

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 #57

イブニングカンファレンス(第57回)

2024.8.29



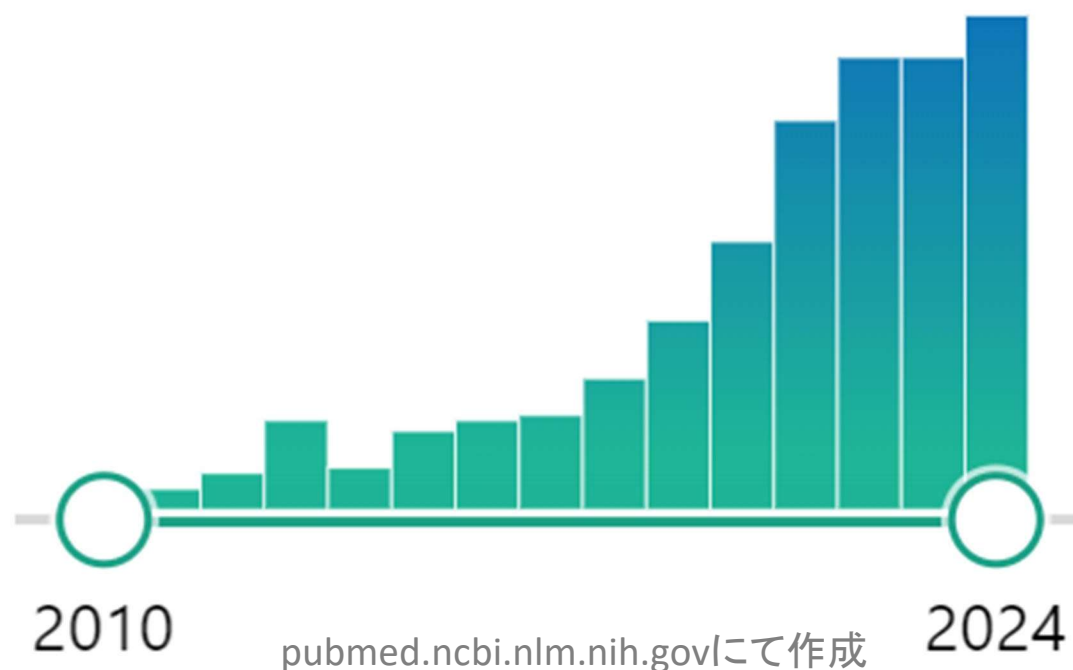
本日の内容

- OHDSI関連論文紹介
- OHDSI Global の話題から
- European Symposium紹介



OHDSI関連論文

Pubmedで“OHDSI or OMOP”を検索



全期間累計：7月447本→8月472本

- 検索に漏れているものがあるため、実際は累計600本を超えている。
- 年間では約100本ペース。

1. [KG-LIME: predicting individualized risk of adverse drug events for multiple sclerosis disease-modifying therapy. J Am Med Inform Assoc. 2024 Aug 1;31\(8\):1693-1703. doi: 10.1093/jamia/ocae155. PMID: 38964369](#)
2. [Advancing Medical Imaging Research Through Standardization: The Path to Rapid Development, Rigorous Validation, and Robust Reproducibility. Invest Radiol. 2024 Jul 11. doi: 10.1097/RLI.0000000000001106. Online ahead of print. PMID: 38985896](#)
3. [Integrating Clinical Data and Medical Imaging in Lung Cancer: Feasibility Study Using the Observational Medical Outcomes Partnership Common Data Model Extension. JMIR Med Inform. 2024 Jul 12;12:e59187. doi: 10.2196/59187. PMID: 38996330 Free article.](#)
4. [alofus: an R package to facilitate use of the All of Us Researcher Workbench. J Am Med Inform Assoc. 2024 Jul 24;ocae198. doi: 10.1093/jamia/ocae198. Online ahead of print. PMID: 39043402](#)
5. [Building Bridges for Federated Learning in Healthcare: Review on Approaches for Common Data Model Development. Stud Health Technol Inform. 2024 Jul 24;315:711-712. doi: 10.3233/SHTI240292. MID: 39049393 Review.](#)
6. [Creating a health informatics data resource for hearing health research. BMC Med Inform Decis Mak. 2024 Jul 29;24\(1\):209. doi: 10.1186/s12911-024-02589-x. PMID: 39075459](#)
7. [Making Metadata Machine-Readable as the First Step to Providing Findable, Accessible, Interoperable, and Reusable Population Health Data: Framework Development and Implementation Study. Online J Public Health Inform. 2024 Aug 1;16:e56237. doi: 10.2196/56237. PMID: 39088253 Free.](#)
8. [Transforming Primary Care Data Into the Observational Medical Outcomes Partnership Common Data Model: Development and Usability Study. JMIR Med Inform. 2024 Aug 13;12:e49542. doi: 10.2196/49542. PMID: 39140273.](#)
9. [How to customize common data models for rare diseases: an OMOP-based implementation and lessons learned. Orphanet J Rare Dis. 2024 Aug 14;19\(1\):298. doi: 10.1186/s13023-024-03312-9. PMID: 39143600.](#)
10. [Harmonizing Norwegian registries onto OMOP common data model: Mapping challenges and opportunities for pregnancy and COVID-19 research. Int J Med Inform. 2024 Aug 14;191:105602. doi: 10.1016/j.ijmedinf.2024.105602. Online ahead of print. PMID: 39153282.](#)
11. [Trends in incidence, prevalence, and survival of breast cancer in the United Kingdom from 2000 to 2021. Sci Rep. 2024 Aug 17;14\(1\):19069. doi: 10.1038/s41598-024-69006-1. PMID: 39153995.](#)
12. [Coverage of physical therapy assessments in the Observational Medical Outcomes Partnership Model common data model. Appl Clin Inform. 2024 Aug 22. doi: 10.1055/a-2401-3688. Online ahead of print. PMID: 39174009](#)
13. [Synthetic Data Generation in Hematology - Paving the Way for OMOP and FHIR Integration. Stud Health Technol Inform. 2024 Aug 22;316:1472-1476. doi: 10.3233/SHTI240692. PMID: 39176482](#)
14. [Towards Efficient Resource Utilization Forecasting in Acute Heat Events Using OMOP. Stud Health Technol Inform. 2024 Aug 22;316:1555-1559. doi: 10.3233/SHTI240715. PMID: 39176503](#)
15. [How to Accurately Detect Renal Replacement Therapy Weaning in Intensive Care: Data Quality and Standardization Considerations for the OMOP Common Data Model. Stud Health Technol Inform. 2024 Aug 22;316:1584-1588. doi: 10.3233/SHTI240724. PMID: 39176511](#)
16. [Intensive Care Quality Indicators Dashboard Using Observational Medical Outcomes Partnership Common Data Model. Stud Health Technol Inform. 2024 Aug 22;316:1605-1606. doi: 10.3233/SHTI240730. PMID: 39176517](#)
17. [A Chatbot-Like Platform to Enhance the Discovery of OMOP CDM Databases. Stud Health Technol Inform. 2024 Aug 22;316:1689-1693. doi: 10.3233/SHTI240748. PMID: 39176535](#)
18. [Synergies Among Health Data Projects with Cancer Use Cases Based on Health Standards. Stud Health Technol Inform. 2024 Aug 22;316:1292-1296. doi: 10.3233/SHTI240649. PMID: 39176618](#)
19. [Real-World Data in Greece: Mapping the Papageorgiou General Hospital Data to the OMOP Common Data Model. Stud Health Technol Inform. 2024 Aug 22;316:1324-1325. doi: 10.3233/SHTI240656. PMID: 39176625](#)
20. [From Syntactic to Semantic Interoperability Using a Hyperontology in the Oncology Domain. Stud Health Technol Inform. 2024 Aug 22;316:1385-1389. doi: 10.3233/SHTI240670. PMID: 39176639](#)
21. [Mapping OMOP-CDM to RDF: Bringing Real-World-Data to the Semantic Web Realm. Stud Health Technol Inform. 2024 Aug 22;316:1406-1410. doi: 10.3233/SHTI240674. PMID: 39176643](#)
22. [Enhancing Healthcare Informatics: Integrating Category Theory Reasoning into OMOP-CDM Ontology Model. Stud Health Technol Inform. 2024 Aug 22;316:1427-1431. doi: 10.3233/SHTI240680. PMID: 39176649](#)
23. [OntoBridge Versus Traditional ETL: Enhancing Data Standardization into CDM Formats Using Ontologies Within the DATOS-CAT Project. Stud Health Technol Inform. 2024 Aug 22;316:1432-1436. doi: 10.3233/SHTI240681. PMID: 39176650](#)
24. [OMOP-on-FHIR: A FHIR Server Development to Facilitate Data Interaction with the OMOP-CDM and FHIR for PGHD. Stud Health Technol Inform. 2024 Aug 22;316:157-158. doi: 10.3233/SHTI240367. PMID: 39176696](#)
25. [DATOS-CAT: OMOP-Common Data Model for the Standardization, Integration and Analysis of Population-Based Biomedical Data in Catalonia. Stud Health Technol Inform. 2024 Aug 22;316:200-201. doi: 10.3233/SHTI240378. PMID: 39176707](#)
26. [Improving Healthcare Quality with a LHS: From Patient-Generated Health Data to Evidence-Based Recommendations. Stud Health Technol Inform. 2024 Aug 22;316:230-234. doi: 10.3233/SHTI240387. PMID: 39176716](#)
27. [Dementia and Alzheimer's Disease Prevalence in Bulgaria During 2018: Nationally Representative Study. Stud Health Technol Inform. 2024 Aug 22;316:262-266. doi: 10.3233/SHTI240394. PMID: 39176723](#)
28. [Converting Entity-Attribute-Value Data Sources to OMOP's CDM: Lessons Learned. Stud Health Technol Inform. 2024 Aug 22;316:356-357. doi: 10.3233/SHTI240419. PMID: 39176749](#)
29. [Application of a Data Quality Framework to Ductal Carcinoma In Situ Using Electronic Health Record Data From the All of Us Research Program. JCO Clin Cancer Inform. 2024 Aug;8:e2400052. doi: 10.1200/CCI.24.00052. PMID: 39178364](#)



