

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 #62

イブニングカンファレンス(第62回)

2025.1.27



## 本日の内容

- OHDSI 論文の紹介
- OHDSI global/APAC から
- 話題
  - LLMを使ったボキャブラリマッピング

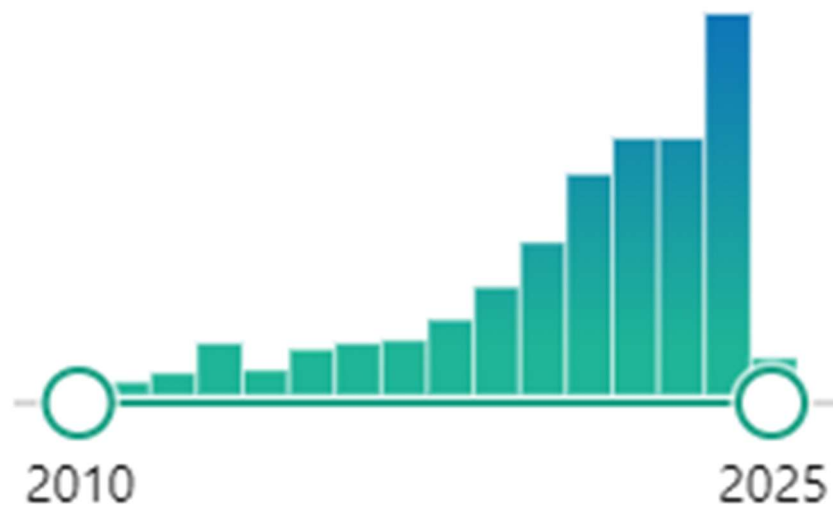


# OHDSI 論文の紹介



## OHDSI関連論文

- Pubmedで”OHDSI or OMOP”を検索



pubmed.ncbi.nlm.nih.govにて作成

- 全期間累計:497報(2024年11月)→514報(2025年1月21日)
- 2024年は約126報





OMOP違い:  
organic  
and metal-  
organic  
polymer

本日、詳細に  
プレゼンテー  
ションいた  
きます

### (前々回積み残し)

1. Berman L, Osthega Y, Giannini J, Anandan LP, Clark E, Spotnitz M, et al. 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.
2. French MA, Hartman P, Hayes HA, Ling L, Magel J, Thackeray A. Coverage of physical therapy assessments in the Observational Medical Outcomes Partnership Model common data model. *Appl Clin Inform.* 2024 Aug 22; [redacted]
3. Hum R, Lane JC, Zhang G, Selles RW, Giladi AM. Observational Health Data Science and Informatics and Hand Surgery Research: Past, Present, and Future. *J Hand Surg Am.* 2024 Oct 17;S0363-5023(24)00433-7.
4. Králik M, Kooš P, Markovič M, Lopatka P. Organic and Metal-Organic Polymer-Based Catalysts-Enfant Terrible Companions or Good Assistants? *Molecules.* 2024 Sep 29;29(19).
5. Mugotitsa B, Bhattacharjee T, Ochola M, Mailosi D, Amadi D, Andeso P, et al. Integrating longitudinal mental health data into a staging database: harnessing DDI-lifecycle and OMOP vocabularies within the INSPIRE Network Datahub. *Front Big Data.* 2024; 7:1435510.
6. Pasha A, Qiao S, Zhang J, Cai R, He B, Yang X, et al. The impact of the COVID-19 pandemic on mental health care utilization among people living with HIV: A real-world data study. *medRxiv : the preprint server for health sciences.* United States; 2024. p. 2024.09.26.24314443.
7. **Kimura E, Kawakami Y, Inoue S, Okajima A. Mapping Drug Terms via Integration of a Retrieval-Augmented Generation Algorithm with a Large Language Model. *Healthc Inform Res.* 2024 Oct;30(4):355-63.**

### (前回積み残し)

1. Tran TC, Schlueter DJ, Zeng C, Mo H, Carroll RJ, Denny JC. PheWAS analysis on large-scale biobank data with PheTK. *Bioinformatics.* 2024 Dec;btac719.
2. Justesen TF, Orhan A, Rosen AW, Gögenur M, Gögenur I. Mismatch Repair Status and Surgical Outcomes in Localized Colorectal Cancer: A Nationwide Cohort Study. *Ann Surg Open.* 2024 Dec;5(4):e499.
3. Kang M, Alvarado-Guzman JA, Rasmussen LV, Starren JB. Evolution of a Graph Model for the OMOP Common Data Model. *Appl Clin Inform.* 2024 Oct;15(5):1056-65.

