

OHDSI内では、実名での活動になります。  
Zoom参加時も「名前は実氏名で」お願いします。



**OHDSI**  
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

オデッセイ  
ジャパン

# OHDSI Japan evening conference #55

イブニング カンファレンス(第55回)

2024.6.28



## 本日の内容

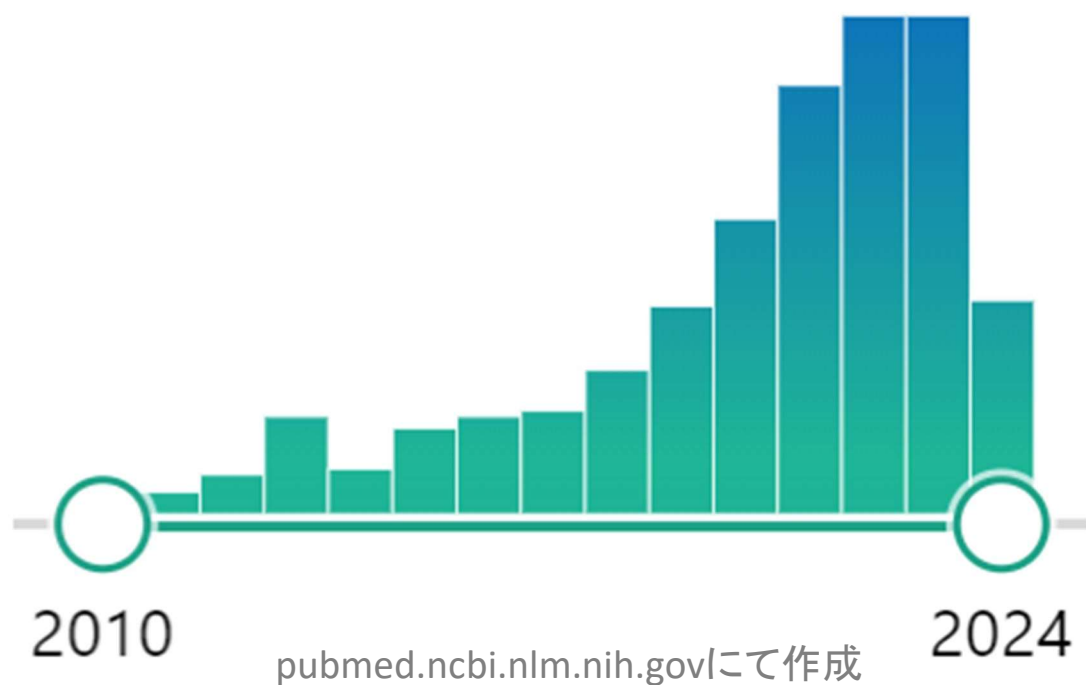
- OHDSI関連論文紹介
- OHDSI Global の話題から
- OHDSI APAC 各地域の状況

CBER BEST Initiative Seminarの報告



# OHDSI関連論文

Pubmedで“OHDSI or OMOP”を検索



全期間累計：5月432本→6月440本

- 検索に漏れているものがあるため、実際は累計500本を超えている。
- 年間では約100本ペース。

1. Towards an Electronic Health Prevention Record Based on HL7 FHIR and the OMOP Common Data Model. *Stud Health Technol Inform.* 2024 Apr 26;313:107-112. doi: 10.3233/SHTI240020. PMID: 38682513
2. The Health Equity Explorer: An open-source resource for distributed health equity visualization and research across common data models. *J Clin Transl Sci.* 2024 Apr 5;8(1):e72. doi: 10.1017/cts.2024.500. eCollection 2024. PMID: 38690224
3. Symptoms and Conditions in Children and Adults up to 90 Days after SARS-CoV-2 Infection: A Retrospective Observational Study Utilizing the Common Data Model. *J Clin Med.* 2024 May 15;13(10):2911. PMID: 38792452
4. Converting OMOP CDM to phenopackets: A model alignment and patient data representation evaluation. *J Biomed Inform.* 2024 May 21;155:104659. PMID: 38777085
5. Calculating daily dose in the Observational Medical Outcomes Partnership Common Data Model. *Pharmacoepidemiol Drug Saf.* 2024 Jun;33(6):e5809. PMID: 38773798
6. Taipei Medical University Clinical Research Database: a collaborative hospital EHR database aligned with international common data standards. *BMJ Health Care Inform.* 2024 May 14;31(1):e100890. PMID: 38749529
7. Calculating daily dose in the Observational Medical Outcomes Partnership Common Data Model. *Pharmacoepidemiol Drug Saf.* 2024 Jun;33(6):e5809. doi: 10.1002/pds.5809. PMID: 38773798
8. Converting OMOP CDM to phenopackets: A model alignment and patient data representation evaluation. *J Biomed Inform.* 2024 Jul;155:104659. doi: 10.1016/j.jbi.2024.104659. Epub 2024 May 21. PMID: 38777085
9. Symptoms and Conditions in Children and Adults up to 90 Days after SARS-CoV-2 Infection: A Retrospective Observational Study Utilizing the Common Data Model. *J Clin Med.* 2024 May 15;13(10):2911. doi: 10.3390/jcm13102911. PMID: 38792452
10. Data harmonization and federated learning for multi-cohort dementia research using the OMOP common data model: A Netherlands consortium of dementia cohorts case study. *J Biomed Inform.* 2024 Jul;155:104661. doi: 10.1016/j.jbi.2024.104661. Epub 2024 May 26. PMID: 38806105
11. MENDS-on-FHIR: leveraging the OMOP common data model and FHIR standards for national chronic disease surveillance. *JAMIA Open.* 2024 May 29;7(2):ooae045. doi: 10.1093/jamiaopen/ooae045. eCollection 2024 Jul. PMID: 38818114
12. OHDSI-compliance: a set of document templates facilitating the implementation and operation of a software stack for real-world evidence generation. *Front Med (Lausanne).* 2024 May 16;11:1378866. doi: 10.3389/fmed.2024.1378866. eCollection 2024. PMID: 38818399
13. Sharing sensitive data in life sciences: an overview of centralized and federated approaches. *Brief Bioinform.* 2024 May 23;25(4):bbae262. doi: 10.1093/bib/bbae262. PMID: 38836701
14. Research protocol for an observational health data analysis to assess the applicability of randomized controlled trials focusing on newly diagnosed metastatic prostate cancer using real-world data: PIONEER IMI's "big data for better outcomes" program. *Int J Surg Protoc.* 2024 Apr 15;28(2):64-72. doi: 10.1097/SP9.0000000000000024. eCollection 2024 Jun. PMID: 38854711
15. Transforming a Large-Scale Prostate Cancer Outcomes Dataset to the OMOP Common Data Model-Experiences from a Scientific Data Holder's Perspective. *Cancers (Basel).* 2024 May 30;16(11):2069. doi: 10.3390/cancers16112069. PMID: 38893186
16. Streamlining intersectoral provision of real-world health data: a service platform for improved clinical research and patient care. *Front Med (Lausanne).* 2024 Jun 5;11:1377209. doi: 10.3389/fmed.2024.1377209. eCollection 2024. PMID: 38903818
17. Learning debiased graph representations from the OMOP common data model for synthetic data generation. *BMC Med Res Methodol.* 2024 Jun 22;24(1):136. doi: 10.1186/s12874-024-02257-8. PMID: 38909216