> [Invest Radiol.](#) 2024 Jul 11. doi: 10.1097/RLI.0000000000001106. Online ahead of print.

Advancing Medical Imaging Research Through Standardization: The Path to Rapid Development, Rigorous Validation, and Robust Reproducibility

Kyulee Jeon ¹, Woo Yeon Park, Charles E Kahn Jr, Paul Nagy, Seng Chan You, Soon Ho Yoon

PMID: 38985896

Affiliation

- ¹ From the Department of Biomedical Systems Informatics, Yonsei University, Seoul, South Korea (K.J., S.C.Y.); Institution for Innovation in Digital Healthcare, Yonsei University, Seoul, South Korea (K.J., S.C.Y.); Biomedical Informatics and Data Science, Johns Hopkins University, Baltimore, MD (W.Y.P., P.N.); Department of Radiology, University of Pennsylvania, Philadelphia, PA (C.E.K.); and Department of Radiology, Seoul National University College of Medicine, Seoul National University Hospital, Seoul, South Korea (S.H.Y.).

Abstract

Artificial intelligence (AI) has made significant advances in radiology. Nonetheless, challenges in AI development, validation, and reproducibility persist, primarily due to the lack of high-quality, large-scale, standardized data across the world. Addressing these challenges requires comprehensive standardization of medical imaging data and seamless integration with structured medical data. Developed by the Observational Health Data Sciences and Informatics community, the OMOP Common Data Model enables large-scale international collaborations with structured medical data. It ensures syntactic and semantic interoperability, while supporting the privacy-protected distribution of research across borders. The recently proposed Medical Imaging Common Data Model is designed to encompass all DICOM-formatted medical imaging data and integrate imaging-derived features with clinical data, ensuring their provenance. The harmonization of medical imaging data and its seamless integration with structured clinical data at a global scale will pave the way for advanced AI research in radiology. This standardization will enable federated learning, ensuring privacy-preserving collaboration across institutions and promoting equitable AI through the inclusion of diverse patient populations. Moreover, it will facilitate the development of foundation models trained on large-scale, multimodal datasets, serving as powerful starting points for specialized AI applications. Objective and transparent algorithm validation on a standardized data infrastructure will enhance reproducibility and interoperability of AI systems, driving innovation and reliability in clinical applications.

概要

人工知能（AI）は放射線医学において著しい進歩を遂げている。それにもかかわらず、AIの開発、検証、再現性には課題が残っており、その主な原因は、高品質で大規模な標準化されたデータが世界的に不足していることである。これらの課題に対処するには、医用画像データの包括的な標準化と、構造化された医療データとのシームレスな統合が必要である。OHDSIコミュニティによって開発されたOMOP共通データモデルは、構造化された医療データとの大規模な国際協力を可能にする。OMOP共通データモデルによって、構造化された医療データを用いた大規模な国際共同研究が可能になり、構文のおよび意味的な相互運用性が確保されるとともに、プライバシーが保護された国境を越えた研究の流通がサポートされる。

最近提案された医用画像共通データモデルは、すべてのDICOMフォーマットの医用画像データを包含し、画像に由来する機能を臨床データと統合し、それらの出所を保証するように設計されている。医用画像データの調和と、世界規模での構造化された臨床データとのシームレスな統合は、放射線医学における高度なAI研究に道を開くでしょう。この標準化により、連携学習が可能になり、医療機関を超えたプライバシー保護された連携が確保され、多様な患者集団を取り込むことで公平なAIが促進される。さらに、大規模なマルチモーダルデータセットで訓練された基礎モデルの開発を促進し、専門的なAIアプリケーションの強力な出発点となる。標準化されたデータ基盤上での客観的で透明性のあるアルゴリズム検証は、AIシステムの再現性と相互運用性を高め、臨床応用における革新と信頼性を促進する。



Yet AnotherなOMOP CDM連携の画像CDM

> JMIR Med Inform. 2024 Jul 12:12:e59187. doi: 10.2196/59187.

Integrating Clinical Data and Medical Imaging in Lung Cancer: Feasibility Study Using the Observational Medical Outcomes Partnership Common Data Model Extension

PMID: 38996330

Hyerim Ji ^{1 2}, Seok Kim ¹, Leonard Sunwoo ³, Sowon Jang ³, Ho-Young Lee ^{1 4}, Sooyoung Yoo ¹

Affiliations

- 1 Office of eHealth Research and Business, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea.
- 2 Department of Health Science and Technology, Graduate School of Convergence Science and Technology, Seoul National University, Seoul, Republic of Korea.
- 3 Department of Radiology, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea.
- 4 Department of Nuclear Medicine, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea.