### (新出)

1. Rao GA, Shoaibi A, Makadia R, et al. CohortDiagnostics: Phenotype evaluation across a network of observational data sources using population-level characterization. *PLoS One.* 2025;20(1):e0310634.
2. Fecho K, Garcia JJ, Yi H, et al. FHIR PIT: a geospatial and spatiotemporal data integration pipeline to support subject-level clinical research. *BMC Med Inform Decis Mak.* 2025;25(1):24.
3. Conover MM, Albogami Y, Hardin J, et al; Observational Health Data Sciences and Informatics (OHDSI) Research Network. Glucagon-Like Peptide 1 Receptor Agonists and Chronic Lower Respiratory Disease Among Type 2 Diabetes Patients: Replication and Reliability Assessment Across a Research Network. *Pharmacoepidemiol Drug Saf.* 2025;34(1):e70087.
4. Kwon YE, Ahn SY, Ko GJ, et al. Impact of Uric Acid Levels on Mortality and Cardiovascular Outcomes in Relation to Kidney Function. *J Clin Med.* 2024;14(1):20.
5. Schuemie MJ, Ostropolets A, Zhuk A, et al. Standardized patient profile review using large language models for case adjudication in observational research. *NPJ Digit Med.* 2025;8(1):18.
6. Hong N, Ko YH, Park JH, et al. A common data model for oral anticoagulants-related risk of spontaneous intracranial hemorrhage. *J Clin Neurosci.* 2025;133:111039.
7. Finster M, Moinat M, Taghizadeh E. ETL: From the German Health Data Lab data formats to the OMOP Common Data Model. *PLoS One.* 2025;20(1):e0311511.
8. de Groot R, Glaser S, Kogan A, et al. ATC-to-RxNorm mappings - A comparison between OHDSI Standardized Vocabularies and UMLS Metathesaurus. *Int J Med Inform.* 2024;195:105777.
9. Basile AO, Verma A, Tang LA, et al. Rapid identification and phenotyping of nonalcoholic fatty liver disease patients using a machine-based approach in diverse healthcare systems. *Clin Transl Sci.* 2025;18(1):e70105.
10. Lee YJ, Kim J, Yu DH, et al. Long-term use of proton pump inhibitors was associated with rapid progression to end stage kidney disease in a Korean nationwide study. *Sci Rep.* 2024;14(1):31477.
11. Yanover C, Magen-Rimon R, Voss EA, et al. Characteristics and Outcomes of Over a Million Patients with Inflammatory Bowel Disease in Seven Countries: Multinational Cohort Study and Open Data Resource. *Dig Dis Sci.* 2024. Epub ahead of print.
12. de Groot R, van der Graaff F, van der Doelen D, et al. Implementing Findable, Accessible, Interoperable, Reusable (FAIR) Principles in Child and Adolescent Mental Health Research: Mixed Methods Approach. *JMIR Ment Health.* 2024;11:e59113.
13. Talvik HA, Oja M, Tamm S, et al. Repeatable process for extracting health data from HL7 CDA documents. *J Biomed Inform.* 2025;161:104765.

今回の対象論文:黄色ハイライト

# オール・オブ・アス研究プログラム： 電子カルテデータを用いた乳管内がんのデータ品質フレームワークへの適用



ORIGINAL REPORTS | August 23, 2024



## Application of a Data Quality Framework to Ductal Carcinoma In Situ Using Electronic Health Record Data From the All of Us Research Program

Authors: [Lew Berman, PhD, MS](#), [Yechiam Ostchega, PhD, RN](#), [John Giannini, PhD](#), [Lakshmi Priya Anandan, MPH](#), [Emily Clark, MPH](#), [Matthew Spotnitz, MD, MPH](#), [Lina Sulleman, PhD](#), [Michael Volynski, PhD](#), and [Andrea Ramirez, MD, MPH](#) | [AUTHORS INFO & AFFILIATIONS](#)

National Institutes of Health, All of Us Research Program, Bethesda, MD., Leidos, Frederick, MD., Vanderbilt University Medical Center, Nashville, TN., InfoPro Systems, Rockville, MD.

