April Olympians: What We Learned & How We Can Use It

OHDSI Community Call
April 30, 2024 • 11 am ET

Upcoming Community Calls

Date	Topic
April 30	April Olympians Update | Presentation: What We Achieved & How You Can Use It
May 7	DevCon 2024 Review
May 14	10-Minute Tutorials
May 21	Open Studies in the OHDSI Community
May 28	Collaborator Showcase Brainstorm
June 4	NO CALL – EUROPEAN SYMPOSIUM
June 11	European Symposium Review
June 18	Application of LLMs In Evidence Generation Process
June 25	Recent OHDSI Publications

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

OHDSI Shoutouts!

Congratulations to the team of **Martin Baumgartner, Karl Kreiner, Aaron Lauschensky, Bernhard Jammerbund, Klaus Donsa, Dieter Hayn, Fabian Wiesmüller, Lea Demelius, Robert Modre-Osprian, Sabrina Neururer, Gerald Slamanig, Sarah Prantl, Luca Brunelli, Bernhard Pfeifer, Gerhard Pölzl, and Günter Schreier** on the publication of **Health data space nodes for privacy-preserving linkage of medical data to support collaborative secondary analyses** in *Frontiers in Medicine.*

frontiers | Frontiers in Medicine

TYPE Original Research
PUBLISHED 10 April 2024
DOI 10.3389/fmed.2024.1301660

Check for updates

OPEN ACCESS

EDITED BY
Gokce Banu Laleci Erturkmen,
Software Research and Development Consulting, Türkiye

REVIEWED BY
Sascha Welten,
RWTH Aachen University, Germany
Stefano Abbate,
University of Naples Federico II, Italy

*CORRESPONDENCE
Martin Baumgartner
martin.baumgartner@ait.ac.at

RECEIVED 25 September 2023
ACCEPTED 21 March 2024
PUBLISHED 10 April 2024

CITATION
Baumgartner M, Kreiner K, Lauschensky A, Jammerbund B, Donsa K, Hayn D, Wiesmüller F, Demelius L, Modre-Osprian R, Neururer S, Slamanig G, Prantl S, Brunelli L, Pfeifer B, Pölzl G and Schreier G (2024) Health data space nodes for privacy-preserving linkage of medical data to support collaborative secondary analyses. *Front. Med.* 11:1301660. doi: 10.3389/fmed.2024.1301660

COPYRIGHT
© 2024 Baumgartner, Kreiner, Lauschensky, Jammerbund, Donsa, Hayn, Wiesmüller, Demelius, Modre-Osprian, Neururer, Slamanig, Prantl, Brunelli, Pfeifer, Pölzl and Schreier. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Health data space nodes for privacy-preserving linkage of medical data to support collaborative secondary analyses

Martin Baumgartner[1,2*], Karl Kreiner[1], Aaron Lauschensky[1], Bernhard Jammerbund[1], Klaus Donsa[1], Dieter Hayn[1,3], Fabian Wiesmüller[1,2,3], Lea Demelius[4,5], Robert Modre-Osprian[6], Sabrina Neururer[7,8], Gerald Slamanig[9], Sarah Prantl[9], Luca Brunelli[10], Bernhard Pfeifer[8,11], Gerhard Pölzl[10] and Günter Schreier[1,2]

[1]Center for Health and Bioresources, AIT Austrian Institute of Technology, Vienna, Austria, [2]Institute of Neural Engineering, Graz University of Technology, Graz, Austria, [3]Ludwig Boltzmann Institute for Digital Health and Prevention, Salzburg, Austria, [4]Institute of Interactive Systems and Data Science, Graz University of Technology, Graz, Austria, [5]Know-Center GmbH, Graz, Austria, [6]telbiomed Medizintechnik und IT Service GmbH, Graz, Austria, [7]Department of Clinical Epidemiology, Tyrolean Federal Institute for Integrated Care, Tirol Kliniken GmbH, Innsbruck, Austria, [8]Division for Digital Health and Telemedicine, UMIT TIROL—Private University for Health Sciences and Technology, Hall in Tyrol, Austria, [9]Tirol Kliniken GmbH, Innsbruck, Austria, [10]Department of Internal Medicine III, Cardiology and Angiology, Medical University of Innsbruck, Innsbruck, Austria, [11]Tyrolean Federal Institute for Integrated Care, Tirol Kliniken GmbH, Innsbruck, Austria

Introduction: The potential for secondary use of health data to improve healthcare is currently not fully exploited. Health data is largely kept in isolated data silos and key infrastructure to aggregate these silos into standardized bodies of knowledge is underdeveloped. We describe the development, implementation, and evaluation of a federated infrastructure to facilitate versatile secondary use of health data based on Health Data Space nodes.

Materials and methods: Our proposed nodes are self-contained units that digest data through an extract-transform-load framework that pseudonymizes and links data with privacy-preserving record linkage and harmonizes into a common data model (OMOP CDM). To support collaborative analyses a multi-level feature store is also implemented. A feasibility experiment was

OHDSI Shoutouts!

Congratulations to the team of **Chungsoo Kim, Dong Han Yu, Hyeran Baek, Jaehyeong Cho, Seng Chan You, and Rae Woong Park** on the publication of **Data Resource Profile: Health Insurance Review and Assessment Service Covid-19 Observational Medical Outcomes Partnership (HIRA Covid-19 OMOP) database in South Korea** in the *International Journal of Epidemiology.*

International Journal of Epidemiology

Issues | More Content | Submit | Purchase | Alerts | About | International Journal of

Volume 53, Issue 3
June 2024
(In Progress)

< Previous

JOURNAL ARTICLE

Data Resource Profile: Health Insurance Review and Assessment Service Covid-19 Observational Medical Outcomes Partnership (HIRA Covid-19 OMOP) database in South Korea Get access

Chungsoo Kim, Dong Han Yu, Hyeran Baek, Jaehyeong Cho, Seng Chan You, Rae Woong Park

International Journal of Epidemiology, Volume 53, Issue 3, June 2024, dyae062, https://doi.org/10.1093/ije/dyae062