Abstract

Background: Digital transformation, particularly the integration of medical imaging with clinical data, is vital in personalized medicine. The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) standardizes health data. However, integrating medical imaging remains a challenge.

Objective: This study proposes a method for combining medical imaging data with the OMOP CDM to improve multimodal research.

Methods: Our approach included the analysis and selection of digital imaging and communications in medicine header tags, validation of data formats, and alignment according to the OMOP CDM framework. The Fast Healthcare Interoperability Resources ImagingStudy profile guided our consistency in column naming and definitions. Imaging Common Data Model (I-CDM), constructed using the entity-attribute-value model, facilitates scalable and efficient medical imaging data management. For patients with lung cancer diagnosed between 2010 and 2017, we introduced 4 new tables-IMAGING_STUDY, IMAGING_SERIES, IMAGING_ANNOTATION, and FILEPATH-to standardize various imaging-related data and link to clinical data.

Results: This framework underscores the effectiveness of I-CDM in enhancing our understanding of lung cancer diagnostics and treatment strategies. The implementation of the I-CDM tables enabled the structured organization of a comprehensive data set, including 282,098 IMAGING_STUDY, 5,674,425 IMAGING_SERIES, and 48,536 IMAGING_ANNOTATION records, illustrating the extensive scope and depth of the approach. A scenario-based analysis using actual data from patients with lung cancer underscored the feasibility of our approach. A data quality check applying 44 specific rules confirmed the high integrity of the constructed data set, with all checks successfully passed, underscoring the reliability of our findings.

Conclusions: These findings indicate that I-CDM can improve the integration and analysis of medical imaging and clinical data. By addressing the challenges in data standardization and management, our approach contributes toward enhancing diagnostics and treatment strategies. Future research should expand the application of I-CDM to diverse disease populations and explore its wide-ranging utility for medical conditions.

概要

背景: デジタルトランスフォーメーション、特に医用画像と臨床データの統合は個別化医療に不可欠である。OMOP共通データモデル（CDM）は医療データを標準化するものである。しかし、医用画像の統合は依然として課題である。

目的: 本研究では、マルチモーダル研究を改善するために、医用画像データをOMOP CDMと統合する方法を提案する。

方法: 我々のアプローチには、デジタル画像と医療通信のヘッダタグの分析と選択、データフォーマットの検証、OMOP CDMフレームワークによるアライメントが含まれる。Fast Healthcare Interoperability Resources ImagingStudyプロファイルは、カラムの命名と定義の一貫性を導くものであった。Imaging Common Data Model（**I-CDM**）は、実体-属性-値モデルを用いて構築され、スケーラブルで効率的な医用画像データ管理を促進する。2010年から2017年の間に診断された肺がん患者について、4つの新しいテーブル-IMAGING_STUDY、IMAGING_SERIES、IMAGING_ANNOTATION、FILEPATH-を導入し、様々な画像関連データを標準化し、臨床データとリンクさせた。

結果: このフレームワークは、肺癌の診断と治療戦略の理解を深める上でI-CDMが有効であることを強調している。I-CDMテーブルの実装により、282,098件のIMAGING_STUDY、5,674,425件のIMAGING_SERIES、48,536件のIMAGING_ANNOTATIONレコードを含む包括的なデータセットの構造的整理が可能となり、このアプローチの広範さと深さが示された。肺がん患者の実際のデータを用いたシナリオに基づく分析により、本アプローチの実現可能性が強調された。44の特定のルールを適用したデータ品質チェックにより、構築されたデータセットの高い完全性が確認され、すべてのチェックが成功裏に通過し、我々の知見の信頼性が強調された。

結論: これらの結果は、I-CDMが医用画像と臨床データの統合と解析を改善できることを示している。データの標準化と管理における課題に対処することで、我々のアプローチは診断と治療戦略の強化に貢献する。今後の研究では、I-CDMの多様な疾患集団への適用を拡大し、医学的状态に対するI-CDMの広範な有用性を探る必要がある。



All of US 利用のためのRパッケージ

> [J Am Med Inform Assoc. 2024 Jul 24;ocae198. doi: 10.1093/jamia/ocae198. Online ahead of print.](#)

allofus: an R package to facilitate use of the All of Us Researcher Workbench

Louisa H Smith^{1 2}, Robert Cavanaugh^{1 2}

Affiliations — collapse

Affiliations

- 1 Roux Institute, Northeastern University, Portland, ME 04101, United States.
- 2 Department of Public Health and Health Sciences, Bouvé College of Health Sciences, Northeastern University, Boston, MA 02115, United States.

PMID: 39043402 DOI: [10.1093/jamia/ocae198](#)

Abstract

Objectives: Despite easy-to-use tools like the Cohort Builder, using All of Us Research Program data for complex research questions requires a relatively high level of technical expertise. We aimed to increase research and training capacity and reduce barriers to entry for the All of Us community through an R package, *allofus*. In this article, we describe functions that address common challenges we encountered while working with All of Us Research Program data, and we demonstrate this functionality with an example of creating a cohort of All of Us participants by synthesizing electronic health record and survey data with time dependencies.

Target audience: All of Us Research Program data are widely available to health researchers. The *allofus* R package is aimed at a wide range of researchers who wish to conduct complex analyses using best practices for reproducibility and transparency, and who have a range of experience using R. Because the All of Us data are transformed into the Observational Medical Outcomes Partnership Common Data Model (OMOP CDM), researchers familiar with existing OMOP CDM tools or who wish to conduct network studies in conjunction with other OMOP CDM data will also find value in the package.

Scope: We developed an initial set of functions that solve problems we experienced across survey and electronic health record data in our own research and in mentoring student projects. The package will continue to grow and develop with the All of Us Research Program. The *allofus* R package can help build community research capacity by increasing access to the All of Us Research Program data, the efficiency of its use, and the rigor and reproducibility of the resulting research.

概要

目的: Cohort Builderのような使いやすいツールがあるにもかかわらず、All of Us Research Programのデータを複雑な研究課題に使用するには、比較的高度な技術的専門知識が必要である。我々はRパッケージであるallofusを通して、All of Usコミュニティの研究・トレーニング能力を向上させ、参入障壁を軽減することを目的とした。この記事では、All of Us Research Programのデータを扱う際に遭遇した一般的な課題に対処する機能について説明し、時間依存性を持つ電子カルテと調査データを統合してAll of Us参加者のコホートを作成する例でこの機能を実証する。

対象読者: All of Us Research Programのデータは、健康研究者が広く利用できる。All of UsのデータはOMOP CDMに変換されるため、既存のOMOP CDMツールに精通している研究者や、他のOMOP CDMデータと組み合わせてネットワーク研究を実施したい研究者も、このパッケージの価値を見出すことができる。

スコープ: 私たちは、私たち自身の研究や学生プロジェクトの指導において、調査データと電子カルテデータ間で経験した問題を解決するための初期機能セットを開発した。このパッケージは、All of Us 研究プログラムと共に成長し、発展し続けます。allofus R パッケージは、All of Us Research Program のデータへのアクセス、その利用の効率化、そして得られた研究の厳密性と再現性を向上させることで、コミュニティの研究能力の構築に貢献することができます。



CDM開発についてロードマップがほしい(?)