## 予防医療のITソリューション構想

➤ [Stud Health Technol Inform.](#) 2024 Apr 26;313:107-112. doi: 10.3233/SHTI240020.

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

Markus Falgenhauer<sup>1</sup>, Aaron Lauschensky<sup>1</sup>, Karl Kreiner<sup>1</sup>, Stefan Beyer<sup>1</sup>, Kristina Reiter<sup>2</sup>, Andreas Ziegl<sup>1 2</sup>, Robert Modre-Osprian<sup>2</sup>, Bernhard Pfeifer<sup>3 4</sup>, Sabrina Neururer<sup>3 4</sup>, Susanne Krestan<sup>3</sup>, Hanna Wagner<sup>3</sup>, Andreas Huber<sup>3</sup>, Sandra Plaikner<sup>5</sup>, Sarah Kuppelwieser<sup>5 6</sup>, Martin Widschwendter<sup>5 6</sup>, Günter Schreier<sup>1</sup>

Affiliations + expand

PMID: 38682513 DOI: [10.3233/SHTI240020](#)

### Affiliations

- <sup>1</sup> AIT Austrian Institute of Technology GmbH, Graz, Austria.
- <sup>2</sup> telbiomed Medizintechnik und IT Service GmbH, Graz, Austria.
- <sup>3</sup> Tyrolean Federal Institute for Integrated Care, Tirol Kliniken, Innsbruck, Austria.
- <sup>4</sup> Division for Digital Health and Telemedicine, UMIT TIROL - Private University for Health Sciences and Health Technology, Hall (Tyrol), Austria.
- <sup>5</sup> EUTOPS Institute, University of Innsbruck, Innsbruck, Austria.
- <sup>6</sup> Tirol Kliniken GmbH, Innsbruck, Austria.

## 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.

**Results:** An IT solution supporting a structured and modular prevention pathway was conceptualized. It contained a personalized management of prevention, risk assessment, diagnostic and preventive measures supported by a modular, interoperable IT infrastructure including a health app, prevention record web-service, decision support modules and a smart prevention registry, separating primary and secondary use of data.

**Conclusion:** A concept was created on how an electronic health prevention record based on HL7 FHIR and the OMOP common data model can be implemented.

## 要旨

**背景:** オーストリアで記録されている全死亡の約40%は行動リスクによるものである。これらのリスクは適切な対策によって回避できる可能性がある。

**目的:** 一次予防と二次予防に焦点を当てた、EHRとEMRの概念の電子予防記録への拡張。

**方法:** P4医学の原則に基づく構造化された予防経路の概念が、多職種による予防ネットワークのために開発された。HL7 FHIRおよびOHDSI OMOP共通データモデルに基づくIT基盤が設計された。

**結果:** 構造化されモジュール化された予防経路をサポートするITソリューションが構想された。このソリューションには、健康アプリ、予防記録ウェブサービス、意思決定支援モジュール、スマート予防レジストリを含む、モジュール化された相互運用可能なIT基盤によってサポートされる予防の個別化管理、リスク評価、診断、予防措置が含まれ、データの一次利用と二次利用を分離した。

**結論:** HL7 FHIRとOMOP共通データモデルに基づく電子健康予防記録の実装方法に関するコンセプトが作成された。

## ※ P4 Medicine

予測を立てて（predictive）個人に合わせた（personalized）形で、予防的（preventive）で参加型（participatory）の処置を行う医療。これからの医療のあり方を示すモデル。





> J Clin Transl Sci. 2024 Apr 5;8(1):e72. doi: 10.1017/cts.2024.500. eCollection 2024.

## The Health Equity Explorer: An open-source resource for distributed health equity visualization and research across common data models

William G Adams <sup>1 2</sup>, Sarah Gasman <sup>1</sup>, Ariel L Beccia <sup>3</sup>, Liza Fuentes <sup>4</sup>

Affiliations + expand

PMID: 38690224 PMCID: PMC11058576 DOI: 10.1017/cts.2024.500

### Affiliations

- 1 Department of Pediatrics, Boston Medical Center, Boston, MA, USA.
- 2 Boston University Clinical and Translational Science Institute, Chobanian & Avedisian School of Medicine, Boston, MA, USA.
- 3 Boston Children's Hospital and Department of Pediatrics, Harvard Medical School, Boston, MA, USA.
- 4 Health Equity Accelerator, Boston Medical Center, Boston, MA, USA.

## Abstract

**Introduction:** There is an urgent need to address pervasive inequities in health and healthcare in the USA. Many areas of health inequity are well known, but there remain important unexplored areas, and for many populations in the USA, accessing data to visualize and monitor health equity is difficult.

**Methods:** We describe the development and evaluation of an open-source, R-Shiny application, the "Health Equity Explorer (H2E)," designed to enable users to explore health equity data in a way that can be easily shared within and across common data models (CDMs).

**Results:** We have developed a novel, scalable informatics tool to explore a wide variety of drivers of health, including patient-reported Social Determinants of Health (SDoH), using data in an OMOP CDM research data repository in a way that can be easily shared. We describe our development process, data schema, potential use cases, and pilot data for 705,686 people who attended our health system at least once since 2016. For this group, 996,382 unique observations for questions related to food and housing security were available for 324,630 patients (at least one answer for all 46% of patients) with 65,152 (20.1% of patients with at least one visit and answer) reporting food or housing insecurity at least once.