Published: 24 April 2024 **Article history**

Cite | Permissions | Share

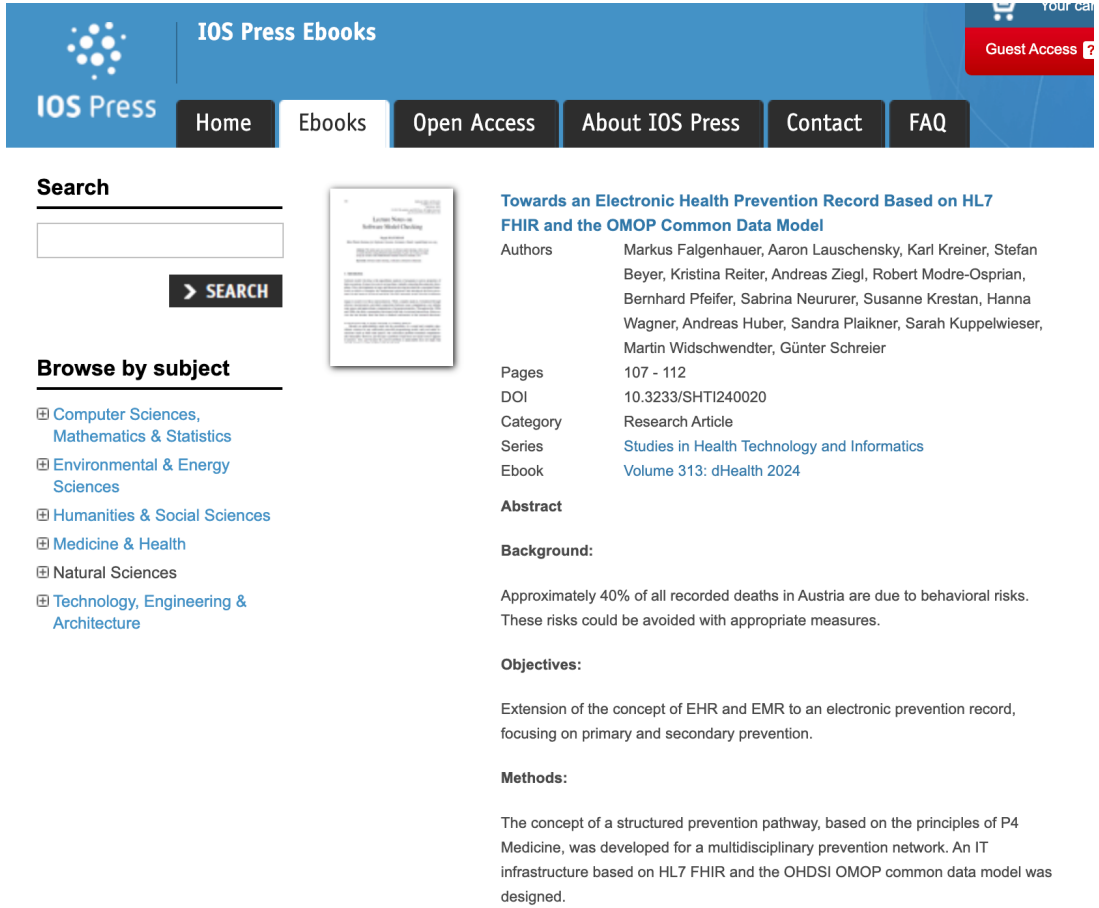
Extract

Data resource basics

The health insurance system in South Korea is known as the National Health Insurance (NHI) system. It is a mandatory insurance system that covers approximately 97% of Korean citizens and residents.[1] The NHI system covers most medical expenses, including prescription drugs, diagnostic tests and medical procedures. The Health Insurance Review and Assessment Service (HIRA) in South Korea plays a crucial role in the health care system by monitoring and assessing the quality and appropriateness of medical services provided to patients under the NHI system.[2] Its primary mission is to ensure that the medical services provided to patients are of high quality and cost-effective, by reviewing and assessing the appropriateness of medical services, evaluating the quality of care and monitoring healthcare use and costs.

OHDSI Shoutouts!

Congratulations to the team of **Markus Falgenhauer, Aaron Lauschensky, Karl Kreiner, Stefan Beyer, Kristina Reiter, Andreas Ziegl, Robert Modre-Osprian, Bernhard Pfeifer, Sabrina Neururer, Susanne Krestan, Hanna Wagner, Andreas Huber, Sandra Plaikner, Sarah Kuppelwieser, Martin Widschwendter, and Günter Schreier** on the publication of **Towards an Electronic Health Prevention Record Based on HL7 FHIR and the OMOP Common Data Model** in *Volume 313 of Studies in Health Technology and Informatics.*



IOS Press Ebooks

Guest Access

IOS Press | Home | Ebooks | Open Access | About IOS Press | Contact | FAQ

Search

SEARCH

Browse by subject

- Computer Sciences, Mathematics & Statistics
- Environmental & Energy Sciences
- Humanities & Social Sciences
- Medicine & Health
- Natural Sciences
- Technology, Engineering & Architecture

Towards an Electronic Health Prevention Record Based on HL7 FHIR and the OMOP Common Data Model

Authors: Markus Falgenhauer, Aaron Lauschensky, Karl Kreiner, Stefan Beyer, Kristina Reiter, Andreas Ziegl, Robert Modre-Osprian, Bernhard Pfeifer, Sabrina Neururer, Susanne Krestan, Hanna Wagner, Andreas Huber, Sandra Plaikner, Sarah Kuppelwieser, Martin Widschwendter, Günter Schreier

Pages: 107 - 112

DOI: 10.3233/SHTI240020

Category: Research Article

Series: Studies in Health Technology and Informatics

Ebook: Volume 313: dHealth 2024

Abstract

Background:

Approximately 40% of all recorded deaths in Austria are due to behavioral risks. These risks could be avoided with appropriate measures.

Objectives:

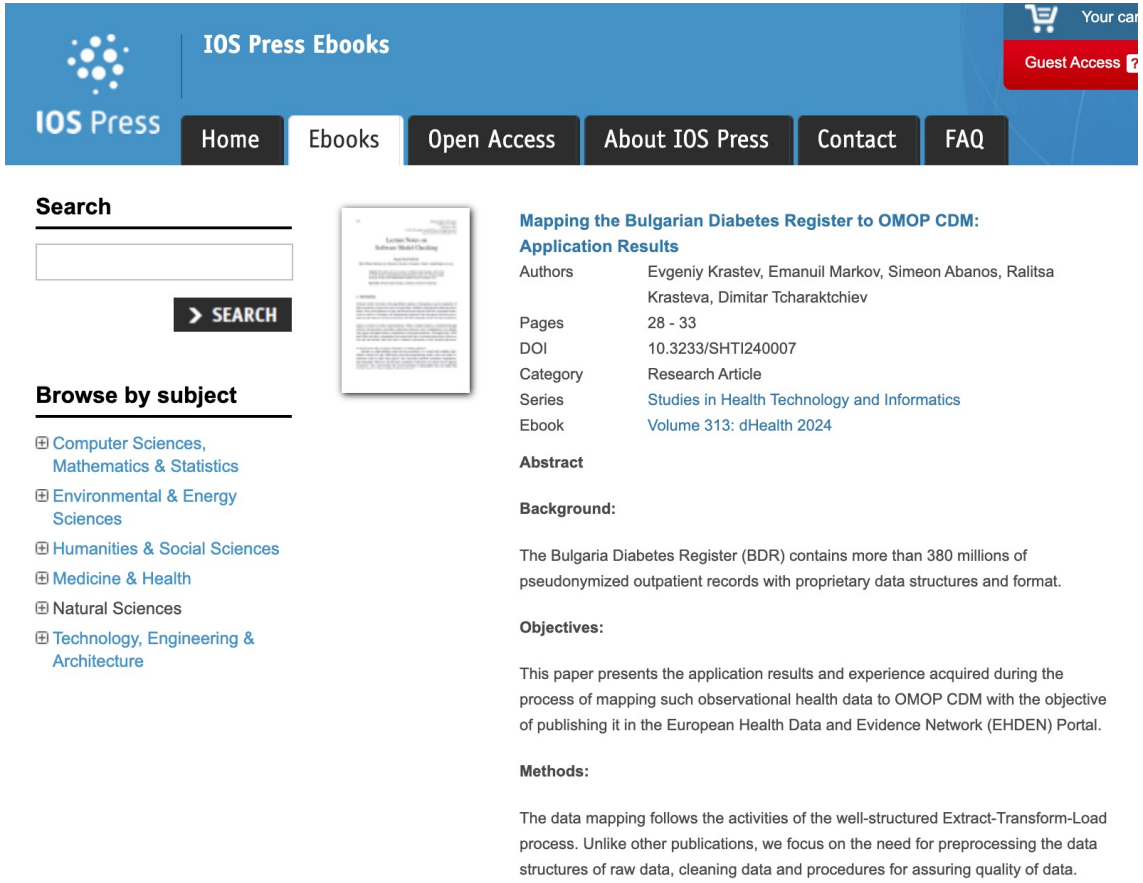
Extension of the concept of EHR and EMR to an electronic prevention record, focusing on primary and secondary prevention.

