a more crucial time to do so again.

Daniel Prieto-Alhambra, Professor of Pharmaco- and Device Epidemiology at Oxford, remembers his OHDSI conversion occurring during one of the afore-mentioned study-a-thon events in 2018. This one had nothing to do with viruses or antibodies; it was about the safety profile of varied knee replacement procedures and ultimately produced a paper published in Lancet Rheumatology.

While that data didn't affect COVID studies 15 months later, the impact of the event stayed with Prieto-Alhambra. 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 and presentations is available at the OHDSI COVID-19 research page). 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 hosted the 2019 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 teams that held varied roles throughout the event.

This setup, for example, enabled a group focused on phenotype development to work collaboratively, while also having the ability to connect with teams inside the characterization, estimation and prediction groups as well. When needed, there were support teams for literature review, data support, study design and more.

Your standard study-a-thon might just send various groups to different areas within a shared space. During these 88 hours, that 'space' might have had collaborators from both

More than 300 people from across 30 countries joined a critical journey during a 4-day study-a-thon in March, 2020, which set the foundation for OHDSI's work around COVID-19. It was the ultimate sign of collaboration through open science.

hemispheres working simultaneously at different points of a 13-hour time period. From breakfast in one part of the world to dinner in another, determined volunteers didn't stop working together to seek answers during a global crisis.

"OHDSI has always been about people working together to solve common goals, and I am proud our team helped to make this event possible," Rijnbeek said. "We brought the OHDSI world to Erasmus MC in person last year, but it was even more important to bring them together virtually right now."

88 hours.

It is unrealistic to think OHDSI's monumental goals could be accomplished in such a limited time. Early work needed to be done to develop an infrastructure for both the meetings and the OHDSI technical platforms, which happened mainly due to the sustained efforts of Lee Evans, Anthony Sena, and James Wiggins. Beyond that, many of the prioritized questions that would become the primary focus of the four days were determined beforehand.

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 national governments, public health agencies, and health-related institutions to learn what the most critical questions were right now. That feedback, as well as a literature review process that began days before the study-a-thon, helped the core team provide a framework for the four days.

There was a clear desire to create a multi-nation characterization study of COVID-positive patients, even if the data size was more limited at the moment.

There was a need to understand the overall safety profile of different drugs being considered in COVID treatment; that included hydroxychloroquine, which became an international fascination after achieving small success in France and then being touted by U.S. President Donald Trump on multiple occasions.

There were crucial prediction questions, which could help healthcare workers make important triage decisions, including which patients would require hospitalization. As each day passed, the challenges facing overwhelmed medical facilities globally were becoming abundantly clear.

Preliminary work with data was necessary as well. Christian Reich led the vocabulary team to develop COVID-related updates on the standardized vocabularies, while Kristin Kostka and Greg Klebanov were among many collaborators working with different sites on either data conversion or analysis support. Seng Chan You and Rae Woong Park collaborated with the South Korean HIRA, 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-

continued on next page

continued from previous page

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 characterization, 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 health-care agencies provided critical efforts, ranging from protocol design to data support.

"The OHDSI community has an open approach to every-thing," 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.

Covid-19 Study-A-Thon Registrants Span The Globe

Argentina	England	Saudi Arabia
Australia	France	Singapore
Belarus	Germany	South Korea
Belgium	Hungary	Spain
Brazil	India	Sweden
Canada	Israel	Switzerland
China	Italy	Taiwan
Colombia	Netherlands	Ukraine
Croatia	New Zealand	UAE
Denmark	Peru	United States