**Conclusions:** H2E can be used to support dynamic and interactive explorations that include rich social and environmental data. The tool can support multiple CDMs and has the potential to support distributed health equity research and intervention on a national scale.

## 要旨

**はじめに:** 米国では、健康と医療における広範な不平等に対処することが急務である。健康格差の多くの領域はよく知られているが、未調査の重要な領域も残っており、米国の多くの集団にとって、健康格差を可視化しモニタリングするためのデータにアクセスすることは困難である。

**方法:** 我々は、オープンソースのR-Shinyアプリケーションである "Health Equity Explorer (H2E) "の開発と評価について述べる。"Health Equity Explorer (H2E) "は、共通のデータモデル (CDM) 内およびCDM間で容易に共有できる方法で、ユーザーが健康の公平性に関するデータを探索できるように設計されている。

**結果:** 我々は、OMOP CDM研究データリポジトリのデータを用いて、患者報告による健康の社会的決定要因 (Social Determinants of Health : SDoH) を含む、健康の多種多様な要因を容易に共有できる方法で探索するための、新規で拡張可能な情報学的ツールを開発した。我々は、開発プロセス、データスキーマ、潜在的なユースケース、および2016年以降に我々の医療システムに少なくとも1回参加した705,686人のパイロットデータについて説明する。このグループについては、324,630人の患者 (患者の46%すべてについて少なくとも1つの回答) に対して、食料および住居の安全保障に関連する質問に対する996,382のユニークな観察が利用可能であり、65,152人 (少なくとも1回の受診および回答を有する患者の20.1%) が少なくとも1回、食料または住居の不安を報告した。

**結論:** H2Eは、豊富な社会・環境データを含むダイナミックでインタラクティブな探索をサポートするために使用できる。このツールは、複数のCDMをサポートすることが可能であり、分散型CDMをサポートする可能性がある。





## 台湾TMUでのOMOP DB

➤ [BMJ Health Care Inform.](#) 2024 May 14;31(1):e100890. doi: 10.1136/bmjhci-2023-100890.

### Taipei Medical University Clinical Research Database: a collaborative hospital EHR database aligned with international common data standards

Phung-Anh Nguyen<sup>1 2 3</sup>, Min-Huei Hsu<sup>4 5</sup>, Tzu-Hao Chang<sup>3 6 7</sup>, Hsuan-Chia Yang<sup>3 6 7 8</sup>,  
Chih-Wei Huang<sup>6 7</sup>, Chia-Te Liao<sup>9 10 11</sup>, Christine Y Lu<sup>12 13 14</sup>, Jason C Hsu<sup>15 2 3 16</sup>

Affiliations + expand

PMID: 38749529 PMCID: [PMC11097871](#) DOI: [10.1136/bmjhci-2023-100890](#)

#### Affiliations

- 1 Clinical Data Center, Office of Data Science, Taipei Medical University, Taipei, Taiwan.
- 2 Research Center of Health Care Industry Data Science, College of Management, Taipei Medical University, Taipei, Taiwan.
- 3 Clinical Big Data Research Center, Taipei Medical University Hospital, Taipei Medical University, Taipei, Taiwan.
- 4 Office of Data Science, Taipei Medical University, Taipei, Taiwan.
- 5 Graduate Institute of Data Science, College of Management, Taipei Medical University, Taipei,

## Abstract

**Objective:** The objective of this paper is to provide a comprehensive overview of the development and features of the Taipei Medical University Clinical Research Database (TMUCRD), a repository of real-world data (RWD) derived from electronic health records (EHRs) and other sources.

**Methods:** TMUCRD was developed by integrating EHRs from three affiliated hospitals, including Taipei Medical University Hospital, Wan-Fang Hospital and Shuang-Ho Hospital. The data cover over 15 years and include diverse patient care information. The database was converted to the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM) for standardisation.

**Results:** TMUCRD comprises 89 tables (eg, 29 tables for each hospital and 2 linked tables), including demographics, diagnoses, medications, procedures and measurements, among others. It encompasses data from more than 4.15 million patients with various medical records, spanning from the year 2004 to 2021. The dataset offers insights into disease prevalence, medication usage, laboratory tests and patient characteristics.

**Discussion:** TMUCRD stands out due to its unique advantages, including diverse data types, comprehensive patient information, linked mortality and cancer registry data, regular updates and a swift application process. Its compatibility with the OMOP CDM enhances its usability and interoperability.

**Conclusion:** TMUCRD serves as a valuable resource for researchers and scholars interested in leveraging RWD for clinical research. Its availability and integration of diverse healthcare data contribute to a collaborative and data-driven approach to advancing medical knowledge and practice.

## 要旨

**目的:** 本論文の目的は、電子カルテ(EHR)やその他の情報源から得られた実世界データ(RWD)のリポジトリである台北医学大学臨床研究データベース(TMUCRD)の開発と特徴の包括的な概要を提供することである。

**方法:** TMUCRDは、台北医学大学附属病院、萬芳病院、双和病院を含む3つの附属病院のEHRを統合して開発された。データは15年以上をカバーし、多様な患者ケア情報を含む。データベースは標準化のためにOMOP CDMに変換された。

**結果:** TMUCRDは、89のテーブル（例えば、各病院の29のテーブルと2つのリンクされたテーブル）から構成され、人口統計、診断、投薬、処置、測定などを含む。TMUCRDは、2004年から2021年までの様々な医療記録を持つ415万人以上の患者のデータを包含している。このデータセットは、疾患の有病率、薬剤の使用状況、臨床検査、患者の特徴に関する洞察を提供する。

**考察:** TMUCRDは、多様なデータタイプ、包括的な患者情報、リンクされた死亡率とがん登録データ、定期的な更新、迅速な申請プロセスなど、そのユニークな利点により際立っている。OMOP CDMとの互換性は、その使いやすさと相互運用性を高めている。

**結論:** TMUCRDは、臨床研究にRWDを活用することに興味を持つ研究者や学者にとって、貴重なリソースとなる。TMUCRDの多様なヘルスケアデータの利用可能性と統合は、医学の知識と実践を進歩させるための共同でデータ駆動型のアプローチに貢献している。