### 研究目的:

- 電子カルテ(EHR)データの品質を保証するための枠組みを開発し、運用すること
- All of Us 研究プログラムのデータを使用して、その枠組みの要素を非浸潤性乳管がん(DCIS)の表現型と治療経路に適用すること
- 枠組みの適用を評価するためのチェックリストを提案し、適用すること

### 研究方法:

- 5つのデータ品質の要素(DQD;完全性、一貫性、適合性、妥当性、時間性(temporality))を検証する枠組みを開発。
- 参加者は質問票で人口統計学的な質問に回答。
- データの内部特性を評価し、記述統計および推定統計を用いてデータを外部ベンチマークと比較。
- DQD チェックリストを開発し、各 DQD について概念の選択、内部検証、外的妥当性を評価。

### 結果:

- DCIS に関する OMOP CDM 概念 ID コードを使用し、18 歳以上の女性 2,209 名のコホートを選択した。
- DQDチェックリストの基準を使用して、(1)概念が選択され、内部で適合性が検証された。(2)概念が選択され、完全性が内部で検証された。(3)概念が選択され、内部で検証され、一貫性について外部で検証された。(4)概念が選択され、内部で検証され、妥当性が外部で検証された。(5)概念が選択され、内部で検証され、時間性について外部で検証された。

### 結論:

- All of Us 研究プログラムの EHR データを使用して、DCIS 表現型のデータ品質に関する洞察が得られた。
- 顕著な臨床指標が概念的枠組み内で選択、適用、運用し、提案されたチェックリストを適用することで使用への適合性を評価できることが示された。

DCIS: Ductal Carcinoma In Situ, OMOP CDM: Observational Medical Outcomes Partnership Common Data

# All of Us 研究プログラム:データの質、有用性、多様性

DESCRIPTOR · Volume 3, Issue 8, 100570, August 12, 2022 · Open Access

[Download Full Issue](#)

## The *All of Us* Research Program: Data quality, utility, and diversity

[Andrea H. Ramirez](#)<sup>1,2,32</sup> · [Lina Sulieman](#)<sup>3</sup> · [David J. Schlueter](#)<sup>4</sup> · ... · [Anthony Philippakis](#)<sup>30</sup> · [Dan M. Roden](#)<sup>1,3,31</sup> · [the All of Us Research Program...](#) [Show more](#)

[Affiliations & Notes](#) ▾ [Article Info](#) ▾

### All of Us Research Program:

- NIHが構築した新しいプラットフォーム
- データパスポートモデルを使用して、身体測定、電子健康記録(EHR)データなどの参加者情報へ分析ツールを用いてアクセスできるようにし、クラウドベースの研究者用ワークベンチとして提供
- 新しいプラットフォームを315,000人の参加者に使用した検証研究の結果も提示。
- 参加者の78%は、これまでの医学研究において過小評価されてきたグループに属する人々で、その中には49%の非白人であると自己申告した人も含まれる。
- 再現されたエビデンスには、うつ病と2型糖尿病患者を対象に、人種による薬物使用パターンの違い、喫煙と既知のがんと関連性の検証、人種による影響の報告による心血管系リスクスコアの算出などが含まれる。
- クラウドベースの研究者用ワークベンチは、大規模なリソースと分析ツールへの幅広い研究者の安全なアクセスを可能にするという点で、重要な進歩を意味する。

# All of Us Research Hub



All of Us Research Hub

NIH National Institutes of Health  
All of Us Research Program

ABOUT

DATA & TOOLS

DISCOVER

SUPPORT

RESEARCHER LOGIN



REGISTER

Home > Data & Tools > Researcher Workbench

## Researcher Workbench

The Researcher Workbench is a cloud-based platform where registered researchers can access Registered and Controlled Tier data. Its powerful tools support data analysis and collaboration. Integrated help and educational resources are provided through the Workbench User Support Hub.



### WORKSPACES

Registered researchers use workspaces to access, store, and analyze data for specific research projects. Workspaces are collaborative and can be shared among other registered researchers within a project team.