What You Should Know About The 2020 OHDSI COVID-19 Study-A-Thon

- **More than 330 people from across 30 countries (six continents) registered for the event.**
- **The event took place over 88 hours between March 26-29, and it was coordinated by the Erasmus University Medical Center.**
- **There were 17 concurrent channels on the overall Teams platform, and those channels hosted more thann 100 collaborator calls.**
- **There were 12 global huddles, spaced out so collaborators from around the world would have a daily opportunity to hear about community progress.**
- **More than 10,000 publications were reviewed both prior and during the event.**
- **There were 13,000+ chat messages that helped design both 355 cohort definitions and nine protocols, as well as the release of 13 study packages.**
- **The closing call has been viewed almost 1,800 times since it was posted to YouTune.**
- **As of August 2021, the OHDSI community has published 10 COVID-19 studies (including in Lancet Rheumatology, Nature Communications, Lancet Digital Health, and The BMJ), and at least nine others that are currently on a preprint server.**

Each person who takes that step strengthens the community.

88 hours.

You've seen that number before? OK, here are a few new ones.

Between the 12 global huddles, there were more than 100 collaborator calls and 13,000 chat messages over 17 concurrent channels (different teams). More than 10,000 publications were reviewed and 355 cohort definitions were defined to lead to the drafting of nine protocols and the release of 13 study packages.

"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. "Reflect-ing 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 good people that you 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. The 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.

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 for peer review two weeks after the final signoff, and more followed.

Generating real-world evidence to improve healthcare has been the OHDSI mission since it officially formed in 2014. This has been a passion project for a global community that expands in both people and analytic capability each year.

Nobody saw this moment coming. But it did, and OHDSI was more ready for it than even the most optimistic collaborator could have imagined.

There were critical discoveries in the first six years, and there are many more to come — including some that will aid global efforts against COVID-19 in the near future.

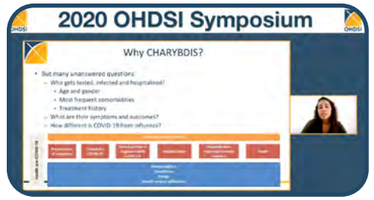
But those 88 hours stand as a defining moment for OHDSI, and they are a glimpse of this community's potential on the journey ahead.

by Craig Sachson
published April 17, 2020

#OHDSI2020 Global Symposium

Less than seven months after our COVID Study-A-Thon, our community shared incredible network research on the pandemic during the annual OHDSI Symposium. Community efforts were not limited to COVID alone, but like the rest of the world in 2020, the pandemic was a major focus. Check out some of the work that was shared, and read more about it in this book, or on OHDSI.org.

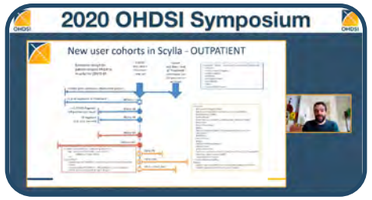
Plenary Session



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

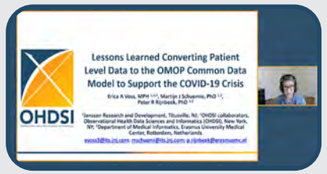
Project SCYLLA (SARS-Cov-2 Large-scale Longitudinal Analyses) focused on assessing comparative effectiveness and safety among treatments administered during hospitalization and prior to intensive services.

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 the COVID-19 pandemic in one way or another. Some of those presentations developed into peer-reviewed studies or are currently under review, while others highlights methods learned in 2020 that can provide road maps for future research. We commend the community for another amazing year of work at a time we needed it most.

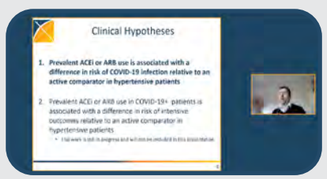
Learn more at ohdsi.org/2020-ohdsi-global-symposium

Lightning Talks



Lessons 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)

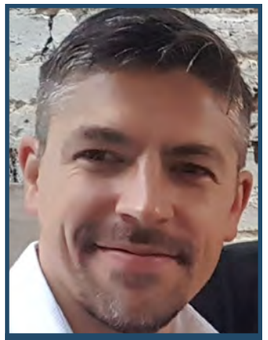


Renin-Angiotensin System Blockers and Susceptibility to COVID-19: a Multi-national Open Science Cohort Study
(Presenter: Daniel Morales)

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

OHDSI Practices Cited In Revision of EMA's ENCePP Guide on Methodological Standards

Jenny Lane



Daniel Morales

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."


EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

EMA/95098/2010 Rev.8

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

(Revision 8)

Both the 2020 and 2021 revisions of the EMA's ENCePP Guide on Methodological Standards in Pharmacoepidemiology cited two OHDSI COVID-19 studies.

The Rise and Fall of Hydroxychloroquine

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 hydroxychloroquine.

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-Uribe.

Largest Global Study on Hydroxychloroquine Safety Finds Increased Cardiovascular Risk with Azithromycin

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 published in Lancet Rheumatology.

In patients with rheumatoid arthritis, HCQ treatment in the short term (30 days) was found to not carry an excess risk of complications associated with its use, but HCQ treatment in the long term had a 65% relative increase in cardiovascular-related mortality, compared to sulfasalazine.

HCQ + AZM had a cardiovascular mortality risk that was more than twice (2.19) as high as the comparative treatment even in the short term based on findings from more than 320,000 users of that combination therapy. This treatment also produced a 15-20% increased rate of angina/chest 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 (ENCePP) 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 16 severe adverse events as compared to a similar RA drug, sulfasalazine (SSZ). Long-term HCQ therapy was associated with a 65% increase in cardiovascular mortality as compared to SSZ.

“Hydroxychloroquine, both alone and in combination with azithromycin, gained strong consideration as a potential COVID treatment without a large-scale study of its overall safety profile,” said Daniel Prieto-Alhambra, MD, MSc, PhD, co-senior author on this study. “We had access to an unprecedented amount of data on this drug, and we were relieved to find no worrying side effects in the short-term use of hydroxychloroquine. However, when prescribed in combination with azithromycin, it may induce

heart failure and cardiovascular mortality and we would urge caution in using the two together.”

This study examined more than 950,000 HCQ users through deidentified electronic health records and administrative claims data over a 20-year period. Records were collected from 14 different databases spanning six nations (Germany, Japan, Netherlands, Spain, United Kingdom, United States) and then mapped to the OMOP Common Data Model to generate this large-scale analysis.

“At medical school we were taught to ‘first do no harm’ and to me, our study focuses on this core belief of modern medicine,” said Jennifer Lane, MD, who served as co-lead author on this study along with Jamie Weaver, MPH, MS. “OHDSI has the power to investigate this question in a very thorough way and to go through rigorous steps. We are looking at patients from the general population, which is why it is so important to look at data from multiple countries. There are reasons why you may get bias from one data source, but if we find a signal in the Netherlands, and we find it in Spain, and we find it in the U.S., then we know we have something.”

“It required a global effort to generate this level of reproducible, reliable real-world evidence to inform decision-making around COVID treatment,” said Patrick Ryan, PhD, co-senior author on this study. “Our community collaborated for years to develop the high-level analytics which set the course for these studies. Standardizing data for nearly 1,000,000 patients on hydroxychloroquine provides confidence in these findings, and we are pleased to see that this study helped make a positive clinical impact as treatment options continue to be evaluated.”

Insufficient Data, Misleading Recommendations Led To Significant Early Heterogeneity In Global COVID-19 Patient Management

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 proclaimed 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 303,000 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.

“At 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-Uribe, a DPhil candidate and Research Assistant in Clinical Epidemiology at the University of Oxford.

“This was also influenced by political and social media pressures that spread misinformation,” said senior author Dani Prieto-Alhambra, Professor of Pharmaco- and Device Epidemiology at the University of Oxford. “Once reliable evidence from well-designed and performed studies came in, the situation quickly improved, and hospitals stopped using the ineffective treatments and turned to more effective ones.”