## 薬物の 1 日投与量を計算する方式

➤ [Pharmacoepidemiol Drug Saf. 2024 Jun;33\(6\):e5809. doi: 10.1002/pds.5809.](#)

### Calculating daily dose in the Observational Medical Outcomes Partnership Common Data Model

Theresa Burkard <sup>1</sup>, Kim López-Güell <sup>1</sup>, Artem Gorbachev <sup>2</sup>, Lucía Bellas <sup>3</sup>, Annika M Jödicke <sup>1</sup>, Edward Burn <sup>1</sup>, Maria de Ridder <sup>4</sup>, Mees Mosseveld <sup>4</sup>, Jasmine Gratton <sup>5</sup>, Sarah Seager <sup>5</sup>, Dina Vojinovic <sup>6</sup>, Miguel Angel Mayer <sup>7 8</sup>, Juan Manuel Ramírez-Anguila <sup>8 9</sup>, Angela Leis Machín <sup>8</sup>, Marek Oja <sup>10</sup>, Raivo Kolde <sup>10</sup>, Klaus Bonadt <sup>11</sup>, Daniel Prieto-Alhambra <sup>1 4</sup>, Chistian Reich <sup>2</sup>, Martí Català <sup>1</sup>

Affiliations + expand

PMID: 38773798 DOI: [10.1002/pds.5809](#)

#### Affiliations

- 1 Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), University of Oxford, Oxford, UK.
- 2 Odysseus Data Services, Cambridge, Massachusetts, USA.
- 3 Clinical Pharmacology Service, Hospital Universitari Vall d'Hebron, Barcelona, Spain.
- 4 Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands.
- 5 IQVIA Ltd, London, UK.
- 6 IQVIA Solutions B.V. Amsterdam, the Netherlands.

## Abstract

**Purpose:** We aimed to develop a standardized method to calculate daily dose (i.e., the amount of drug a patient was exposed to per day) of any drug on a global scale using only drug information of typical observational data in the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM) and a single reference table from Observational Health Data Sciences And Informatics (OHDSI).

**Materials and methods:** The OMOP DRUG\_STRENGTH reference table contains information on the strength or concentration of drugs, whereas the OMOP DRUG\_EXPOSURE table contains information on patients' drug prescriptions or dispensations/claims. Based on DRUG\_EXPOSURE data from the primary care databases Clinical Practice Research Datalink GOLD (United Kingdom) and Integrated Primary Care Information (IPCI, The Netherlands) and healthcare claims from PharMetrics® Plus for Academics (USA), we developed four formulas to calculate daily dose given different DRUG\_STRENGTH reference table information. We tested the dose formulas by comparing the calculated median daily dose to the World Health Organization (WHO) Defined Daily Dose (DDD) for six different ingredients in those three databases and additional four international databases representing a variety of healthcare settings: MAITT (Estonia, healthcare claims and discharge summaries), IQVIA Disease Analyzer Germany (outpatient data), IQVIA Longitudinal Patient Database Belgium (outpatient data), and IMASIS Parc Salut (Spain, hospital data). Finally, in each database, we assessed the proportion of drug records for which daily dose calculations were possible using the suggested formulas.

**Results:** Applying the dose formulas, we obtained median daily doses that generally matched the WHO DDD definitions. Our dose formulas were applicable to >85% of drug records in all but one of the assessed databases.

**Conclusion:** We have established and implemented a standardized daily dose calculation in OMOP CDM providing reliable and reproducible results.

## 概要

**目的:** OMOP CDMの典型的な観察データの薬剤情報とOHDSIの単一の参照表のみを用いて、あらゆる薬剤の1日投与量（すなわち、患者が1日に曝露した薬剤の量）を世界規模で算出する標準的な方法を開発することを目的とした。

**材料と方法:** OMOP DRUG\_STRENGTH参照表には薬剤の強度または濃度に関する情報が含まれ、OMOP DRUG\_EXPOSURE表には患者の薬剤処方または調剤/請求に関する情報が含まれる。プライマリケアデータベースであるClinical Practice Research Datalink GOLD（英国）とIntegrated Primary Care Information（オランダ）のDRUG\_EXPOSUREデータと、PharMetrics® Plus for Academics（米国）の医療費請求に基づいて、異なるDRUG\_STRENGTH参照テーブル情報が与えられた場合の1日投与量を計算するための4つの公式を開発した。これらの3つのデータベースと、様々な医療現場を代表する国際的な4つのデータベースで、6つの異なる成分について、計算された1日用量の中央値を世界保健機関（WHO）の定義1日用量（DDD）と比較することにより、用量計算式をテストした：MAITT（エストニア、医療費請求および退院サマリー）、IQVIA Disease Analyzer Germany（外来患者データ）、IQVIA Longitudinal Patient Database Belgium（外来患者データ）、IMASIS Parc Salut（スペイン、病院データ）である。最後に、各データベースにおいて、提案された計算式を用いて1日投与量の計算が可能であった薬剤記録の割合を評価した。

**結果:** 用量計算式を適用した結果、WHOのDDD定義に概ね一致する1日用量の中央値が得られた。われわれの用量公式は、評価したデータベースのうち1つを除くすべての薬剤記録の85%以上に適用可能であった。

**結論:** 我々は、OMOP CDMにおける標準化された1日投与量計算を確立し、実施した。





## OMOPからPhenopacketsへの新たな変換モデル

> J Biomed Inform. 2024 Jul;155:104659. doi: 10.1016/j.jbi.2024.104659. Epub 2024 May 21.

# Converting OMOP CDM to phenopackets: A model alignment and patient data representation evaluation

Kayla Schiffer-Kane<sup>1</sup>, Cong Liu<sup>1</sup>, Tiffany J Callahan<sup>1</sup>, Casey Ta<sup>1</sup>, Jordan G Nestor<sup>2</sup>, Chunhua Weng<sup>3</sup>

Affiliations + expand

PMID: 38777085 PMCID: PMC11181468 (available on 2025-07-01) DOI: 10.1016/j.jbi.2024.104659

### Affiliations

- <sup>1</sup> Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, NY, USA.
- <sup>2</sup> Department of Medicine, Division of Nephrology, Columbia University Irving Medical Center, New York, NY, USA.
- <sup>3</sup> Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, NY, USA. Electronic address: cw2384@cumc.columbia.edu.