Methods:

The concept of a structured prevention pathway, based on the principles of P4 Medicine, was developed for a multidisciplinary prevention network. An IT infrastructure based on HL7 FHIR and the OHDSI OMOP common data model was designed.

OHDSI Shoutouts!

Congratulations to the team of **Evgeniy Krastev, Emanuil Markov, Simeon Abanos, Ralitsa Krasteva, and Dimitar Tcharaktchiev** on the publication of **Towards an Electronic Health Prevention Record Based on HL7 FHIR and the OMOP Common Data Model** in *Volume 313 of Studies in Health Technology and Informatics.*



Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

Upcoming Workgroup Calls

Date	Time (ET)	Meeting
Tuesday	12 pm	Common Data Model Vocabulary Subgroup
Wednesday	8 am	Psychiatry
Wednesday	3 pm	Joint Vulcan/OHDSI Meeting
Wednesday	7 pm	Medical Imaging
Thursday	9:30 am	Themis
Thursday	11 am	Industry
Thursday	1 pm	OMOP CDM Vocabulary/Development Subgroup
Thursday	7 pm	Dentistry
Friday	10 am	GIS-Geographic Information System
Friday	11:30 am	Steering Group
Monday	9 am	Vaccine Vocabulary
Monday	10 am	Healthcare Systems Interest Group
Tuesday	9 am	Atlas/WebAPI
Tuesday	10 am	Registry
Tuesday	10 am	Common Data Model

DevCon 2024 Videos Are Posted

Morning Agenda

9:00 am – Introduction (Adam Black, Paul Nagy)

9:15 am – Developers Panel and Lightning Talks (Katy Sadowski)

- *OHDSI/OMOP – The hard way is the easy way!* (Vishnu V Chandrabalan)
- *Moving OMOP to the Cloud With DBT and Snowflake* (Roger Carlson)
- *Use cases for ORMs in OMOP* (Georgina Kennnedy)
- *Carrot: code-free OMOP ETL without full data access* (Sam Cox)
- *Rabbit-in-a-blender - an ETL pipeline to transform your EMR data into OMOP* (Pieter-jan Lammertyn)

10:45 am – Darwin EU® Developers Update (Adam Black)

- *CDMConnector, PatientProfiles, CohortCharacteristics, CohortSurvival*

12:00 pm – Break

Afternoon Agenda

12:30 pm – OHDSI Ecosystem Updates

- TAB Update (Frank DeFalco)
- Strategus Update (Anthony Sena)
- Broadsea Update (Lee Evans)
- Kheiron Updates (Paul Nagy)

1:15 pm – JACKALOPE PLUS The Power of ML for Healthcare Data Mapping & Management (Denys Kaduk)

2:00 pm - An Introduction to Knowledge Graphs using PheKnowLator and OMOP2OBO with Example Applications in Drug Surveillance and Computational Phenotyping (Tiffany Callahan)

#OHDSI2024 Registration Is Open!

Registration is now OPEN for the 2024 OHDSI Global Symposium, which will be held Oct. 22-24 at the Hyatt Regency Hotel in New Brunswick, N.J., USA.

Tuesday: Tutorials
Wednesday: Plenary/Showcase
Thursday: Workgroup Activities

ohdsi.org/OHDSI2024

#OHDSI2024 Collaborator Showcase

Submissions are now being accepted for the 2024 Global Symposium Collaborator Showcase.

All submissions are due by 8 pm ET on Friday, June 21.

Notification of acceptance will be made by Tuesday, Aug. 20.

ohdsi.org/OHDSI2024

Maternal Health Data Science Fellowship

This program is designed to empower clinical investigators to leverage emerging technologies for improved maternal and neonatal care while reducing morbidity and mortality.

Three main components of this program:

1) Career Development (create evidence, leverage data models, build skills on network studies)

2) Practice (design effective observational research protocols, master tools, write papers/grants)

3) Networking (build relationships with mentors, learners, coordinate with global OHDSI collaborators)

Application deadline: May 15

Want to build your career?

Generate reproducible evidence by leading multi-institutional studies!

Learn more & apply!

RWE Workshop at AIME24: Call for Submissions!

Workshop: **AI for Reliable and Equitable Real-World Evidence Generation in Medicine**

https://medicine.utah.edu/dbmi/aime/ai-reliable

Organizing Committee
Linying Zhang
Adam Wilcox
Yves Lussier

Scientific Program Committee
Peter Rijnbeek
Larry Han
Xiaoqian Jiang
Mattia Prosperi
Xia Ning
Yifan Peng

Opening Keynote
George Hripcsak

IMPORTANT DATES

May 31, 2024 | Submission Deadline

June 14, 2024 | Notice of Acceptance

July 12, 2024 | Workshop

#OHDSISocialShowcase This Week

MONDAY

Opportunity and Challenge of Implementing the OHDSI System in Indonesia

(Dian Tri Wiyanti, Daniel C.A. Nugroho, Yudha Eri Saputra, Septi Melisa, Phan Thanh-Phuc, Nguyen Phung-Anh, Jason C. Hsu, Min-Huei Hsu)