**USES:** Organizing research projects, collaboration



### ANALYSIS TOOLS

Researchers with R, Python, or SAS experience can perform queries and analyses within the *All of Us* datasets using our integrated cloud-based analysis tools, including Jupyter Notebook, RStudio, and SAS Studio.

**USES:** Analysis, queries

<https://researchallofus.org/data-tools/workbench/>

# Observational Medical Outcomes Partnership 共通データモデルにおける理学療法の評価の適用範囲

Research Article 1003

Department of Physical Therapy and Athletic Training, University of Utah, Salt Lake City, Utah, United States.

## Coverage of Physical Therapy Assessments in the Observational Medical Outcomes Partnership Common Data Model

Margaret A. French<sup>1</sup> Paul Hartman<sup>1</sup> Heather A. Hayes<sup>1</sup> Leah Ling<sup>1</sup> John Magel<sup>1</sup> Anne Thackeray<sup>1</sup><sup>1</sup>Department of Physical Therapy and Athletic Training, University of Utah, Salt Lake City, Utah, United States

Address for correspondence: Margaret A. French, PT, DPT, PhD, Department of Physical Therapy and Athletic Training, University of Utah, 520 Wakara Way, Suite 302, Salt Lake City, UT 84108, United States (e-mail: Maggie.french@utah.edu).

Appl Clin Inform 2024;15:1003-1012.

### 研究目的:

- 神経学的、整形外科的状态に対して使用される理学療法の評価が OMOP CDMにどの程度含まれているかを調査すること

### 研究方法:

- 評価を特定した後、2つのレビューチームが独立して神経学的および整形外科的评价をOMOP CDMにマッピングした。
- レビューチーム内の合意は、両方のレビュー者によってマッピングされた評価の数、一方のレビュー者によってマッピングされたがもう一方のレビュー者によってマッピングされなかった評価の数、またはどちらのレビュー者によってもマッピングされなかった評価の数によって評価した。
- レビューチームは不一致点を調整し、合意内容と評価ごとの平均的な概念ID番号を評価した。

### 結果:

- 81件の神経学的評価のうち、照合後、46.9%は両方のレビュー者によってマッピングされ、53.1%はマッピングされなかった。
- 70件の整形外科的评价のうち、照合後、48.1%は両方のレビュー者によってマッピングされ、51.9%はマッピングされなかった。
- マッピングされた評価のほとんどは、1つ以上の概念ID番号を持っていた(神経学的評価:  $2.2 \pm 1.3$ 、整形外科的评价:  $4.3 \pm 4.4$ )

### 結論:

- OMOP CDMには、神経学的および整形外科的疾患での使用が推奨されている一部の理学療法評価が含まれているが、多くの評価には複数の概念IDが割り当てられている。
- OMOP CDMにさらに多くの機能評価を含め、マッピングのガイドラインを作成すれば、大規模データセットに機能データを組み込む能力が向上するだろう。

OMOP CDM: Observational Medical Outcomes Partnership Common Data Model



## 観察的な医療データ科学と情報学および 手の外科研究:過去、現在、そして未来



Articles in Press | Publish | Topics | Multimedia | CME | About | Contact

THE HAND SURGERY LANDSCAPE · Articles in Press, October 17, 2024

### Observational Health Data Science and Informatics and Hand Surgery Research: Past, Present, and Future

Richard Hum, MS\* · Jennifer CE. Lane, DPhil, FRCS (T&O)† · Gongliang Zhang, PhD, MS‡§ · Ruud W. Selles, PhD ||¶  
Aviram M. Giladi, MD, MS †‡✉

1. Georgetown University School of Medicine, Washington, DC.
2. 2Barts Bone & Joint Health, Blizard Institute, Queen Mary University of London, London, UK.
3. The Curtis National Hand Center, MedStar Union Memorial Hospital, Baltimore, MD; MedStar Health Research Institute, Hyattsville, MD.
4. Department of Rehabilitation Medicine, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands; Department of Plastic and Reconstructive Surgery and Hand Surgery, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands.
5. The Curtis National Hand Center, MedStar Union Memorial Hospital, Baltimore, MD.