## Abstract

**Objective:** This study aims to promote interoperability in precision medicine and translational research by aligning the Observational Medical Outcomes Partnership (OMOP) and Phenopackets data models. Phenopackets is an expert knowledge-driven schema designed to facilitate the storage and exchange of multimodal patient data, and support downstream analysis. The first goal of this paper is to explore model alignment by characterizing the common data models using a newly developed data transformation process and evaluation method. Second, using OMOP normalized clinical data, we evaluate the mapping of real-world patient data to Phenopackets. We evaluate the suitability of Phenopackets as a patient data representation for real-world clinical cases.

**Methods:** We identified mappings between OMOP and Phenopackets and applied them to a real patient dataset to assess the transformation's success. We analyzed gaps between the models and identified key considerations for transforming data between them. Further, to improve ambiguous alignment, we incorporated Unified Medical Language System (UMLS) semantic type-based filtering to direct individual concepts to their most appropriate domain and conducted a domain-expert evaluation of the mapping's clinical utility.

**Results:** The OMOP to Phenopacket transformation pipeline was executed for 1,000 Alzheimer's disease patients and successfully mapped all required entities. However, due to missing values in OMOP for required Phenopacket attributes, 10.2 % of records were lost. The use of UMLS-semantic type filtering for ambiguous alignment of individual concepts resulted in 96 % agreement with clinical thinking, increased from 68 % when mapping exclusively by domain correspondence.

**Conclusion:** This study presents a pipeline to transform data from OMOP to Phenopackets. We identified considerations for the transformation to ensure data quality, handling restrictions for successful Phenopacket validation and discrepant data formats. We identified unmappable Phenopacket attributes that focus on specialty use cases, such as genomics or oncology, which OMOP does not currently support. We introduce UMLS semantic type filtering to resolve ambiguous alignment to Phenopacket entities to be most appropriate for real-world interpretation. We provide a systematic approach to align OMOP and Phenopackets schemas. Our work facilitates future use of Phenopackets in clinical applications by addressing key barriers to interoperability when deriving a Phenopacket from real-world patient data.



## 要旨

**目的:** 本研究は、OMOPとPhenopacketsのデータモデルを整合させることにより、精密医療とトランスレーショナル研究における相互運用性を促進することを目的とする。Phenopacketsは、マルチモーダルな患者データの保存と交換を容易にし、下流の分析をサポートするために設計されたエキスパート知識駆動型スキーマである。本論文の第一の目的は、新たに開発したデータ変換プロセスと評価方法を用いて共通のデータモデルを特徴付けることにより、モデルの整合性を探ることである。次に、OMOPで正規化された臨床データを用いて、実際の患者データのPhenopacketsへのマッピングを評価する。実臨床症例に対する患者データ表現としてのPhenopacketsの適合性を評価する。

**方法:** OMOPとPhenopackets間のマッピングを特定し、実際の患者データセットに適用して変換の成功を評価した。両モデル間のギャップを分析し、両モデル間でデータを変換するための重要な考慮事項を特定した。さらに、曖昧なアライメントを改善するために、Unified Medical Language System (UMLS) のセマンティックタイプに基づくフィルタリングを組み込んで、個々の概念を最も適切なドメインに誘導し、マッピングの臨床的有用性についてドメインの専門家による評価を実施した。

**結果:** OMOPからPhenopacketへの変換パイプラインは、1,000人のアルツハイマー病患者に対して実行され、すべての必要なエンティティのマッピングに成功した。しかし、必要な Phenopacket 属性の値が OMOP で欠落していたため、10.2 % のレコードが失われた。個々の概念の曖昧なアライメントにUMLS-semantic type filteringを使用した結果、臨床思考との一致率は96%に達し、ドメイン対応のみでマッピングした場合の68%から増加した。

**結論:** 本研究は、OMOPからPhenopacketsへデータを変換するパイプラインを提示した。我々は、データ品質を確保するための変換の考慮事項、Phenopacketの検証を成功させるための制限事項、および不一致のデータ形式を取り扱った。OMOPが現在サポートしていない、ゲノムや腫瘍学などの専門的なユースケースに特化した、適用不可能なPhenopacket属性を特定した。UMLS のセマンティック・タイプ・フィルタリングを導入し、Phenopacket エンティティへの曖昧なアライメントを解決することで、実世界での解釈に最適なアライメントを行う。OMOPとPhenopacketsのスキーマを整合させる体系的なアプローチを提供する。実世界の患者データから Phenopacket を導き出す際の相互運用性に対する主要な障壁に対処することで、臨床アプリケーションにおける Phenopacket の今後の利用を促進する。



## COVID感染後の持続症状

> J Clin Med. 2024 May 15;13(10):2911. doi: 10.3390/jcm13102911.

HIRA OMOPデータベース  
1000万人サンプル

# Symptoms and Conditions in Children and Adults up to 90 Days after SARS-CoV-2 Infection: A Retrospective Observational Study Utilizing the Common Data Model

Minjung Han<sup>1 2</sup>, Taehee Chang<sup>1</sup>, Hae-Ryoung Chun<sup>1</sup>, Suyoung Jo<sup>3</sup>, Yeongchang Jo<sup>4</sup>,  
Dong Han Yu<sup>5</sup>, Sooyoung Yoo<sup>6</sup>, Sung-Il Cho<sup>1 3</sup>

Affiliations + expand

PMID: 38792452 PMCID: PMC11122571 DOI: 10.3390/jcm13102911

### Affiliations

- 1 Graduate School of Public Health, Seoul National University, Seoul 08826, Republic of Korea.
- 2 Chaum Life Center, CHA University School of Medicine, Seoul 06062, Republic of Korea.
- 3 Institute of Health and Environment, Seoul National University, Seoul 08826, Republic of Korea.
- 4 Department of Preventive Medicine, Yonsei University College of Medicine, Seoul 03722, Republic of Korea.
- 5 Big Data Department, Health Insurance Review and Assessment Service, Wonju 26465, Republic of Korea.
- 6 Healthcare ICT Research Center, Seoul National University Bundang Hospital, Seongnam 13620, Republic of Korea.