Opportunity and Challenge of Implementing the OHDSI System in Indonesia

Dian Tri Wiyanti[1,2],Daniel C.A. Nugroho[3,4], Yudha Eri Saputra[1], Septi Melisa[1], Phan Thanh-Phuc[1],Nguyen Phung-Anh[5,6,7], Jason C.Hsu[1,8], Min-Huei Hsu[1,8,9,10,11]*

[1]International Ph.D. Program in Biotech and Healthcare Management, Taipei Medical University, Taiwan
[2]Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Indonesia
[3]Graduate Institute of Biomedical Informatics, Taipei Medical University
[4]Faculty of Medicine, Duta Wacana Christian University, Indonesia
[5]Clinical Data Center, Office of Data Science, Taipei Medical University, Taiwan
[6]Research Center of Health Care Industry Data Science, College of Management, Taipei Medical University, Taiwan
[7]Clinical Big Data Research Center, Taipei Medical University Hospital, Taipei Medical University, Taiwan
[8]Graduate Institute of Data Science, College of Management, Taipei Medical University, Taipei, Taiwan
[9]Professional Master Program in Artificial Intelligence in Medicine, Taipei Medical University, Taipei, Taiwan
[10]Taipei Medical University Shuang Ho Hospital, Taipei Medical University, Taipei, Taiwan
[11]TMU Research Center of Artificial Intelligence in Medicine and Health, Taipei Medical University, Taipei, Taiwan
*Corresponding author

Background

Indonesia, a densely populated nation with diverse ethnicities, possesses medical institutions like UKDW (Universitas Kristen Duta Wacana) that have well-established healthcare systems. These institutions publish research utilizing Electronic Health Records (EHR), making them valuable for the OHDSI (Observational Health Data Sciences and Informatics) community. This study aims to evaluate the potential for implementing OHDSI in Indonesia, with UNNES (Universitas Negeri Semarang) offering fresh insights. Despite challenges, the establishment of a new health faculty presents an opportunity to integrate OHDSI into the curriculum, foster research collaborations, and develop a strong data platform. Overcoming obstacles will necessitate careful planning, resource allocation, and active involvement of stakeholders. The study assesses healthcare preparedness, technical prerequisites, market landscape, and identifies obstacles and prospects for implementation.

Methods

Table 1. Variable Comparison across Indonesia BPJS, South Korea NHIS-NSC OMOP CDM Conversion, TMUCRD, and Indonesian Hospital Record Datasets

Indonesia BPJS	South Korea NHIS-NSC OMOP CDM Conversion	TMUCRD	Indonesian Hospital Record
NHI ID	PERSON	ID_No	NHI ID
-	DEATH	Death	DEATH
PHC VISIT Hospital Visit	VISIT	OPD/IPD FEE_No	Visit
PHC DIAGNOSIS Hospital Diagnosis	CONDITION	ICD9_Code/ICD10_Code	Diagnosis
Indonesian Case Based Group	DRUG	MED_Code	Drug
	PROCEDURE	OPD/IPD GROUP_Code/EXPER_No	Procedure
	DEVICE	-	-
	MEASUREMENT	EXPER_Code/R_Item	-
	OBSERVATION	-	-
	COST	TOT_AMT	-

Results

The dataset of the Indonesian Health Social Security Agency includes primary services (registration, outpatient procedures, and discharge procedures) and administrative tasks (planning, procurement, inventory maintenance, asset management, HR management, and financial management). It encompasses 111 variables. These variables align with the OMOP CDM categories (Person, Death, Visit, Condition, Drug, Procedure, Device, Measurement, Observation, and Cost) and serve as classifiers for the Indonesian BPJS dataset.

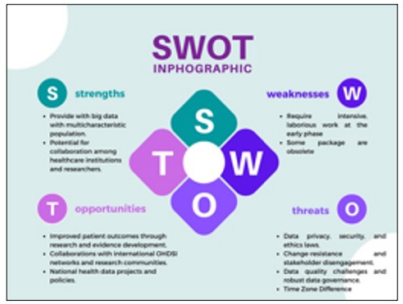


Figure 1. Analyzing the Opportunities and Challenges: Unleashing the Potential of OHDSI Implementation in Indonesia

Conclusions

Overall, OHDSI in Indonesia brings potential and problems. The report emphasises UKDW's infrastructure, knowledge, and stakeholder engagement in OHDSI acceptability. With UKDW's worldwide collaborations, UNNES and UKDW want to implement OHDSI. Evidence-based decision-making, innovation, and thorough awareness campaigns are effective implementation techniques. OHDSI implementation is simplified by aligning national health insurance data format with OMOP-CDM. Effective resource management and UNNES training may help adapt. For OHDSI integration, UKDW's IT and EHR competence is essential. Both UKDW and UNNES expansion activities help support OHDSI in Indonesia. Despite minimal efficacy testing, collaboration and improvement are stressed. Integration needs careful planning, resource allocation, stakeholder participation, and problem-solving. UKDW-UNNES collaboration improves OHDSI implementation in Indonesia.

References:

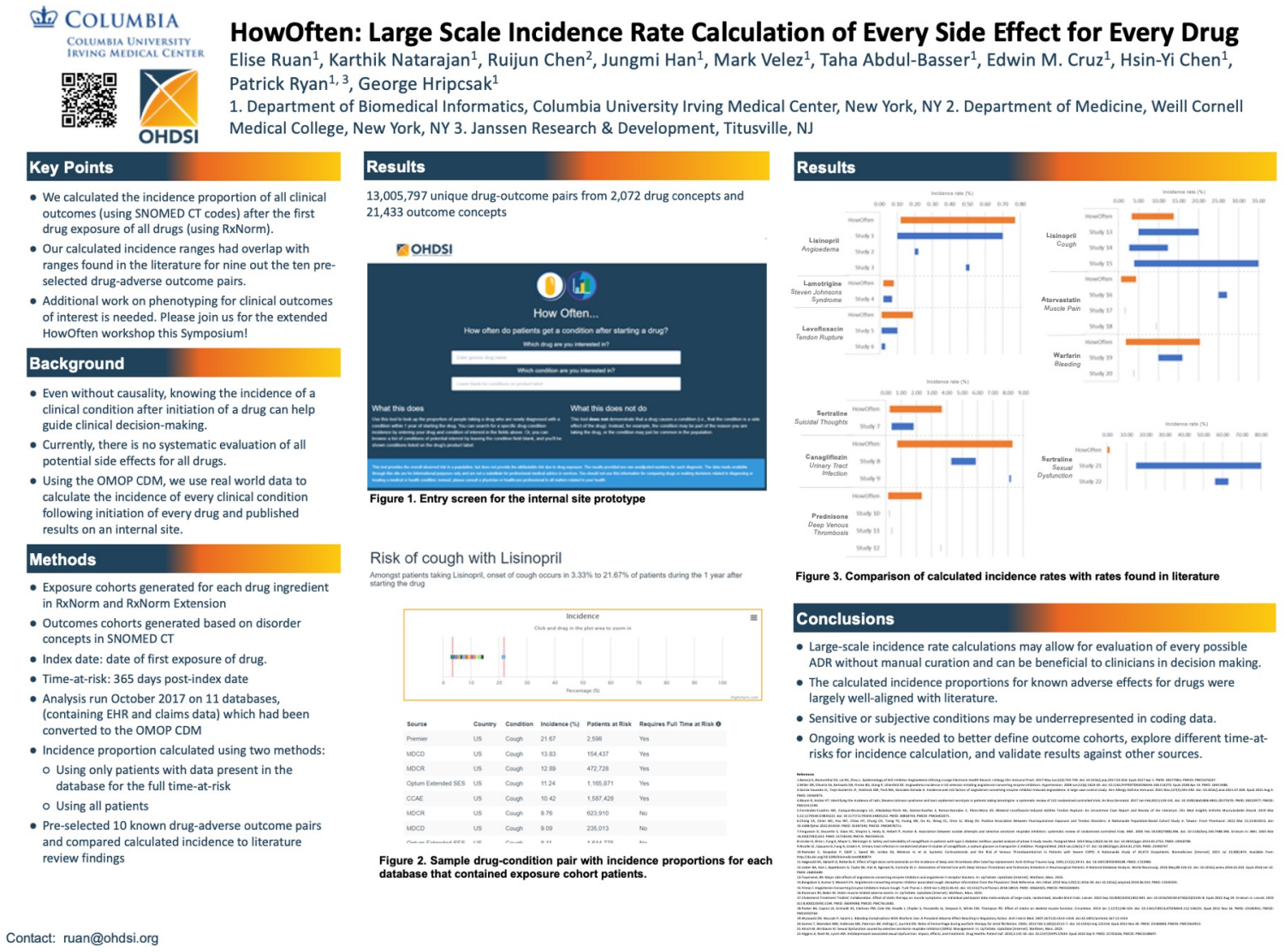
1. R. Padma and R. T. Pinzon, "A poor prognostic factor for clinical outcomes in ischemic stroke with dysphagia," J. Neurol. Sci., vol. 381, p. 643, Oct. 2017, doi: 10.1016/j.jns.2017.08.1810.
2. R. P. Gitasari, T. A. R. Kusumosih, D. C. A. Nugroho, and H. Rahardjo, "Factors Associated with Retained Placenta: A Case-Control at Bethesda Hospital Yogyakarta," Berk. Ilm. Kedokt. Duta Wacana, vol. 7, no. 1, p. 11, Aug. 2022, doi: 10.21460/bikdw.v7i1.356.
3. S. Lee et al., "OMOP-CDM Conversion and Anonymization of National Health Insurance Service-National Sample Cohort," in OHDSI Symposium, 2019, pp. 1–3.
4. S. Umaiyah, "UKDW Akan Bangun Kampus II di Bantul," Tribunjogja.com, 2019. https://jogja.tribunnews.com/2019/11/01/ukdw-akan-bangun-kampus-ii-di-bantul.
5. W. Pras, "Rencana Peruntukkan Lahan di Bantul Yang Telah Dapatkan Izin Prinsip," Rumah Jogja Indonesia.com, 2020. https://www.rumahjogjaindonesia.com/isi-majalah/rencana-peruntukkan-lahan-di-bantul-yang-telah-dapatkan-izin-prinsip.html.
6. UKDW, "Publication Ethics," BIKDW. https://bikdw.ukdw.ac.id/index.php/bikdw/pages/view/ethics.
7. G. Y. Sanjaya, A. Fuad, W. A. Baros, E. Dhanalvin, and N. Albab, "Database Riset Bersumber Data Jaminan Kesehatan Nasional untuk Mendukung Penelitian dan Publikasi Ilmiah di Indonesia," J. Inf. Syst. Public Heal., vol. 2, no. 2, pp. 18–29, 2017.
8. C. R. Yusuf, D. C. A. Nugroho, K. Adisaputro, and M. A. Sigilipoe, "Analysis of Technical Readiness and Telehealth Service Usage in Bantul District Health Centers," 2023.
9. I. A. Triastuti and D. C. A. Nugroho, "Evaluation of Multidisciplinary Learning: A Case Study from The Application of Medical Technology Course," 2022.

#OHDSISocialShowcase This Week

TUESDAY

HowOften: Large Scale Incidence Rate Calculation of Every Side Effect for Every Drug

(Elise Ruan, Karthik Natarajan, Ruijun Chen, Jungmi Han, Mark Velez, Taha Abdul-Basser, Edwin M. Cruz, Cindy Hsin-Yi Chen, Patrick Ryan, George Hripcsak)



COLUMBIA
COLUMBIA UNIVERSITY IRVING MEDICAL CENTER

OHDSI

HowOften: Large Scale Incidence Rate Calculation of Every Side Effect for Every Drug

Elise Ruan[1], Karthik Natarajan[1], Ruijun Chen[2], Jungmi Han[1], Mark Velez[1], Taha Abdul-Basser[1], Edwin M. Cruz[1], Hsin-Yi Chen[1], Patrick Ryan[1, 3], George Hripcsak[1]

1. Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, NY 2. Department of Medicine, Weill Cornell Medical College, New York, NY 3. Janssen Research & Development, Titusville, NJ

Key Points

- We calculated the incidence proportion of all clinical outcomes (using SNOMED CT codes) after the first drug exposure of all drugs (using RxNorm).
- Our calculated incidence ranges had overlap with ranges found in the literature for nine out the ten pre-selected drug-adverse outcome pairs.
- Additional work on phenotyping for clinical outcomes of interest is needed. Please join us for the extended HowOften workshop this Symposium!

Background

- Even without causality, knowing the incidence of a clinical condition after initiation of a drug can help guide clinical decision-making.
- Currently, there is no systematic evaluation of all potential side effects for all drugs.
- Using the OMOP CDM, we use real world data to calculate the incidence of every clinical condition following initiation of every drug and published results on an internal site.

Methods

- Exposure cohorts generated for each drug ingredient in RxNorm and RxNorm Extension
- Outcomes cohorts generated based on disorder concepts in SNOMED CT
- Index date: date of first exposure of drug.
- Time-at-risk: 365 days post-index date
- Analysis run October 2017 on 11 databases, (containing EHR and claims data) which had been converted to the OMOP CDM
- Incidence proportion calculated using two methods:
  - Using only patients with data present in the database for the full time-at-risk
  - Using all patients
- Pre-selected 10 known drug-adverse outcome pairs and compared calculated incidence to literature review findings

Results

13,005,797 unique drug-outcome pairs from 2,072 drug concepts and 21,433 outcome concepts

OHDSI

How Often...

How often do patients get a condition after starting a drug?

Which drug are you interested in?

Which condition are you interested in?