- 単一施設研究では、バイアス、一般化可能性と変動性の欠如、まれな状態の研究が不可能などの限界があるが、多施設共同観察研究は、非常に限られる手の外科において、これらの懸念の多くに対処できる可能性がある。しかし、規制上の問題、共通用語の欠如、データセット構造のばらつきなど、数多くの障壁がある。
- OHDSIプログラムは、複数の施設にまたがる大規模な共同研究を可能にすることで、これらの限界を克服することを目的としている。OHDSIは、OMOP CDMを使用してヘルスケアデータを共通言語に標準化し、一貫性のある信頼性の高い分析を可能にする。
- OMOP CDMは、複数のデータベースを単一の語彙による標準化されたコードに変換し、複数のデータセットを一貫して分析可能にするという点で画期的なもの。OMOP CDMを基盤として、OHDSIはデータ抽出から統計モデリングまで、あらゆる研究段階に対応する広範なオープンソースツールを提供している。機密データをローカルに保管し、要約統計のみを共有することで、プライバシー規制への準拠を確保しながら、大規模な分析を可能にする。OHDSIによって、手の外科の研究を深め、結果、リスク要因、合併症、機器の性能に関する理解を深めることができ、最終的にはより良い患者ケアにつながる。

# 縦断的な精神保健データを病期分類データベースに統合:INSPIRE ネットワーク・データハブにおけるDDIライフサイクルとOMOP用語集の活用

frontiers | Frontiers in Big Data

ORIGINAL RESEARCH  
PUBLISHED 11 October 2024  
DOI 10.3389/fdata.2024.1435510

Check for updates

## OPEN ACCESS

EDITED BY  
Chang-Tien Lu,  
Virginia Tech, United States

REVIEWED BY  
Alladoubaye Nguellibaye,  
Shenzhen University, China  
M. Eduard Tudoreanu,  
University of Arkansas at Little Rock,  
United States

\*CORRESPONDENCE  
Bylah Mugotitsa  
✉ [bmugotitsa@aphrc.org](mailto:bmugotitsa@aphrc.org)

RECEIVED 20 May 2024  
ACCEPTED 09 September 2024  
PUBLISHED 11 October 2024

CITATION  
Mugotitsa B, Bhattacharjee T, Ochola M,  
Mailosi D, Amadi D, Andeso P, Kuria J,  
Momanyi B, Omondi F, Kajungu D, Todd J

## Integrating longitudinal mental health data into a staging database: harnessing DDI-lifecycle and OMOP vocabularies within the INSPIRE Network Datahub

Bylah Mugotitsa<sup>1,2\*</sup>, Tathagata Bhattacharjee<sup>3</sup>, Michael Ochola<sup>1</sup>, Dorothy Mailosi<sup>4</sup>, David Amadi<sup>5</sup>, Pauline Andeso<sup>1</sup>, Joseph Kuria<sup>1</sup>, Reinpeter Momanyi<sup>1</sup>, Evans Omondi<sup>1,5</sup>, Dan Kajungu<sup>6</sup>, Jim Todd<sup>3</sup>, Agnes Kiragga<sup>1,7</sup> and Jay Greenfield<sup>4</sup>

1. African Population and Health Research Center (APHRC), Nairobi, Kenya
2. Strathmore University Business School, Strathmore University, Nairobi, Kenya
3. Department of Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom
4. Artificial Intelligence and Machine Learning (AI and ML), CODATA-Committee on Data of the International Science Council, Paris, France
5. Institute of Mathematical Sciences, Strathmore University, Nairobi, Kenya
6. Iganga Mayuge Health and Demographic Surveillance Site (IMHDSS), Makerere University Centre for Health and Population Research (MUCHAP), Kampala, Uganda
7. Infectious Diseases Institute, College of Health Sciences, Makerere University, Kampala, Uganda.

CDM: Common Data Model, DDI:Data Documentation Initiative, OMOP: Observational Medical Outcomes Partnership

Mugotitsa B, Bhattacharjee T, Ochola M, Mailosi D, Amadi D, Andeso P, et al. Integrating longitudinal mental health data into a staging database: harnessing DDI-lifecycle and OMOP vocabularies within the INSPIRE Network Datahub. *Front Big Data*. 2024; 7:1435510.