## Abstract

**Background/Objectives:** There have been widespread reports of persistent symptoms in both children and adults after SARS-CoV-2 infection, giving rise to debates on whether it should be regarded as a separate clinical entity from other postviral syndromes. This study aimed to characterize the clinical presentation of post-acute symptoms and conditions in the Korean pediatric and adult populations. **Methods:** A retrospective analysis was performed using a national, population-based database, which was encoded using the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). We compared individuals diagnosed with SARS-CoV-2 to those diagnosed with influenza, focusing on the risk of developing prespecified symptoms and conditions commonly associated with the post-acute sequelae of COVID-19. **Results:** Propensity score matching yielded 1,656 adult and 343 pediatric SARS-CoV-2 and influenza pairs. Ninety days after diagnosis, no symptoms were found to have elevated risk in either adults or children when compared with influenza controls. Conversely, at 1 day after diagnosis, adults with SARS-CoV-2 exhibited a significantly higher risk of developing abnormal liver function tests, cardiorespiratory symptoms, constipation, cough, thrombophlebitis/thromboembolism, and pneumonia. In contrast, children diagnosed with SARS-CoV-2 did not show an increased risk for any symptoms during either acute or post-acute phases. **Conclusions:** In the acute phase after infection, SARS-CoV-2 is associated with an elevated risk of certain symptoms in adults. The risk of developing post-acute COVID-19 sequelae is not significantly different from that of having postviral symptoms in children in both the acute and post-acute phases, and in adults in the post-acute phase. These observations warrant further validation through studies, including the severity of initial illness, vaccination status, and variant types.

## 要旨

**背景/目的:** SARS-CoV-2感染後の小児および成人における症状の持続が広く報告されており、他のウイルス後症候群とは別の臨床像とみなすべきかどうか議論されている。本研究の目的は、韓国の小児および成人集団における急性感染後の症状および病態の特徴を明らかにすることである。

**方法:** レトロスペクティブ解析は、OMOP CDM を用いて符号化された全国的な集団ベースのデータベースを用いて行われた。SARS-CoV-2と診断された患者とインフルエンザと診断された患者を比較し、COVID-19の急性後遺症によく関連する症状や病態の発症リスクに焦点を当てた。

**結果:** 傾向スコアマッチングにより、成人SARS-CoV-2とインフルエンザのペア1,656例、小児SARS-CoV-2とインフルエンザのペア343例が得られた。診断から90日後、インフルエンザの対照群と比較した場合、成人、小児ともにリスクが上昇する症状は認められなかった。逆に、診断後1日目に、SARS-CoV-2と診断された成人は、肝機能検査異常、心肺症状、便秘、咳嗽、血栓性静脈炎/血栓塞栓症、肺炎を発症するリスクが有意に高かった。一方、SARS-CoV-2と診断された小児では、急性期、急性期後のいずれの症状においてもリスクの増加は認められなかった。

**結論:** 感染後の急性期において、SARS-CoV-2は成人の特定の症状のリスク上昇と関連している。COVID-19の急性後遺症を発症するリスクは、急性期、急性期後ともに小児において、また急性期後では成人において、ウイルス感染後の症状を発症するリスクと有意差はない。これらの観察結果は、初発疾患の重症度、ワクチン接種の有無、変異型を含む研究によってさらに検証する必要がある。



## コホートの連合学習ネットワーク

> J Biomed Inform. 2024 Jul;155:104661. doi: 10.1016/j.jbi.2024.104661. Epub 2024 May 26.

# Data harmonization and federated learning for multi-cohort dementia research using the OMOP common data model: A Netherlands consortium of dementia cohorts case study

Pedro Mateus<sup>1</sup>, Justine Moonen<sup>2</sup>, Magdalena Beran<sup>3</sup>, Eva Jaarsma<sup>4</sup>, Sophie M van der Landen<sup>2</sup>, Joost Heuvelink<sup>5</sup>, Mahlet Birhanu<sup>6</sup>, Alexander G J Harms<sup>6</sup>, Esther Bron<sup>6</sup>, Frank J Wolters<sup>7</sup>, Davy Cats<sup>8</sup>, Hailiang Mei<sup>8</sup>, Julie Oomens<sup>9</sup>, Willemijn Jansen<sup>9</sup>, Miranda T Schram<sup>10</sup>, Andre Dekker<sup>11</sup>, Inigo Bermejo<sup>11</sup>

### Affiliations

PMID: 38806105

- 1 Department of Radiation Oncology (Maastr), GROW School for Oncology and Reproduction, Maastricht University Medical Centre+, Maastricht, Netherlands. Electronic address: pedro.mateus@maastro.nl.
- 2 Alzheimer Center Amsterdam, Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC location VUmc, Amsterdam, Netherlands; Amsterdam Neuroscience, Neurodegeneration, Amsterdam, Netherlands.
- 3 Department of Internal Medicine, School for Cardiovascular Diseases (CARIM), Maastricht University, Maastricht, Netherlands; Department of Epidemiology and Global Health, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands.
- 4 Center for Nutrition, Prevention, and Health Services, National Institute for Public Health and

## Abstract

**Background:** Establishing collaborations between cohort studies has been fundamental for progress in health research. However, such collaborations are hampered by heterogeneous data representations across cohorts and legal constraints to data sharing. The first arises from a lack of consensus in standards of data collection and representation across cohort studies and is usually tackled by applying data harmonization processes. The second is increasingly important due to raised awareness for privacy protection and stricter regulations, such as the GDPR. Federated learning has emerged as a privacy-preserving alternative to transferring data between institutions through analyzing data in a decentralized manner.

**Methods:** In this study, we set up a federated learning infrastructure for a consortium of nine Dutch cohorts with appropriate data available to the etiology of dementia, including an extract, transform, and load (ETL) pipeline for data harmonization. Additionally, we assessed the challenges of transforming and standardizing cohort data using the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) and evaluated our tool in one of the cohorts employing federated algorithms.

**Results:** We successfully applied our ETL tool and observed a complete coverage of the cohorts' data by the OMOP CDM. The OMOP CDM facilitated the data representation and standardization, but we identified limitations for cohort-specific data fields and in the scope of the vocabularies available. Specific challenges arise in a multi-cohort federated collaboration due to technical constraints in local environments, data heterogeneity, and lack of direct access to the data.