Deidentified patient data from 11 databases across three continents (Asia, Europe and North America) showed that more than 3,400 different medicines were used in the treatment of COVID-19 patients. Among the most popular in the earliest stages of the pandemic was hydroxychloroquine, which was heavily promoted without the backing of reliable evidence and later revoked from emergency approval status following both randomized controlled trials (RCTs) and related studies, including an OHDSI study showing dangerous risk of combining hydroxychloroquine with another early popular prescribed COVID-19 therapeutic, azithromycin.

Heterogeneity in drug therapy choice was dramatic across databases around the world. For example, lopinavir-ritonavir was used 50% of the time in one Spanish setting (HM Hospitals), 35% of the time in a South Korean setting (HIRA), and 0% of the time in a U.S. setting (Department of Veterans Affairs).

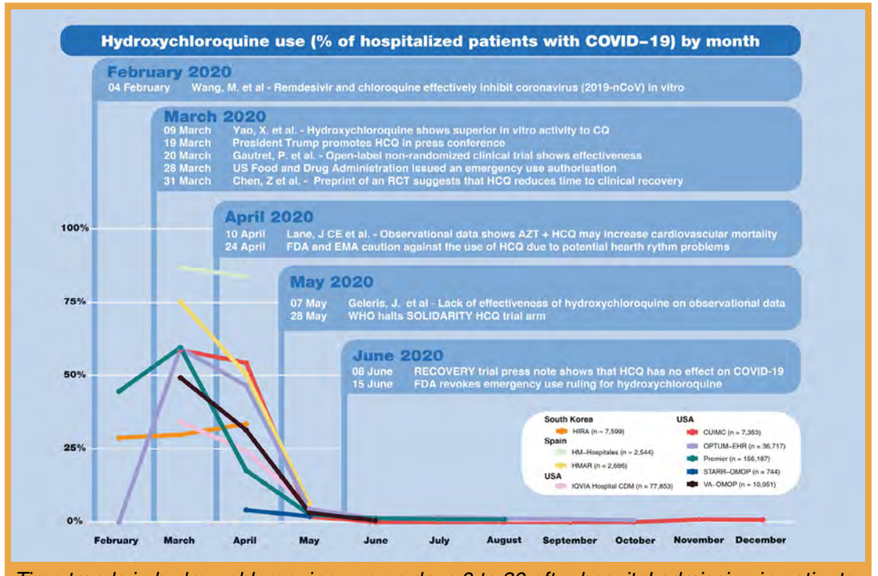
Albert Prats-Uribe

Adjunctive therapies developed into popular forms of management for supportive care, with the most recognized being corticosteroids and anti-cytokines, both of which have been shown to reduce mortality in more serious cases.

While these were lightly used early in the pandemic, results from the RECOVERY RCT showed efficacy in reducing death on hospitalized patients with severe respiratory disease.

Scientific discovery through observational data often reversed false information being distributed through political channels and/or social media. This study highlights the role observational studies can fit into informing clinical decision-making moving forward.

“The use of ineffective medicines and potentially harmful combinations started with information from promising in vitro analyses, and were fueled by poorly performed observational studies, as well as misinformation campaigns in social and traditional media with clearly political intentions,” Prats-Uribe said. “This would have taken a long time to counter in the traditional scientific timings. 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.”



Time trends in hydroxychloroquine use on days 0 to 30 after hospital admission in patients with a positive test result for or diagnosis of covid-19 by month. ***(Graphic courtesy of Albert Prats-Uribe)***

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.

You can't build a house without a foundation. Of course, you also don't live on the foundation either. At that point, the work had only begun.

"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, symptoms and outcomes of interest among individuals with COVID-19 overall and stratified by sex, age and specific comorbidities

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

Building on a study led by Ed Burn and published by Nature Communications entitled "Deep phenotyping of 34,128 adult patients hospitalized with COVID-19 in an international network study," which generated real-world evidence on the natural disease history of COVID-19, CHARYBDIS co-leads Kostka, Talita Duarte-Salles, and Albert Prats-Uribe led a community-wide effort to investigate deeper.