## 縦断的な精神保健データを病期分類データベースに統合:INSPIRE ネットワーク・データハブにおけるDDIライフサイクルとOMOP用語集の活用

### 研究目的:

- 精神疾患の経時的な進行を理解するには縦断的研究が不可欠だが、うつ病、不安、精神病などの状態を評価するために異なる方法で収集されたデータを組み合わせることは、大きな課題。
- DDI Lifecycle および OMOP CDM 標準を活用し、多様な縦断的データを標準化されたステージングデータベースに変換するマッピング技術を紹介すること。

### 研究方法:

- 「INSPIRE」プロジェクトでは、アフリカの研究から得られた縦断的データを、スノーフレイク・スキーマで構成されたメタデータ文書化標準を使用してステージングデータベースに統合し、OMOP CDMへのデータ統合のための抽出・変換・読込(ETL)スクリプトの開発を容易にした。
- ステージングデータベースのスキーマは、研究プロトコルの変更や、データ収集のタイミングで異なる操作の使用など、縦断的研究の動的な性質を捉えるように設計された。

### 結果:

- 本マッピング方法を利用することで、ステージングデータベースへのデータ移行プロセスを合理化し、OMOP CDMへの統合が可能となった。
- メタデータ標準に準拠することで、データの質が保証され、相互運用性が促進され、精神保健研究におけるデータ共有の機会が拡大した。

### 結論:

- ステージングデータベースは、単なるデータホスティングにとどまらず、包括的な研究記述子として機能する、縦断的な精神保健データの管理における革新的なツール。
- 各研究段階の詳細な洞察を提供し、OMOP CDMへのデータの標準化と統合のためのデータサイエンス基盤を確立する。

CDM: Common Data Model, DDI:Data Documentation Initiative, OMOP: Observational Medical Outcomes Partnership

Mugotitsa B, Bhattacharjee T, Ochola M, Mailosi D, Amadi D, Andeso P, et al. Integrating longitudinal mental health data into a staging database: harnessing DDI-lifecycle and OMOP vocabularies within the INSPIRE Network Datahub. *Front Big Data*. 2024; 7:1435510.



# HIV感染者のメンタルヘルスケア利用における新型コロナウイルス感染症 (COVID-19)パンデミックの影響:リアルデータ研究

medRxiv

THE PREPRINT SERVER FOR HEALTH SCIENCES



BMJ Yale

H

Follow this preprint

**The impact of the COVID-19 pandemic on mental health care utilization among people living with HIV: A real-world data study**

Atana Pasha, Shan Qiao, Jijia Zhang, Ruilie Cai, Buwei He, Xueying Yang, Chen Liang, Sharon Weissman, Xiaoming Li

doi: <https://doi.org/10.1101/2024.09.26.24314443>

1. Department of Health Promotion, Education and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA.
2. South Carolina SmartState Center for Healthcare Quality, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA.
3. Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA.
4. Department of Health Services Policy and Management, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA.
5. Department of Internal Medicine, School of Medicine, University of South Carolina, Columbia, SC, USA.

CDM: Common Data Model, Incidence Rate ratio, OMOP: Observational Medical Outcomes Partnership, PLWH: people living with HIV

Pasha A, Qiao S, Zhang J, Cai R, He B, Yang X, et al. The impact of the COVID-19 pandemic on mental health care utilization among people living with HIV: A real-world data study. medRxiv : the preprint server for health sciences. United States; 2024. p. 2024.09.26.24314443.

# HIV感染者のメンタルヘルスケア利用における新型コロナウイルス感染症 (COVID-19)パンデミックの影響:リアルデータ研究

## 研究目的:

- 本研究では、All of Usプログラムの電子健康記録(EHR)と基礎調査データを利用して、COVID-19パンデミック中のHIV感染者(PLWH)におけるメンタルヘルスケアの利用と関連要因を調査した。

## 研究方法:

- レトロスペクティブコホートデザインを用いて、2018年3月から2022年3月までのAll of Usプログラムで関連するOMOP CDMコンセプトセットに基づき、コンピューショナル・フェノタイピングにより、4,575人のPLWHを特定し、対象者とした。
- メンタルヘルスケアの利用状況は、メンタルヘルスケア受診の年間カウント数を用いて測定し、利用パターンはパンデミック前(2018年~2020年)とパンデミック中(2020年~2022年)で比較した。
- メンタルヘルスケアの利用、COVID-19感染歴、人口統計学的要因、既存の慢性疾患(高血圧、糖尿病など)、社会経済的地位の関連性を調べるために、ポアソン一般化線形混合モデル(GLMM)による発生率比(IRR)を使用した。