**Conclusion:** In this article, we describe the solutions to these challenges and limitations encountered in our study. Our study shows the potential of federated learning as a privacy-preserving solution for multi-cohort studies that enhance reproducibility and reuse of both data and analyses.

## 要旨

**背景:** コホート研究間の共同研究の確立は、健康研究の進歩にとって基本である。しかし、このような共同研究は、コホート間の異質なデータ表現とデータ共有に対する法的制約によって妨げられている。前者は、コホート研究間のデータ収集と表現の基準におけるコンセンサスの欠如から生じ、通常はデータ調和プロセスを適用することで対処される。2つ目は、プライバシー保護に対する意識の高まりとGDPRのような規制強化により、ますます重要になってきている。フェデレーテッド・ラーニングは、分散型方法でデータを分析することにより、機関間でデータを転送するためのプライバシーを保護する選択肢として登場した。

**方法:** 本研究では、認知症の病因に適切なデータが利用可能な9つのオランダのコホートのコンソーシアムのために、データ整合化のための抽出・変換・ロード（ETL）パイプラインを含む連携学習インフラを構築した。さらに、OMOP CDM を用いてコホートデータの変換と標準化の課題を評価し、連合アルゴリズムを採用したコホートの1つで我々のツールを評価した。

**結果:** ETLツールの適用に成功し、OMOP CDMがコホートデータを完全にカバーしていることが確認された。OMOP CDMはデータ表現と標準化を容易にしたが、コホート特有のデータフィールドや利用可能な語彙の範囲には限界があることがわかった。複数のコホートが連携した共同研究においては、ローカル環境における技術的制約、データの異種性、データへの直接アクセスの欠如などにより、特有の課題が生じる。

**結論:** 本稿では、我々の研究で遭遇したこれらの課題に対する解決策と限界について述べる。我々の研究は、データと解析の両方の再現性と再利用性を高めるマルチコホート研究のためのプライバシー保持ソリューションとしての連合学習の可能性を示している。





## OMOP → FHIR → MENDS

➤ JAMIA Open. 2024 May 29;7(2):ooae045. doi: 10.1093/jamiaopen/ooae045. eCollection 2024 Jul.

# MENDS-on-FHIR: leveraging the OMOP common data model and FHIR standards for national chronic disease surveillance

Shahim Essaid <sup>1</sup>, Jeff Andre <sup>2</sup>, Ian M Brooks <sup>1 3</sup>, Katherine H Hohman <sup>4</sup>, Madelyne Hull <sup>3</sup>, Sandra L Jackson <sup>5</sup>, Michael G Kahn <sup>1 3</sup>, Emily M Kraus <sup>6 7</sup>, Neha Mandadi <sup>1 3</sup>, Amanda K Martinez <sup>4</sup>, Joyce Y Mui <sup>1 3</sup>, Bob Zambarano <sup>2</sup>, Andrey Soares <sup>8</sup>

PMID: 38818114

### Affiliations

- 1 Department of Biomedical Informatics, University of Colorado Anschutz Medical Campus, Denver, CO 80045, United States.
- 2 Commonwealth Informatics Inc, Waltham, MA 02451, United States.
- 3 Health Data Compass, University of Colorado Anschutz Medical Campus, Denver, CO 80045, United States.
- 4 National Association of Chronic Disease Directors (NACDD), Decatur, GA 30030, United States.
- 5 National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention (CDC), Atlanta, GA 30333, United States.
- 6 Kraushold Consulting, Denver, CO 80120, United States.
- 7 Public Health Informatics Institute, Decatur, GA 30030, United States.
- 8 Department of Medicine, University of Colorado Anschutz Medical Campus, Denver, CO 80045, United States.



## Abstract

**Objectives:** The Multi-State EHR-Based Network for Disease Surveillance (MENDS) is a population-based chronic disease surveillance distributed data network that uses institution-specific extraction-transformation-load (ETL) routines. MENDS-on-FHIR examined using Health Language Seven's Fast Healthcare Interoperability Resources (HL7® FHIR®) and US Core Implementation Guide (US Core IG) compliant resources derived from the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) to create a standards-based ETL pipeline.

**Materials and methods:** The input data source was a research data warehouse containing clinical and administrative data in OMOP CDM Version 5.3 format. OMOP-to-FHIR transformations, using a unique JavaScript Object Notation (JSON)-to-JSON transformation language called Whistle, created FHIR R4 V4.0.1/US Core IG V4.0.0 conformant resources that were stored in a local FHIR server. A REST-based Bulk FHIR \$export request extracted FHIR resources to populate a local MENDS database.

**Results:** Eleven OMOP tables were used to create 10 FHIR/US Core compliant resource types. A total of 1.13 trillion resources were extracted and inserted into the MENDS repository. A very low rate of non-compliant resources was observed.

**Discussion:** OMOP-to-FHIR transformation results passed validation with less than a 1% non-compliance rate. These standards-compliant FHIR resources provided standardized data elements required by the MENDS surveillance use case. The Bulk FHIR application programming interface (API) enabled population-level data exchange using interoperable FHIR resources. The OMOP-to-FHIR transformation pipeline creates a FHIR interface for accessing OMOP data.

**Conclusion:** MENDS-on-FHIR successfully replaced custom ETL with standards-based interoperable FHIR resources using Bulk FHIR. The OMOP-to-FHIR transformations provide an alternative mechanism for sharing OMOP data.