"Ed's characterization study is the foundation of how we got to this spot," Kostka said. "We quickly figured out we needed an over-arching frame to put everything into, so a lot of ideas that 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 DuVall and Duarte-Salles provided critical leadership with their work around the VA and SIDIAP 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 4.53 million patients diagnosed or tested positive for COVID-19
– more than 886,000 patients hospitalized with COVID-19

This level of work takes time. Let's take one for example. The study "Characteristics and outcomes of 627,044 COVID-19 patients with and without obesity in the United States, Spain, and the United Kingdom," led by Elena Roel, Martina Recalde, and Duarte-Salles, 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.

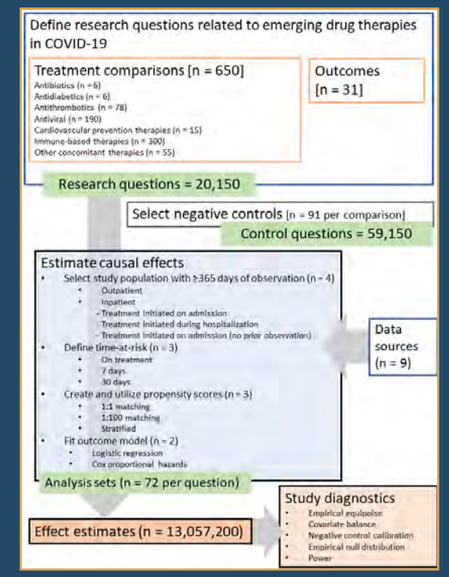
The SCYLLA Study

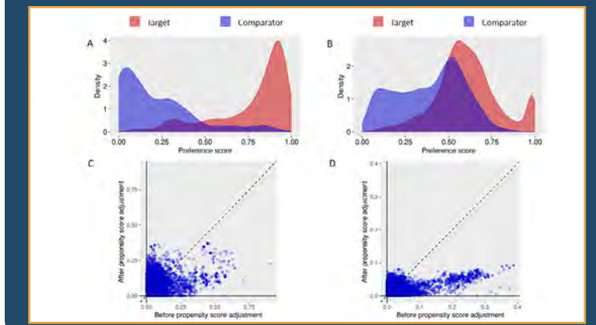
While Project CHARYBDIS studied the natural disease history of COVID-19, the OHDSI community recognized the need for real-world evidence around the different treatments being used around the world. In a world before vaccines, understanding both the safety and effectiveness of these treatments was of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

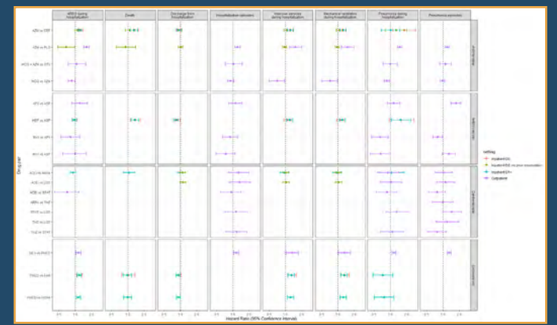
The SCYLLA (SARS-Cov-2 Large-scale Longitudinal Analyses) Study set out to do that work.

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.

The team shared several graphics to provide greater perspective of this study.







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.

"How do we differentiate what we are doing so people know this is one of the biggest things you can tap into, even if other people publish a paper first?" Kostka said. "We have more diversity in terms of geography, we have larger sample sizes, we've done more curation of reliability of the information. 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 police makers. 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.

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.

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 (AESI) being monitored in vaccine surveillance efforts identified that these rates vary substantially by age, sex, and database.