Review > Stud Health Technol Inform. 2024 Jul 24:315:711-712. doi: 10.3233/SHTI240292.

Building Bridges for Federated Learning in Healthcare: Review on Approaches for Common Data Model Development

Hanna von Gerich¹, Taridzo Chomutare², Laura-Maria Peltonen^{1 3}

Affiliations – collapse

Affiliations

- 1 Department of Nursing Science, University of Turku, Finland.
- 2 Norwegian Centre for E-health Research, Norway.
- 3 Turku University Hospital, Finland.

PMID: 39049393 DOI: [10.3233/SHTI240292](https://doi.org/10.3233/SHTI240292)

Abstract

Common data models provide a standardized way to represent data used in federated learning tasks. The aim of this review was to explore the development and use of common data models to harmonize electronic health record data in health research. The data search yielded 724 records, of which 19 were included for this study. None of the research focused on nursing specific topics. All studies either utilized the Observational Medical Outcomes Partnership (OMOP) common data model, or developed a model partly based on the OMOP. A roadmap to guide research for the development of common data models for federated learning are warranted.

概要

共通データモデルは、連合学習タスクで使用されるデータを表現するための標準化された方法を提供する。このレビューの目的は、健康研究において電子カルテデータを調和させるための共通データモデルの開発と利用について調査することであった。データ検索により724件の記録が得られ、そのうち19件が本研究の対象となった。どの研究も看護に特化したテーマではなかった。すべての研究は、OMOP共通データモデルを利用するか、OMOPに部分的に基づいたモデルを開発していた。連携学習のための共通データモデル開発のための研究の指針となるロードマップが必要である。



聴覚障害研究のためのOMOP基盤

> BMC Med Inform Decis Mak. 2024 Jul 29;24(1):209. doi: 10.1186/s12911-024-02589-x.

Creating a health informatics data resource for hearing health research

PMID: 39075459

Nishchay Mehta^{1 2}, Baptiste Briot Ribeyre^{1 3}, Lilia Dimitrov^{1 2}, Louise J English^{1 3}, Colleen Ewart⁴, Antje Heinrich^{5 6}, Nikhil Joshi^{1 2}, Kevin J Munro^{5 6}, Gail Roadknight⁷, Luis Romao^{1 3}, Anne Gm Schilder^{1 2}, Ruth V Spriggs^{8 9}, Ruth Norris^{5 10}, Talisa Ross^{11 12 13}, George Tilston^{5 10}

Affiliations

- 1 NIHR University College London Hospitals Biomedical Research Centre, London, UK.
- 2 Royal National ENT Hospital, UCLH Trust, London, UK.
- 3 Institute of Health Informatics, Faculty of Population Health Sciences, University College London, London, UK.
- 4 NIHR Health Informatics Collaborative Hearing Health, Patient and Public Engagement Group, London, UK.
- 5 NIHR Manchester Biomedical Research Centre, Manchester, UK.
- 6 Manchester Centre for Audiology and Deafness (ManCAD), School of Health Sciences, The University of Manchester, Manchester, UK.
- 7 Oxford University Hospitals NHS Foundation Trust: Oxford, Oxfordshire, UK.
- 8 NIHR Biomedical Research Centre School of Medicine University of Nottingham Nottingham

Abstract

Background: The National Institute of Health and Social Care Research (NIHR) Health Informatics Collaborative (HIC) for Hearing Health has been established in the UK to curate routinely collected hearing health data to address research questions. This study defines priority research areas, outlines its aims, governance structure and demonstrates how hearing health data have been integrated into a common data model using pure tone audiometry (PTA) as a case study.

Methods: After identifying key research aims in hearing health, the governance structure for the NIHR HIC for Hearing Health is described. The Observational Medical Outcomes Partnership (OMOP) was chosen as our common data model to provide a case study example.

Results: The NIHR HIC Hearing Health theme have developed a data architecture outlining the flow of data from all of the various siloed electronic patient record systems to allow the effective linkage of data from electronic patient record systems to research systems. Using PTAs as an example, OMOPification of hearing health data successfully collated a rich breadth of datapoints across multiple centres.

Conclusion: This study identified priority research areas where routinely collected hearing health data could be useful. It demonstrates integration and standardisation of such data into a common data model from multiple centres. By describing the process of data sharing across the HIC, we hope to invite more centres to contribute and utilise data to address research questions in hearing health. This national initiative has the power to transform UK hearing research and hearing care using routinely collected clinical data.

概要

背景：英国では、国立保健社会医療研究機構（NIHR）の聴覚保健情報共同体（HIC）が設立され、日常的に収集される聴覚保健データを管理して研究課題に取り組んでいる。本研究では、優先研究分野を定義し、その目的とガバナンス構造を概説し、純音聴力測定（PTA）をケーススタディとして聴覚健康データがどのように共通のデータモデルに統合されたかを示す。

方法：聴覚保健における主要な研究目的を特定した後、NIHR HIC for Hearing Healthのガバナンス構造を説明する。OMOPを共通データモデルとして選び、ケーススタディの例とした。

結果：NIHR HIC Hearing Healthのテーマは、電子患者記録システムから研究システムへの効果的なデータ連携を可能にするために、様々なサイロ化された電子患者記録システムすべてからのデータの流れをアウトライン化したデータアーキテクチャを開発した。PTAを例にとると、聴覚健康データのOMOP化は、複数のセンターにわたる豊富なデータポイントの照合に成功した。

結論：本研究は、日常的に収集される聴覚健康データが有用な優先研究分野を特定した。また、このようなデータを複数のセンターから共通のデータモデルに統合し、標準化することを実証した。HIC全体でのデータ共有のプロセスを説明することで、より多くのセンターがデータを提供し、聴覚保健の研究課題に取り組むために活用することを期待している。この全国的なイニシアチブは、日常的に収集される臨床データを用いて、英国の聴覚研究と聴覚ケアを変革する力を持っている。



メタデータフレームワークを作った

> [Online J Public Health Inform.](#) 2024 Aug 1:16:e56237. doi: 10.2196/56237.

Making Metadata Machine-Readable as the First Step to Providing Findable, Accessible, Interoperable, and Reusable Population Health Data: Framework Development and Implementation Study

PMID: 39088253

Erratum in PMID: 39141912

David Amadi¹, Sylvia Kiwuwa-Muyingo², Tathagata Bhattacharjee¹, Amelia Taylor³, Agnes Kiragga², Michael Ochola², Chifundo Kanjala⁴, Arofan Gregory⁵, Keith Tomlin¹, Jim Todd¹, Jay Greenfield⁵

Affiliations

- 1 Department of Population Health, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom.
- 2 African Population and Health Research Center, Nairobi, Kenya.
- 3 Malawi University of Business and Applied Science, Blantyre, Malawi.
- 4 Independent researcher, Lilongwe, Malawi.
- 5 Committee on Data (CODATA), Paris, France.