What this does

What this does not do

Figure 1. Entry screen for the internal site prototype

Risk of cough with Lisinopril

Amongst patients taking Lisinopril, onset of cough occurs in 3.33% to 21.67% of patients during the 1 year after starting the drug

Incidence

Click and drag in the plot area to zoom in

Percentage (%)

Source	Country	Condition	Incidence (%)	Patients at Risk	Requires Full Time at Risk
Premier	US	Cough	21.67	2,598	Yes
MDCD	US	Cough	13.83	154,437	Yes
MDCR	US	Cough	12.89	472,728	Yes
Optum Extended SES	US	Cough	11.24	1,165,871	Yes
CCAE	US	Cough	10.42	1,587,428	Yes
MDCR	US	Cough	9.76	623,910	No
MDCD	US	Cough	9.09	235,013	No

Figure 2. Sample drug-condition pair with incidence proportions for each database that contained exposure cohort patients.

Results

Lisinopril Angioedema
Lamotrigine Steven Johnsons Syndrome
Levofloxacin Tendon Rupture
Lisinopril Cough
Atorvastatin Muscle Pain
Warfarin Bleeding
Sertraline Suicidal Thoughts
Canagliflozin Urinary Tract Infection
Prednisone Deep Venous Thrombosis
Sertraline Sexual Dysfunction

Figure 3. Comparison of calculated incidence rates with rates found in literature

Conclusions

- Large-scale incidence rate calculations may allow for evaluation of every possible ADR without manual curation and can be beneficial to clinicians in decision making.
- The calculated incidence proportions for known adverse effects for drugs were largely well-aligned with literature.
- Sensitive or subjective conditions may be underrepresented in coding data.
- Ongoing work is needed to better define outcome cohorts, explore different time-at-risks for incidence calculation, and validate results against other sources.

Contact: ruan@ohdsi.org

#OHDSISocialShowcase This Week

WEDNESDAY

Agreement between measurement and diagnosis-based phenotype algorithms

(Azza Shoaibi, Gowtham Rao, Dmytro Dymshyts, Anna Ostropolets, Patrick Ryan)

Title: Agreement between measurement and diagnosis-based phenotype algorithms

PRESENTER: **Azza Shoaibi**

INTRO:

- Various types of clinical information, including diagnoses, medications, and procedures can be used to identify a specific clinical condition or event in observational data.
- Previous research indicates that the accuracy of phenotype algorithms can improve when multiple data types are incorporated.
- The aim of this paper is to compare diagnosis-based phenotype algorithms with those that are based on clinical measurements across five different clinical conditions in seven separate data sources.

METHODS:

- We selected five condition phenotypes: rhabdomyolysis, neutropenia, thrombocytopenia, pancytopenia, and end-stage renal disease
- We developed two different types of algorithms for all five conditions. Measurement-based phenotype algorithms used defined value thresholds for diagnostic markers. Diagnosis-based algorithms relied on the occurrence of at least one diagnosis.
- The cohorts were generated and evaluated across a range of data sources (listed in table 1); including 3 claims-based database, 1 general practitioner records and 2 health record (EHR) . ALL data sources contained measurement data from outpatient and/or inpatient encounters with at least partial coverage

The utility of measurements vary by clinical condition and by data source. Measurements should be empirically evaluated when developing phenotype algorithms.

	Thrombocytopenia					Neutropenia				
Database Name	**D Only**	**M Only**	**Both**	**overlap**	**N**	**D Only**	**M Only**	**Both**	**overlap**	**N**
CCAE	88.4%	8.7%	2.9%	25.0%	1,209,577	86.3%	11.4%	2.2%	16.2%	1,226,238
IQVIA® Ambulatory EMR	14.4%	78.3%	7.3%	8.5%	1,075,889	50.5%	37.7%	11.8%	23.8%	573,655
IQVIA® LPD in Australia	17.7%	79.2%	3.1%	3.8%	5,244	100.0%	0	0	NA	1,805
Medicare	88.6%	6.3%	5.2%	45.2%	501,880	94.3%	3.7%	2.0%	34.5%	239,815
Optum's DOD	50.8%	29.5%	19.7%	40.0%	2,388,152	61.6%	27.5%	10.9%	28.4%	1,431,307
Optum® EHR	4.1%	72.6%	23.4%	24.4%	6,361,599	98.4%	0.8%	0.7%	46.7%	828,438
Premier	74.8%	17.8%	7.3%	29.1%	7,891,028	85.0%	10.8%	4.2%	28.0%	2,094,195
	Pancytopenia					**End-stage renal disease**				
	D Only	**M Only**	**Both**	**overlap**	**N**	**D Only**	**M Only**	**Both**	**overlap**	**N**
CCAE	97.7%	0.9%	1.5%	65.2%	234,617	100.0%	0.0%	<0.0%	NA	250,770
IQVIA® Ambulatory EMR	47.4%	31.1%	21.6%	50.0%	83,100	97.6%	0.9%	1.5%	62.5%	253,075
IQVIA® LPD in Australia	100.0%	0.0%	0.0%		84	39.6%	57.4%	3.0%	5.0%	197
Medicare	97.2%	0.7%	2.1%	75.0%	139,462	100.0%	0.0%	0	NA	223,704
Optum's DOD	85.4%	5.2%	9.4%	64.4%	406,264	99.0%	0.1%	0.9%	90.0%	527,668
Optum® EHR	24.5%	18.5%	57.0%	75.5%	433,971	100.0%	0.0%	0.0%	NA	341,753
Premier	88.4%	3.4%	8.2%	70.7%	1,527,731	100.0%	0.0%	0.0%	NA	2,700,277
	Rhabdomyolysis					**D Only**: Proportion of patients identified by diagnosis-based phenotype only among total patients identified by either approach **M Only:** Proportion of patients identified by measurement-based phenotype only among total patients identified by either approach **Both:** Proportion of patients identified by both approaches among total patients identified by either approach **Overlap:** Proportion of patients identified in both approaches among those identified by the measurement-based algorithm				
	D Only	**M Only**	**Both**	**overlap**	**N**					
CCAE	NA	NA	NA	NA	NA					
IQVIA® Ambulatory EMR	NA	NA	NA	NA	NA					
IQVIA® LPD in Australia	NA	NA	NA	NA	NA					
Medicare	NA	NA	NA	NA	NA					
Optum's DOD	NA	NA	NA	NA	NA					
Optum® EHR	99.5%	0.2%	0.3%	60.0%	161,254					
Premier	100.0%	0.0%	0.0%	NA	879,951					

Table 1 -The overlap between the diagnosis based and measurement-based phenotype algorithms by phenotype in each data source.