## 要旨

**目的:** Multi-State EHR-Based Network for Disease Surveillance (MENDS) は、施設固有の抽出-変換-ロード (ETL) ルーチンを使用する集団ベースの慢性疾患サーベイランス分散データネットワークである。MENDS-on-FHIRでは、HL7® FHIR® と US Core Implementation Guide (US Core IG) に準拠したリソースを使用して、OMOPのCDMから標準ベースのETLパイプラインを作成することを検討した。

**材料と方法:** 入力データソースは、OMOP CDM Version 5.3形式の臨床データおよび管理データを含む研究データウェアハウスであった。Whistleと呼ばれる独自のJavaScript Object Notation (JSON) -to-JSON変換言語を使用したOMOP-to-FHIR変換により、FHIR R4 V4.0.1/US Core IG V4.0.0適合リソースが作成され、ローカルのFHIRサーバーに保存された。RESTベースのBulk FHIR \$exportリクエストは、ローカルのMENDSデータベースに入力するFHIRリソースを抽出した。

**結果:** 10のFHIR/US Core準拠リソースタイプを作成するために11のOMOPテーブルが使用された。合計1兆1300億のリソースが抽出され、MENDSリポジトリに挿入された。非準拠リソースの割合は非常に低かった。

**考察:** OMOPからFHIRへの変換結果は、1%未満の非準拠率で検証に合格しました。これらの標準準拠のFHIRリソースは、MENDSサーベイランスのユースケースで必要とされる標準化されたデータ要素を提供した。一括 FHIR アプリケーション・プログラミング・インターフェース (API) は、相互運用可能な FHIR リソースを使用した集団レベルのデータ交換を可能にした。OMOP-to-FHIR変換パイプラインは、OMOPデータにアクセスするためのFHIRインターフェースを作成する。

**結論:** MENDS-on-FHIRは、Bulk FHIRを使用して、カスタムETLを標準ベースの相互運用可能なFHIRリソースに置き換えることに成功した。OMOP-to-FHIR変換は、OMOPデータを共有するための代替メカニズムを提供する。



## 組織内規制遵守のための 文書テンプレートを作ったよ

➤ [Front Med \(Lausanne\)](#). 2024 May 16;11:1378866. doi: 10.3389/fmed.2024.1378866. eCollection 2024.

## OHDSI-compliance: a set of document templates facilitating the implementation and operation of a software stack for real-world evidence generation

Felix N Wirth <sup>1</sup>, Hammam Abu Attieh <sup>1</sup>, Fabian Prasser <sup>1</sup>

Affiliations — collapse

### Affiliation

<sup>1</sup> Berlin Institute of Health at Charité - Universitätsmedizin Berlin, Center of Health Data Science, Berlin, Germany.

PMID: 38818399 PMCID: [PMC11137233](#) DOI: [10.3389/fmed.2024.1378866](#)

## Abstract

**Introduction:** The open-source software offered by the Observational Health Data Science and Informatics (OHDSI) collective, including the OMOP-CDM, serves as a major backbone for many real-world evidence networks and distributed health data analytics platforms. While container technology has significantly simplified deployments from a technical perspective, regulatory compliance can remain a major hurdle for the setup and operation of such platforms. In this paper, we present OHDSI-Compliance, a comprehensive set of document templates designed to streamline the data protection and information security-related documentation and coordination efforts required to establish OHDSI installations.

**Methods:** To decide on a set of relevant document templates, we first analyzed the legal requirements and associated guidelines with a focus on the General Data Protection Regulation (GDPR). Moreover, we analyzed the software architecture of a typical OHDSI stack and related its components to the different general types of concepts and documentation identified. Then, we created those documents for a prototypical OHDSI installation, based on the so-called Broadsea package, following relevant guidelines from Germany. Finally, we generalized the documents by introducing placeholders and options at places where individual institution-specific content will be needed.

**Results:** We present four documents: (1) a record of processing activities, (2) an information security concept, (3) an authorization concept, as well as (4) an operational concept covering the technical details of maintaining the stack. The documents are publicly available under a permissive license.

**Discussion:** To the best of our knowledge, there are no other publicly available sets of documents designed to simplify the compliance process for OHDSI deployments. While our documents provide a comprehensive starting point, local specifics need to be added, and, due to the heterogeneity of legal requirements in different countries, further adoptions might be necessary.

## 要旨

**はじめに:** OMOP-CDMを含む OHDSI集団が提供するオープンソースソフトウェアは、多くの実世界のエビデンスネットワークや分散型健康データ分析プラットフォームの主要なバックボーンとして機能している。コンテナ技術は技術的な観点から展開を大幅に簡素化したが、規制遵守はこのようなプラットフォームのセットアップと運用にとって大きなハードルであり続ける。この論文では、OHDSIの設置に必要なデータ保護と情報セキュリティ関連の文書化と調整作業を合理化するために設計された包括的な文書テンプレートセットであるOHDSI-Complianceを紹介する。

**方法:** 関連する一連の文書テンプレートを決定するために、まず一般データ保護規則（GDPR）を中心に法的要件と関連ガイドラインを分析した。さらに、典型的なOHDSIスタックのソフトウェア・アーキテクチャを分析し、そのコンポーネントを、特定されたさまざまな一般的タイプの概念と文書に関連付けた。次に、いわゆるBroadseaパッケージに基づいて、ドイツの関連ガイドラインに従ったOHDSIのプロトタイプの実装について、これらの文書を作成しました。最後に、個々の機関固有の内容が必要とされる箇所にプレースホルダーやオプションを導入することで、文書を一般化した。

**結果:** 4つの文書を紹介する：(1)処理活動の記録、(2)情報セキュリティの概念、(3)認可の概念、および(4)スタック維持の技術的詳細をカバーする運用の概念である。これらの文書は、寛容なライセンスの下で一般に公開されている。

**議論:** 我々の知る限りでは、OHDSI 導入のためのコンプライアンス・プロセスを簡素化するように設計された、一般に入手可能な文書群は他にない。我々の文書が包括的な出発点を提供する一方で、地域特有のものを追加する必要があり、また国によって法的要件が異なるため、さらなる採用が必要かもしれない。



# 6月の OHDSI Global

## ●Global Community Call テーマ

- June 11 Europe Symposium Review  
+ OHDSI Evidence Network Update
- June 18 Application of Large Language Models in the  
Evidence Generation Process
- June 25 Recent OHDSI Publications

## ●APAC Call テーマ

- June 6 Scientific Forum  
Demonstration of OHDSI ETL Tools
- June 20 Community Call  
Regional Chapter Mid-Year Updates



# 今後のOHDSI イベント

- OHDSI US/Global

The 10<sup>th</sup> 2024 OHDSI Global Symposium

会期 10/22-24 @ニュージャージー

- OHDSI APAC

2024 OHDSI APAC Symposium

会期 12/4-8 (Main 5-6) @シンガポール

2日間はマリーナベイ・サンズで開催されます。





## OHDSI APAC 各地域の状況

- オーストラリア
- 日本
- 韓国
- 台湾
- シンガポール