Abstract

Background: Metadata describe and provide context for other data, playing a pivotal role in enabling findability, accessibility, interoperability, and reusability (FAIR) data principles. By providing comprehensive and machine-readable descriptions of digital resources, metadata empower both machines and human users to seamlessly discover, access, integrate, and reuse data or content across diverse platforms and applications. However, the limited accessibility and machine-interpretability of existing metadata for population health data hinder effective data discovery and reuse.

Objective: To address these challenges, we propose a comprehensive framework using standardized formats, vocabularies, and protocols to render population health data machine-readable, significantly enhancing their FAIRness and enabling seamless discovery, access, and integration across diverse platforms and research applications.

Methods: The framework implements a 3-stage approach. The first stage is Data Documentation Initiative (DDI) integration, which involves leveraging the DDI Codebook metadata and documentation of detailed information for data and associated assets, while ensuring transparency and comprehensiveness. The second stage is Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) standardization. In this stage, the data are harmonized and standardized into the OMOP CDM, facilitating unified analysis across heterogeneous data sets. The third stage involves the integration of Schema.org and JavaScript Object Notation for Linked Data (JSON-LD), in which machine-readable metadata are generated using Schema.org entities and embedded within the data using JSON-LD, boosting discoverability and comprehension for both machines and human users. We demonstrated the implementation of these 3 stages using the Integrated Disease Surveillance and Response (IDSR) data from Malawi and Kenya.

Results: The implementation of our framework significantly enhanced the FAIRness of population health data, resulting in improved discoverability through seamless integration with platforms such as Google Dataset Search. The adoption of standardized formats and protocols streamlined data accessibility and integration across various research environments, fostering collaboration and knowledge sharing. Additionally, the use of machine-interpretable metadata empowered researchers to efficiently reuse data for targeted analyses and insights, thereby maximizing the overall value of population health resources. The JSON-LD codes are accessible via a GitHub repository and the HTML code integrated with JSON-LD is available on the Implementation Network for Sharing Population Information from Research Entities website.

Conclusions: The adoption of machine-readable metadata standards is essential for ensuring the FAIRness of population health data. By embracing these standards, organizations can enhance diverse resource visibility, accessibility, and utility, leading to a broader impact, particularly in low- and middle-income countries. Machine-readable metadata can accelerate research, improve health care decision-making, and ultimately promote better health outcomes for populations worldwide.

概要

背景: メタデータは、他のデータを記述し、そのコンテキストを提供するものであり、ファインダビリティ、アクセシビリティ、インターオペラビリティ、リユーザビリティ（FAIR）のデータ原則を実現する上で極めて重要な役割を果たす。デジタルリソースの包括的で機械可読な記述を提供することで、メタデータは、機械と人間のユーザーの双方に、多様なプラットフォームやアプリケーション間でデータやコンテンツをシームレスに発見、アクセス、統合、再利用する力を与える。しかし、集団健康データに関する既存のメタデータのアクセシビリティと機械的解釈可能性は限られており、効果的なデータの発見と再利用を妨げている。

目的: これらの課題に対処するため、標準化されたフォーマット、語彙、プロトコルを用いた包括的なフレームワークを提案し、集団健康データを機械可読化することで、そのFAIR性を大幅に向上させ、多様なプラットフォームや研究アプリケーション間でのシームレスな発見、アクセス、統合を可能にする。

方法: このフレームワークは3段階のアプローチを実施する。**第一段階**はData Documentation Initiative (DDI)の統合で、DDI Codebookメタデータを活用し、透明性と包括性を確保しつつ、データと関連資産の詳細情報を文書化する。**第二段階**はOMOP共通データモデル（CDM）の標準化である。この段階では、データを OMOP CDM に整合化・標準化し、異種データセット間の統一的な解析を容易にする。**第三段階**では、Schema.orgとJavaScript Object Notation for Linked Data (JSON-LD)の統合が行われ、Schema.orgエンティティを使用して機械可読メタデータが生成され、JSON-LDを使用してデータ内に埋め込まれる。マラウイとケニアの統合疾病サーベイランス・レスポンス（IDSR）データを用いて、これら3段階の実装を実証した。

結果: 我々のフレームワークの実装により、集団保健データのFAIR性が大幅に向上し、Google Dataset Searchなどのプラットフォームとのシームレスな統合により、発見性が改善された。標準化されたフォーマットとプロトコルの採用により、様々な研究環境におけるデータアクセシビリティと統合が合理化され、コラボレーションと知識の共有が促進された。さらに、機械的に解釈可能なメタデータを使用することで、研究者は的を絞った分析や洞察のためにデータを効率的に再利用できるようになり、集団健康資源の全体的な価値が最大化された。JSON-LDコードはGitHubリポジトリからアクセスでき、JSON-LDと統合されたHTMLコードはImplementation Network for Sharing Population Information from Research Entitiesのウェブサイトから入手できる。

結論: 機械可読のメタデータ標準の採用は、集団健康データの公平性を確保するために不可欠である。これらの標準を採用することで、組織は多様なリソースの可視性、アクセシビリティ、有用性を高めることができ、特に低・中所得国において、より広範な影響につながる。機械可読なメタデータは、研究を加速し、医療の意思決定を改善し、最終的には世界中の人々の健康状態の改善を促進することができる。



プライマリーケアデータのOMOP化

› [JMIR Med Inform. 2024 Aug 13:12:e49542. doi: 10.2196/49542.](#)

Transforming Primary Care Data Into the Observational Medical Outcomes Partnership Common Data Model: Development and Usability Study

PMID: 39140273

Mathilde Fruchart¹, Paul Quindroit¹, Chloé Jacquemont², Jean-Baptiste Beuscart¹,
Matthieu Calafiore^{1 2}, Antoine Lamer^{1 3}

Affiliations

- 1 Univ Lille, CHU Lille, ULR 2694 - METRICS: Évaluation des Technologies de santé et des Pratiques médicales, 2 Place de Verdun, Lille, F-59000, France.
- 2 Département de Médecine Générale, University of Lille, Lille, France.
- 3 F2RSM Psy - Fédération régionale de recherche en psychiatrie et santé mentale Hauts-de-France, Saint-André-Lez-Lille, France.

Abstract

Background: Patient-monitoring software generates a large amount of data that can be reused for clinical audits and scientific research. The Observational Health Data Sciences and Informatics (OHDSI) consortium developed the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) to standardize electronic health record data and promote large-scale observational and longitudinal research.

Objective: This study aimed to transform primary care data into the OMOP CDM format.

Methods: We extracted primary care data from electronic health records at a multidisciplinary health center in Watrelos, France. We performed structural mapping between the design of our local primary care database and the OMOP CDM tables and fields. Local French vocabularies concepts were mapped to OHDSI standard vocabularies. To validate the implementation of primary care data into the OMOP CDM format, we applied a set of queries. A practical application was achieved through the development of a dashboard.

Results: Data from 18,395 patients were implemented into the OMOP CDM, corresponding to 592,226 consultations over a period of 20 years. A total of 18 OMOP CDM tables were implemented. A total of 17 local vocabularies were identified as being related to primary care and corresponded to patient characteristics (sex, location, year of birth, and race), units of measurement, biometric measures, laboratory test results, medical histories, and drug prescriptions. During semantic mapping, 10,221 primary care concepts were mapped to standard OHDSI concepts. Five queries were used to validate the OMOP CDM by comparing the results obtained after the completion of the transformations with the results obtained in the source software. Lastly, a prototype dashboard was developed to visualize the activity of the health center, the laboratory test results, and the drug prescription data.