Outcomes by sex	1-5 years	6-17 years	18-34 years	35-54 years	55-64 years	65-74 years	75-84 years	≥85 years
Non-haemorrhagic stroke								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Acute myocardial infarction								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Deep vein thrombosis								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Haemorrhagic stroke								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Pulmonary embolism								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Appendicitis								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Bell's palsy								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Anaphylaxis								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Immune thrombocytopenia								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Myocarditis or pericarditis								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Disseminated intravascular coagulation								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Encephalomyelitis								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Narcolepsy								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Guillain-Barré syndrome								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Transverse myelitis								
Female	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Male	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Incidence rate per 100 000 person years (95% prediction interval)

Council of International Organizations of Medical Sciences frequency classification

Led by researchers at Oxford University, Columbia University, Erasmus MC, UCLA, and Janssen, an international team of collaborators within the OHDSI network provided a timely reference of the background rates of AESIs in the study "Characterising the background incidence rates of adverse events of special interest for covid-19 vaccines in eight countries: multinational network cohort study" published June 14 by The BMJ.

There were significant differences in the observed rates of AESIs based on the age groups and sex of more than 126 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 Dani Prieto-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 benchmark 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.

"We found significant heterogeneity in background rates between age and sex," said co-lead author Xintong Li, DPhil 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."

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 populations who are more likely to suffer these AESIs (like the older man in this example) 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 in the Columbia University Department of Biomedical Informatics.

The variability between different databases could reflect numerous factors, ranging from the process of data capture to population differences such as socioeconomic status and comorbidities.

"The observed heterogeneity between databases was more than I expected," Prieto-Alhambra said. "As a consequence, vaccine safety surveillance should be conducted using the same database for both post-vaccine and background rates."

X. Join The Journey

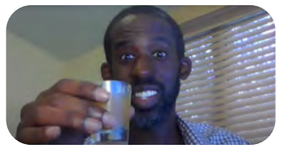
Artwork by OHDSI collaborator Sarah Seager

Cheers, From The OHDSI Community!

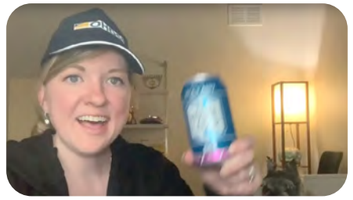
2020 threatened to pull people apart, but the OHDSI community came closer together. Volunteer researchers from around the globe joined forces to study COVID-19 and other critical healthcare concerns. Collaboration in the spirit of open science drove us to do far more together than anybody could have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



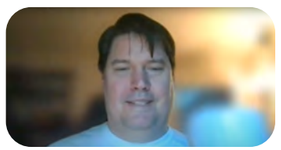
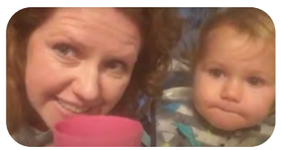


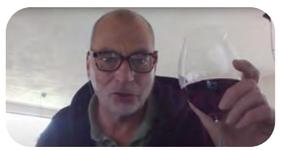






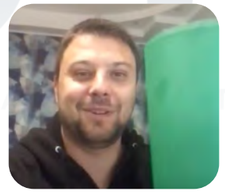
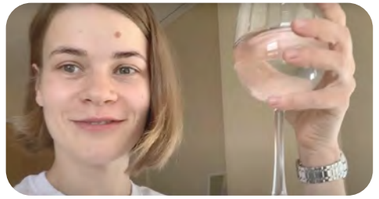






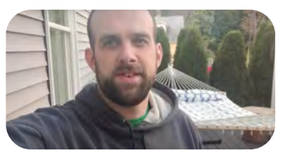









A favorite part of every OHDSI Symposium is the closing talk, given by Patrick Ryan. Naturally, we figured the appropriate way to end this OHDSI report was a closing letter from Patrick. Read along and check out some memorable images from past closings — ranging from Dr. Seuss to Hamilton, and baseball cards to cake.

Join The Journey

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 health-care 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 in 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.

JOIN THE JOURNEY

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.

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!

OHDSI

Join The Journey

@OHDSI OHDSI.org OHDSI