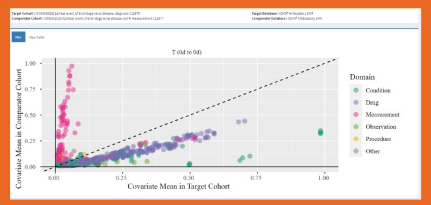


Figure 1: The covariate distribution among patients with thrombocytopenia identified through diagnosis compared to those identified through measurement on index date in IQVIA® LPD Australia

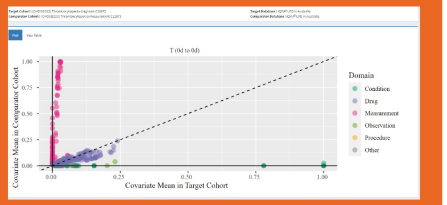


Figure 2: The covariate distribution among patients with End-stage renal disease identified through diagnosis compared to those identified through measurement on index date in IQVIA® Ambulatory EMR

Take a picture to view the abstract

METHODS, cont'd.

- For each condition we report on:
  1. the proportion of patients identified solely by each approach.
  2. the proportion identified by both approaches, and the overlap of identified patients.
- In addition, we conducted a comparison of covariate distributions among individuals who met each definition to evaluate the agreement in patient characteristics.
- We utilized the CohortDaignistic R package 3 to generate all results.

RESULTS

- A substantial heterogeneity in results was observed across data source and by condition.
- Creatine kinase (ck) and glomerular filtration rate (GFR) measurements identified a relatively small number of patients of rhabdomyolysis and end-stage renal disease respectively.
- A considerable number of patients was identified using measurement among the blood disorders.
- Among blood disorders, the characteristics of patients who met each definition were comparable (figure 1).

DISCUSSION

- Using measurements can significantly impact the sensitivity and specificity of the phenotype algorithm .
- We provide a framework for evaluating the utility of measurements for defining a phenotype within a given data source.

Azza Shoaibi, Gowtham Rao, Dmytro Dymshyts Anna Ostropolets, Patrick Ryan

Johnson & Johnson Epidemiology
OFFICE OF THE CHIEF MEDICAL OFFICER

OHDSI

#OHDSISocialShowcase This Week

THURSDAY

Comparing Patient Self-Reported Symptoms with SNOMED/ICD-10-CM Codes at Primary Care Visits

(Victor M. Castro, Danielle M. Crookes, Vivian Gainer, Shawn N. Murphy, Justin Manjourides)

Comparing Patient Self-Reported Symptoms and SNOMED/ICD-10-CM Codes at Primary Care Visits

PRESENTER: **Victor M. Castro**

INTRO

- **Measurement error** in RWD study outcomes increases bias and compromises study validity.
- Many **symptom** outcomes in RWD studies rely on **SNOMED and ICD-10-CM codes.**
- In this study we aim to **assess the sensitivity of clinician-recorded symptom SNOMED and ICD-10-CM codes compared to patient self-reports.**

METHODS

- Over **400K primary care patients completed a symptom screening questionnaire prior to their visits** between 2019 and 2021.
- We compared symptom self-report rates to clinician recorded symptoms in the EHR.

RESULTS

- Of the 15 symptoms evaluated, 13 were reported by patients much higher than recorded by clinicians in the EHR (Figure 1).
- Anxiety and depression were coded by physicians at a higher than reported by patients.

CONCLUSIONS

- Symptom outcomes defined by SNOMED or ICD-10-CM codes alone are likely to have poor sensitivity.

Primary care patients **self-report symptoms** up to **twelve times more** than **clinicians record** with structured codes in the EHR

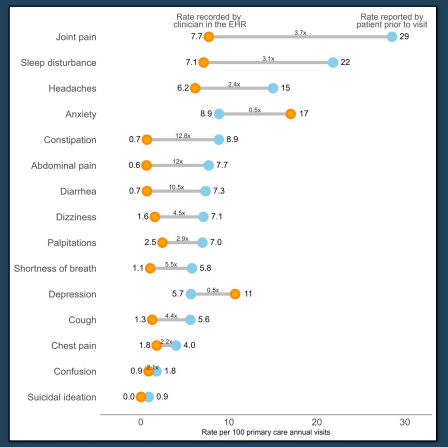


Figure 1. Case rates of symptom patient self-report (blue points) compared to clinician-recorded symptoms with ICD-10-CM or SNOMED codes (orange dots). Delta from self-report to clinician-recorded is also given above the line segment. Values below zero indicate clinicians record these symptoms more often the patients report at their visit.

Take a picture to download the full paper

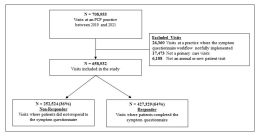


Figure 2. Study flow diagram

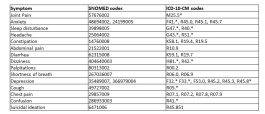


Table 1. Symptom SNOMED and ICD-10-CM codes

Table 2. Concordance between patient self-reported symptoms and clinical record symptom using SNOMED and ICD-10-CM codes.

Victor M. Castro, MS (1, 2), Danielle M. Crookes, PhD (1), Vivian Gainer, MS (2), Shawn N. Murphy, MD, PhD (2), Justin Manjourides, PhD (1, 3)

(1) Bouvé College of Health Sciences, Northeastern University, Boston, MA.,
(2) Research Information Science and Computing, Mass General Brigham, Somerville, MA.
(3) OHDSI Center at The Roux Institute, Northeastern University, Portland, ME.

Northeastern University

OHDSI

Mass General Brigham

#OHDSISocialShowcase This Week

FRIDAY

Using the Informatics for Integrating Biology and the Bedside Platform to Query OMOP Data in the OHDSI Ecosystem

(Jeffrey G. Klann, Griffin M. Weber, Michele Morris, Michael Mendis, Diane Keogh, Shawn N. Murphy)

OHDSI Broadsea Evolution

"Broadsea is the easiest way to install (& upgrade) the OHDSI tools"

10k+ Downloads

docker hub

- **v1.0** — Atlas/WebAPI & RStudio Docker images on Mac/Linux/Windows
- **v2.0** — Pre-populated demo postgres database image & Traefik reverse proxy
- **v3.0** — Docker profiles for a-la-carte services, more Traefik networking, environment variable driven deployment, new OHDSI apps, build from Git

```
docker-compose --profile default up -d
```

https://github.com/OHDSI/Broadsea

Opening: Junior Research Software Engineer, Tufts

Tufts | CTSI Tufts Clinical and Translational Science Institute

REDCap | TUFTS CTSI I LEARN | TUFTS PROFILES | CAREERS @ TUFTS CTSI

Request Services | Calendar | Contact Us

Research Services | Education | Funding Opportunities | Our Impact | Faculty & Staff | About Us

INFORMATICS

- Research Services
- COVID-19 Information and Resources
- Data and Safety Monitoring Board (DSMB) Program
- Center for Clinical Trials (CCT)
- Program Evaluation
- Qualitative and Mixed Methods Service
- Clinical Trial Design Labs
- Dissemination and Implementation (D&I) Core
- Science Communications

"Our Informatics team can help you collect and manage research data, develop databases, and identify study participants. We'll find the best data collection solution for your study. To get started, please submit a request below."