Conclusions: Primary care data from a French health care facility have been implemented into the OMOP CDM format. Data concerning demographics, units, measurements, and primary care consultation steps were already available in OHDSI vocabularies. Laboratory test results and drug prescription data were mapped to available vocabularies and structured in the final model. A dashboard application provided health care professionals with feedback on their practice.

概要

背景: 患者モニタリングソフトウェアは、臨床監査や科学研究に再利用可能な大量のデータを生成する。OHDSI コンソーシアムは、電子カルテデータを標準化し、大規模な観察研究および縦断的研究を促進するために、OMOP 共通データモデル (CDM) を開発した。

目的: 本研究は、プライマリケアデータをOMOP CDM形式に変換することを目的とした。

方法: フランスのWattrelosにある集学的医療センターの電子カルテからプライマリケアデータを抽出した。地域のプライマリケアデータベースの設計とOMOP CDMのテーブルおよびフィールドの構造マッピングを行った。フランス語のローカル語彙の概念をOHDSI標準語彙にマッピングした。プライマリケアデータをOMOP CDMフォーマットに実装することを検証するために、一連のクエリーを適用した。ダッシュボードの開発を通じて、実用的なアプリケーションを実現した。

結果: 18,395人の患者のデータがOMOP CDMに実装され、20年間の592,226人の診察に対応した。合計18のOMOP CDMテーブルが実装された。プライマリ・ケアに関連する語彙として、患者の特徴（性別、居住地、出生年、人種）、測定単位、生体測定値、検査結果、病歴、薬剤処方に対応する17のローカル語彙が特定された。セマンティックマッピングでは、10,221のプライマリケア概念が標準的なOHDSI概念にマッピングされた。5つのクエリーを用いて、変換完了後に得られた結果をソースソフトウェアで得られた結果と比較し、OMOP CDMを検証した。最後に、ダッシュボードのプロトタイプを開発し、保健センターの活動、検査結果、薬剤処方データを可視化した。

結論: フランスの医療施設からのプライマリケアデータがOMOP CDMフォーマットに実装された。人口統計、単位、測定値、プライマリケアの診察ステップに関するデータは、すでにOHDSI語彙で利用可能であった。臨床検査結果と薬剤処方データは、利用可能な語彙にマッピングされ、最終モデルで構造化された。ダッシュボード・アプリケーションは、医療従事者の診療に関するフィードバックを提供した。



稀少疾患CDMの意義

> Orphanet J Rare Dis. 2024 Aug 14;19(1):298. doi: 10.1186/s13023-024-03312-9.

How to customize common data models for rare diseases: an OMOP-based implementation and lessons learned

PMID: 39143600

Najia Ahmadi¹, Michele Zoch², Oya Guengoeze³, Carlo Facchinello³, Antonia Mondorf³, Katharina Stratmann³, Khader Musleh³, Hans-Peter Erasmus³, Jana Tchertov², Richard Gebler², Jannik Schaaf⁴, Lena S Frischen⁵, Azadeh Nasirian⁶, Jiabin Dai², Elisa Henke², Douglas Tremblay⁷, Andrew Srisuwananukorn⁸, Martin Bornhäuser⁹, Christoph Röllig⁹, Jan-Niklas Eckardt^{9 10}, Jan Moritz Middeke^{9 10}, Markus Wolfien^{2 11}, Martin Sedlmayr²

Affiliations

- 1 Institute for Medical Informatics and Biometry, Carl Gustav Carus Faculty of Medicine, TUD Dresden University of Technology, Fetscherstraße 74, 01307, Dresden, Germany. najia.ahmadi@tu-dresden.de.
- 2 Institute for Medical Informatics and Biometry, Carl Gustav Carus Faculty of Medicine, TUD Dresden University of Technology, Fetscherstraße 74, 01307, Dresden, Germany.
- 3 Department of Internal Medicine I, University Hospital Frankfurt, Goethe University, Frankfurt, Germany.
- 4 Goethe University Frankfurt, University Hospital, Institute of Medical Informatics, Frankfurt, Germany.
- 5 University Hospital Frankfurt, Goethe University, Executive Department for Medical IT-Systems

Abstract

Background: Given the geographical sparsity of Rare Diseases (RDs), assembling a cohort is often a challenging task. Common data models (CDM) can harmonize disparate sources of data that can be the basis of decision support systems and artificial intelligence-based studies, leading to new insights in the field. This work is sought to support the design of large-scale multi-center studies for rare diseases.

Methods: In an interdisciplinary group, we derived a list of elements of RDs in three medical domains (endocrinology, gastroenterology, and pneumonology) according to specialist knowledge and clinical guidelines in an iterative process. We then defined a RDs data structure that matched all our data elements and built Extract, Transform, Load (ETL) processes to transfer the structure to a joint CDM. To ensure interoperability of our developed CDM and its subsequent usage for further RDs domains, we ultimately mapped it to Observational Medical Outcomes Partnership (OMOP) CDM. We then included a fourth domain, hematology, as a proof-of-concept and mapped an acute myeloid leukemia (AML) dataset to the developed CDM.

Results: We have developed an OMOP-based rare diseases common data model (RD-CDM) using data elements from the three domains (endocrinology, gastroenterology, and pneumonology) and tested the CDM using data from the hematology domain. The total study cohort included 61,697 patients. After aligning our modules with those of Medical Informatics Initiative (MII) Core Dataset (CDS) modules, we leveraged its ETL process. This facilitated the seamless transfer of demographic information, diagnoses, procedures, laboratory results, and medication modules from our RD-CDM to the OMOP. For the phenotypes and genotypes, we developed a second ETL process. We finally derived lessons learned for customizing our RD-CDM for different RDs.

Discussion: This work can serve as a blueprint for other domains as its modularized structure could be extended towards novel data types. An interdisciplinary group of stakeholders that are actively supporting the project's progress is necessary to reach a comprehensive CDM.

Conclusion: The customized data structure related to our RD-CDM can be used to perform multi-center studies to test data-driven hypotheses on a larger scale and take advantage of the analytical tools offered by the OHDSI community.

概要

背景: 希少疾病（RDs）の地理的な希少性を考慮すると、コホートの構築はしばしば困難な課題である。共通データモデル（CDM）は、意思決定支援システムや人工智能に基づく研究の基礎となる異種のデータソースを調和させることができ、この分野の新たな知見につながる。この研究は、希少疾患の大規模多施設共同研究のデザインを支援することを目的としている。

方法: 学際的なグループにおいて、3つの医療領域（内分泌学、消化器病学、肺臓病学）におけるRDの要素リストを、専門家の知識と臨床ガイドラインに従って反復プロセスで導き出した。次に、すべてのデータ要素に適合する RDs データ構造を定義し、その構造を共同 CDM に転送するための ETL（Extract, Transform, Load）プロセスを構築した。開発したCDMの相互運用性を確保し、その後さらにRDsドメインで使用するために、最終的にOMOP CDMにマッピングした。その後、概念実証として4つ目のドメインである血液学を追加し、急性骨髄性白血病（AML）データセットを開発したCDMにマッピングした。

結果: 3つのドメイン（内分泌学、消化器学、肺医学）のデータ要素を用いてOMOPベースの希少疾患共通データモデル（RD-CDM）を開発し、血液学ドメインのデータを用いてCDMをテストした。全研究コホートには61,697人の患者が含まれた。我々のモジュールをMedical Informatics Initiative (MII) Core Dataset (CDS)のモジュールと整合させた後、そのETLプロセスを活用した。これにより、RD-CDMからOMOPへの人口統計情報、診断、処置、検査結果、投薬モジュールのシームレスな転送が容易になった。表現型と遺伝子型については、2番目のETLプロセスを開発した。最後に、異なるRDのためにRD-CDMをカスタマイズするための教訓を得た。

考察: この研究は、モジュール化された構造を新しいデータタイプに拡張することが可能であるため、他のドメインのための青写真として役立つ。包括的なCDMに到達するためには、プロジェクトの進展を積極的に支援する学際的な利害関係者のグループが必要である。

結論: 我々のRD-CDMに関連するカスタマイズされたデータ構造は、データ駆動型の仮説をより大規模に検証するための多施設研究を実施し、OHDSIコミュニティが提供する分析ツールを活用するために使用することができる。



8月の OHDSI Global

- Global Community Call テーマ

July 30 Patient-Level Prediction

Aug. 6 Building the OHDSI Evidence Network, Session III -
OHDSI Vocabularies

Aug. 13 A Look Inside The OHDSI 2024 Global Symposium

Aug. 20 Building The OHDSI Evidence Network Sprint,
Final Session

- APAC Call テーマ

Aug. 1 Scientific Forum
Community-Wide ETL Project Kickoff

Aug.14 Scientific Forum
Community-Wide ETL Project Sprint 1 Review

Aug.15 Community Call
Evidence Network & Workgroup Updates



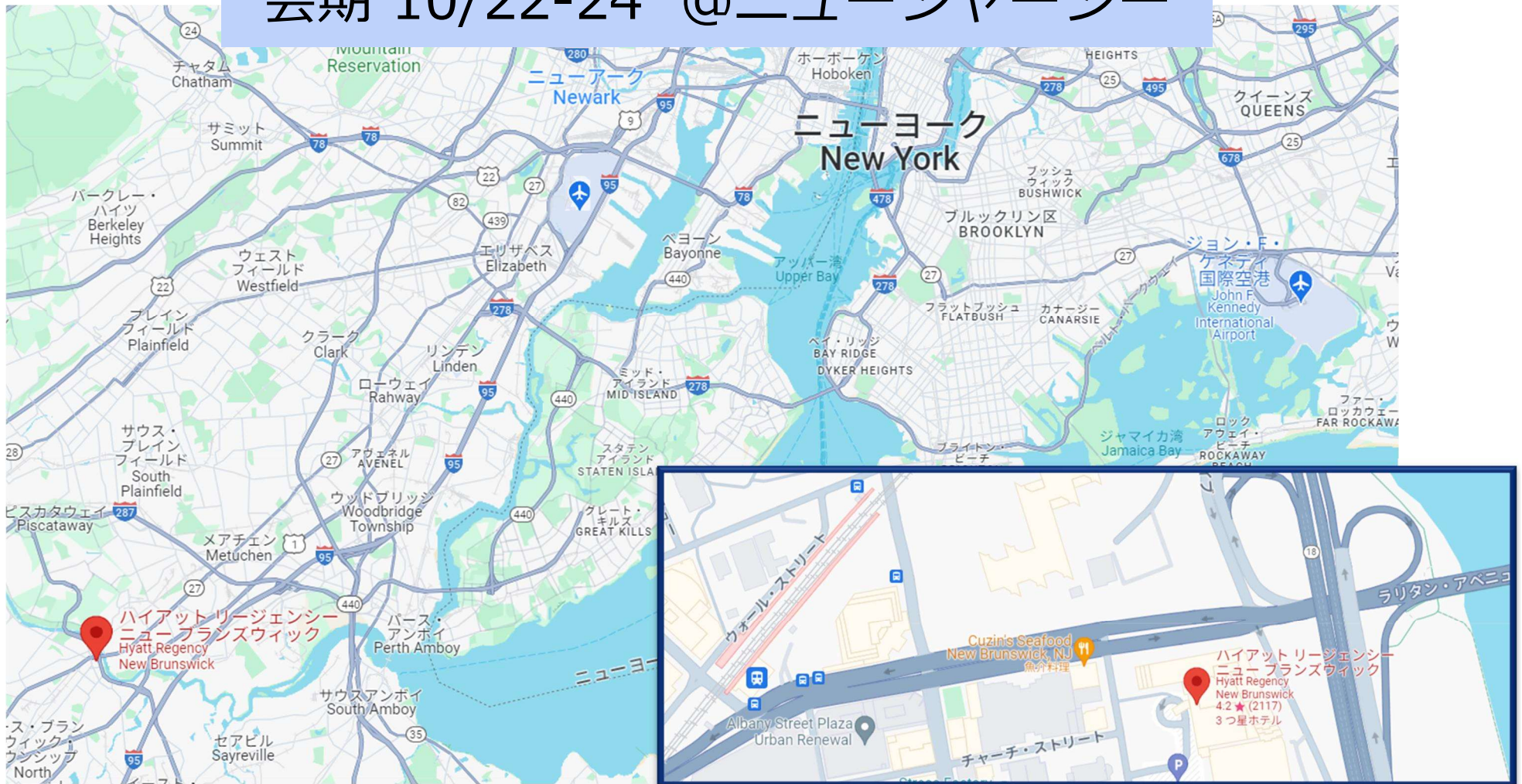
今後のOHDSI イベント

<https://ohdsi.org/ohdsi2024/>

2024 OHDSI Global Symposium

Oct. 22-24 • New Brunswick, N.J. • Hyatt Regency Hotel

会期 10/22-24 @ニュージャージー





今後のOHDSI イベント

<https://ohdsi.org/apac2024/>

2024 OHDSI APAC Symposium

December 4-8 • Marina Bay Sands & National University of Singapore (NUS)

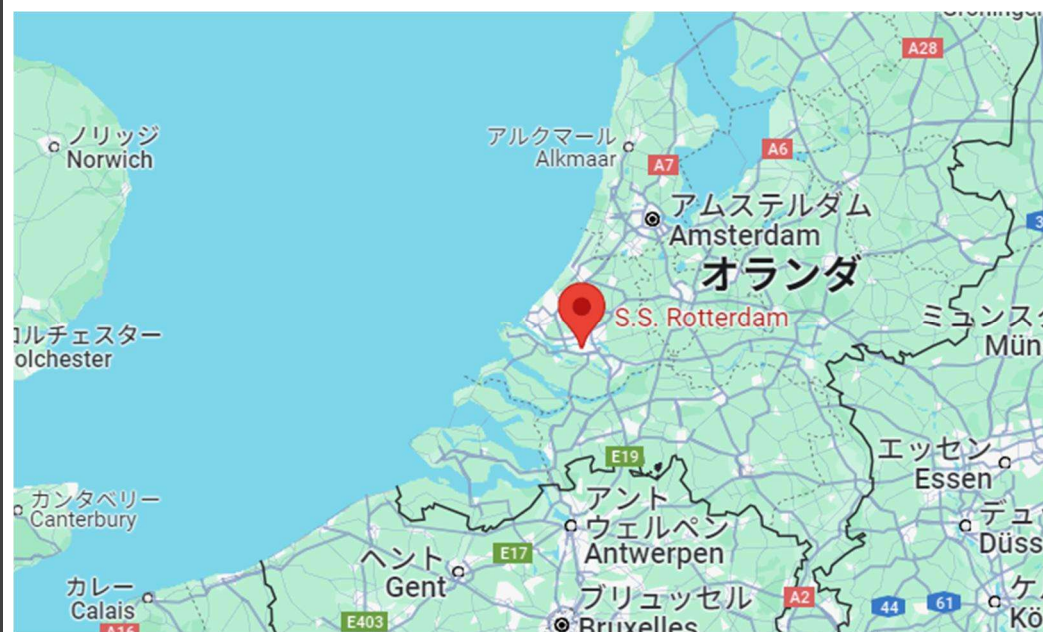
会期 12/4-8 (Main 5-6) @シンガポール
Main2日間はマリーナベイ・サンズで開催されます。