William Harvey, MD, MSc, FACR
Co-Director, Informatics and Tufts Medical Center CMIO

Overview

We participate in development of a robust institutional informatics infrastructure, enabling research teams to maintain their focus on scientific discovery and analyses rather than on data wrangling. Our infrastructure and support systems are dynamic, to keep pace with the changing and interdependent fields of health informatics, bioinformatics, statistics, and data science; expandable, to accommodate new data types and analytic methods; and scalable, to support efficient and methodologically rigorous multisite/institution research. These defining traits allow us to elucidate novel methods and operational principles, harmonize datasets, and create pipelines for data sharing and analytics.

Opening: Research Assistant, University of Oxford

UNIVERSITY OF OXFORD

UK date and time: 23-April-2024 15:25

Applicant Options

- New Search
- Login
- Job Details
- Help
- Terms of Use & Privacy Policy

Job Details

Research Assistant in Health Data Sciences

Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Botnar Research Centre, Windmill Road, Oxford, OX3 7LD

We have an exciting opportunity for a Research Assistant in Health Data Sciences to join the Pharmaco- and Device epidemiology research group led by Professor Daniel Prieto-Alhambra at the Botnar Research Centre, NDORMS, University of Oxford. The NDORMS Pharmaco- and Device epidemiology research group is involved in a number of national and international studies exploring the conditions of use (adherence, compliance, off and on-label use) of a number of licensed drugs, devices, and vaccines for the prevention and treatment of human disease in 'real world' (routine practice) conditions.

As a Research Assistant in Health Data Sciences you will contribute to the programming of analytical pipelines for the analysis of routinely collected data mapped to the OMOP Common Data Model. You will analyse real world data to address regulatory questions related to the prevalence/incidence of disease, use of medicines/vaccines, and the risks or benefits of medicines/vaccines or devices. You will prepare analytical packages to run a number of pre-specified analyses, contribute to wider project planning, including ideas for new research projects and gather, analyse, and present scientific data from a variety of sources.

You will hold a relevant BA or MSc degree in Mathematics, Engineering, or a related field. Knowledge of medical statistics and experience analysing large datasets, experience in biostatistics and/or health data sciences and experience in the programming of R packages are essential. Experience in propensity scores, overlap weighting, inverse probability weighting and/or similar methods, expertise in pharmaco or vaccine epidemiology and experience of working with electronic medical records/routinely collected data are desirable.

This is a full-time fixed-term appointment for 2 years.

The closing date for this position is 12 noon on 10 May 2024. You will be required to upload a CV and supporting statement as part of your online application.

Contact Person :	HR Team, NDORMS	**Vacancy ID :**	172348
Contact Phone :		**Closing Date & Time :**	10-May-2024 12:00
Pay Scale :	STANDARD GRADE 6	**Contact Email :**	hr@ndorms.ox.ac.uk
Salary (£) :	£32,332 - £38,205 p.a		

Opening: Biomedical Informatics Data Scientist at Stanford

Stanford MEDICINE Health Care

Who We Are | What We Offer | How We Hire | Career Areas | Search

Biomedical Informatics Data Scientist

1.0 FTE • Full time • Day - 08 Hour • R2335119 • Hybrid • 84866 IT RESEARCH • Technology & Digital Solutions • 455 Broadway,REDWOOD CITY,California •

1.0 FTE Full time Day - 08 Hour R2335119 Hybrid 84866 IT RESEARCH Technology & Digital Solutions 455 Broadway,REDWOOD CITY,California

If you're ready to be part of our legacy of hope and innovation, we encourage you to take the first step and explore our current job openings. Your best is waiting to be discovered.

Day - 08 Hour (United States of America)

This is a Stanford Health Care job.

A Brief Overview

The Biomedical Informatics Data Scientist will partner with researchers and clinicians to enable effective and efficient use of data and resources available via Stanford's research clinical data repository (STARR) including the Electronic Health Records in the OMOP Common Data Model, radiology and cardiology imaging data and associated metadata, and new data types as they get integrated along with their databases and respective cohort query tools and interfaces e.g., OHDSI ATLAS. This individual will enable researchers to maximize their understanding, interpretation and use of these clinical and research tools for more informed and productive research, clinical trials, patient care and quality outcome projects.

Clean, extract, transform and analyze various kinds of clinical data to create analysis-ready datasets that follow the FAIR (Findable, Accessible, Interoperable and Re-usable) principles. Partner with researchers and clinicians to enable effective and efficient use of Stanford Clinical data and resources for the advancement of research and the educational mission.

(I²DB)

Postdoc/Senior Data Analyst Opening at WashU

The Zhang Lab at Washington University School of Medicine in St. Louis has **one postdoct/senior data analyst position** to work on **causal machine learning** and **responsible AI** for reliable real-world evidence generation.

PI: Linying Zhang, PhD

- More details at https://linyingzhang.com
  - Postdoc: https://linyingzhang.com/files/Postdoc.pdf
  - Data analyst: https://linyingzhang.com/files/Analyst.pdf
- If interested, please send CV and cover letter to linyingz@wustl.edu

Washington University School of Medicine in St. Louis

Director, RWE at Gilead

Director, RWE - Data Science - OHDSI

Apply

Responsibilities:

Collaborate with researchers and data scientists to understand project requirements and translate them into OHDSI-compatible solutions. Work with databases, ensuring data integrity and optimization for OHDSI-related queries and analyses. Perform data analyses in OHDSI-related tools like ATLAS. Customize and extend OHDSI tools and applications to meet specific project needs. Collaborate with cross-functional teams to troubleshoot and resolve technical issues related to OHDSI implementations. Stay informed about OHDSI community updates, best practices, and emerging trends in observational health data research. Contribute to the development and documentation of data standards and conventions within the OHDSI community.

About Us

GILEAD

Gilead Sciences, Inc. is a biopharmaceutical company that has pursued and achieved breakthroughs in medicine for more than three decades, with the goal of creating a healthier world for all people. The company is committed to advancing innovative medicines to prevent and treat life-threatening diseases, including HIV, viral hepatitis and cancer. Gilead operates in more than 35 countries worldwide, with headquarters in Foster City, California.

Where Are We Going?

Any other announcements of upcoming work, events, deadlines, etc?

Three Stages of The Journey

Where Have We Been?

Where Are We Now?

Where Are We Going?

April 30: April Olympians Advancements

Clair Blacketer

Director, Observational Health Data Analytics

Janssen Research & Development

Melanie Philofsky

Senior Business Analyst and Project Manager

Odysseus Data Services, Inc.

The weekly OHDSI community call is held every Tuesday at 11 am ET.

Everybody is invited!

Links are sent out weekly and available at: ohdsi.org/community-calls