第5回 欧州シンポジウム
2024年6月1～3日
@ロッテルダム

<https://www.ohdsi-europe.org/index.php/symposium/45-archive-symposium-2024>



SESSION ONE

00:00 - Welcome to the European OHDSI Journey.

Peter Rijnbeek, PhD, Chair, Department of Medical Informatics, Erasmus MC

14:00 - Journey of OHDSI: Where Have We Been and Where We Can Go Together?

Patrick Ryan, PhD, Janssen Research and Development, Department of Biomedical Informatics, Columbia University Medical Center

42:45 - Selection of European Initiatives Using the OMOP CDM

Moderator: Renske Los, PhD, Assistant Professor of Medical Informatics, Department of Medical Informatics, Erasmus MC



Selection of European Initiatives Using the OMOP CDM

- OHDSI Europe National Nodes building opportunities through collaboration - **Renske Los**, Erasmus MC, The Netherlands
- ONCOVALUE: Can Real-World Data Shape the Future of Health Technology Assessment in Oncology? - **Andreas Henriksen**, Copenhagen University Hospital, Denmark
- DigiONE: technical challenges and solutions to the European cancer OMOP conversions from hospital EHR - **Piers Mahon**, DIGICORE, UK
- The PHEMS Project: New Strategies in Health Data Sharing - **Sofia Bazakou**, The Hyve, The Netherlands
- Ecraid: European Clinical Research Alliance on Infectious Diseases - **Ankur Krishnan**, Heidelberg University Hospital, Germany
- PHederation - the Federated Network of Pulmonary Hypertension Registries - **Eva-Maria Didden**, Actelion, a Johnson & Johnson Company, Belgium
- Q&A

SESSION TWO

09:00 - Collaborator Showcase: Rapid Fire Presentation



Moderator: **Katia Verhamme**, MD, Associate Professor of Use and Analysis of Observational Data, Department of Medical Informatics, Erasmus MC, Rotterdam.

1. Adoption of the OMOP Common Data Model in the UK
Speaker: **Alex Knight**, Health Data Research UK
2. Piloting the Transformation of Multiple Sclerosis Real-World Data to the OMOP CDM: Lessons Learned
Speaker: **Tina Parciak**, UHasselt
3. Annotation-preserving machine translation of English corpora to validate Dutch clinical concept extraction tools. Speaker: **Tom Seinen**, Erasmus MC
4. Beyond Diagnostic Codes: A Weakly Supervised learning Framework for Accurate Multimorbidity Identification in Electronic Health Records. Speaker: **Luz Saúde**, Portugal
5. OHDSI meets Flowise to Streamline Biomedical Data Discovery and Analysis
Speaker: **João Almeida**
6. ReportGenerator: Automating study reports and visualization apps for DARWIN EU® research
Speaker: **Cesar Barboza Gutierrez**, Erasmus MC
7. Analysis of Lung Cancer Patient Treatment with Immune Checkpoint Inhibitors Using Natural Language Processing for Data Extraction from Electronic Health Records
Speaker: **Clara L. Oeste and Annelies Verbiest**, Lynxcare
8. An Exploration of Ovarian Cancer Therapy Sequence Utilization in Treatment-naive Women from 2008-2020. Speaker: **Whitney Burton**, Taipei Medical University
9. Baseline Characterization and Treatment Pathways of Patients With Alport Syndrome Across Geographies: Exploring a Rare Disease in Multi-Database Retrospective Cohort Study
Speaker: **Katrin Manlik**, Bayer AG

SESSION THREE

06:55 - Large Scale Evidence Generation in EHDEN and DARWIN EU®

Moderators: **Daniel Prieto-Alhambra** and **Katia Verhamme**, Department of Medical Informatics, Erasmus MC

1. Introduction to EHDEN - **Daniel Prieto Alhambra**
2. Predicting long term cancer survival for Health Technology Assessment: A Multinational Cohort Study Across Europe - **Jeremy Dietz**, National Institute for Health and Care Excellence
3. Trends over time in medicines with suggested shortages in Europe - **Marta Pineda-Moncusi**, Oxford University
4. Q&A
5. Introduction to DARWIN EU® - **Katia Verhamme**
6. DARWIN EU®- Trend of prescription opioid use in Europe - **Annika Jödicke**, Oxford University
7. DARWIN EU®- Treatments of multiple myeloma in Europe from 2012-2022: a population-based network cohort study - **Talita Duarte Salles**, Erasmus MC
8. Q&A



Vocabulary Workshop Part 1 and 2

What you need to know about OHDSI Vocabularies to do phenotyping?

OHDSI Standardized Vocabularies(標準語彙)はOMOP CDMの必須参照標準であり、OHDSIネットワークにおける標準化された強固な研究を可能にする。OHDSI標準語彙は、対象患者を特定し、特徴を構築し、研究結果をナビゲートするのに役立つ。これらをフルに活用するためには、研究者はその操作方法と欠点を軽減する方法を知る必要がある。このチュートリアルでは、phenotypingにおけるVocabulariesの使用方法について説明する。

- 文献中のコードリストからコンセプトセットを作成する
- コンセプトセットを包括的なものにする
- 階層構造（SNOMED、RxNorm、ATC、その他の階層構造）によるコンセプトセット構築の簡素化
- これらすべてにツールを使用する。

Workshops

■ Introduction to OHDSI tutorial

- OHDSIの歴史と理念
- コミュニティはどのように機能しているのか
- OHDSIツールで現在できること
- ツールを使えるようになるには何が必要か
- どこで、どのようにOHDSIについて学ぶことができるか

■ Open-source Development Workshop

OMOP CDMとOHDSIコミュニティへの貢献方法を紹介する。参加者は、OHDSIエコシステムに関連する実践的なスキルと、進行中のプロジェクトに有意義に貢献する方法を学ぶ。ワークショップ終了時には、参加者はOHDSIコミュニティに積極的に参加するための知識を身につけ、実世界の健康アウトカムに影響を与えるツールと研究の両方を強化することができる。

■ Phenotyping Workshop

- Phenotype とは何か？
- なぜ phenotype が必要なのか？
- ATLASで phenotype を作成する方法

■ HADES Workshop

HADESツールを使用してデータベースを検索し、コホートを作成し、簡単な特性解析を実行する方法についてのチュートリアル。