## 結果:

- パンデミック期間におけるメンタルヘルスケアの年間受診回数は、パンデミック前の期間と比較して有意に減少した(IRR = 0.89)。
- COVID-19感染歴がある場合、メンタルヘルスケアの利用回数が多いことが判明した(IRR = 1.35)。
- 30~39歳、40~49歳、50~64歳のグループは、最も若いグループ(18~29歳)と比較して、受診回数が有意に高かった。
- 黒人/アフリカ系アメリカ人の参加者は、白人参加者と比較して受診回数が少ない傾向にあった(IRR = 0.71)。
- メディケイド(低所得者向け医療扶助制度)は受診回数の増加と関連していた(IRR = 1.29)。
- 雇用者または労働組合が提供する保険は、メンタルヘルスケアの利用減少と関連していた(IRR = 0.54)。
- 既存の併存疾患は、メンタルヘルスケアの利用回数を大幅に増加させた(併存疾患1つ:IRR = 5.49、2つ以上:IRR = 10.4)。

CDM: Common Data Model, Incidence Rate ratio, OMOP: Observational Medical Outcomes Partnership, PLWH: people living with HIV

# PheTKによる大規模バイオバンクデータのPheWAS解析

JOURNAL ARTICLE

Bioinformatics

## PheWAS analysis on large-scale biobank data with PheTK [Get access >](#)

Tam C Tran ✉, David J Schlueter, Chenjie Zeng, Huan Mo, Robert J Carroll,

Joshua C Denny

1. National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, 20892, United States.
2. University of Toronto, ON, M5S 1A1, Canada.
3. Vanderbilt University School of Medicine, Nashville, TN, 37240, United States.
4. All of Us Research Program, National Institutes of Health, Bethesda, MD, 20892, United States.

PheWAS: phenome-wide association studies

### 要約:

- 膨大なコホートの電子カルテ(EHR)データと関連付けられた遺伝子データの急速な増加に伴い、大規模なフェノーム全体関連解析(PheWAS)は、生物医学研究における強力な発見ツールとなっている。
- PheWASは、経時的なEHRデータを利用して表現型関連を研究する解析手法で、これまでは、主に小規模なデータセットと初期のPheWASアプローチで開発されていた。
- PheTKは分析を簡素化し、バイオバンク規模のデータを効率的に処理できるように設計されている。マルチスレッドを使用し、OMOPデータベースやHailマトリクス表からのデータ抽出、phecodeバージョン1.2およびphecodeXの両方に対するPheWAS分析を含む、PheWASの完全なワークフローをサポートしている。
- ベンチマークの結果、PheTKは同じワークフローを完了するのに、R PheWASパッケージよりも64%少ない時間で処理できることが示された。
- PheTKはローカルでも、All of Us Researcher Workbench(All of Us)やUK Biobank(UKB) Research Analysis Platform(RAP)などのクラウドプラットフォームでも実行できる。

# 限局性大腸がんにおけるミスマッチ修復状態と外科的治療成績： 全国規模のコホート研究

ANNALS OF  
SURGERY OPEN

OPEN

Original Study

## Mismatch Repair Status and Surgical Outcomes in Localized Colorectal Cancer

### A Nationwide Cohort Study

Tobias Freyberg Justesen, MD,\* Adile Orhan, MD,\* Andreas Weinberger Rosen, MD,\* Mikail Gögenur, MD,\* and Ismail Gögenur, MD, DMSc†

1. From the Department of Surgery, Center for Surgical Science, Zealand University Hospital, Køge, Denmark.
2. Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark.

### 研究目的：

- 限局性大腸がん患者における、ミスマッチ修復欠損(dMMR)とミスマッチ修復能正常(pMMR)の状態と、全生存期間および無病生存期間との関連性を調査すること。

### 研究方法：


- 本コホート研究では、2009年から2020年の間に限局性大腸がんの根治手術を受けた患者を対象とした。
- 患者はデンマーク大腸がんグループのデータベースで特定し、患者レベルのデータを6つの登録データベースから抽出した。
- 対象患者を抽出した後、推定傾向スコアを用いてdMMR状態の患者とpMMR状態の患者を1:1でマッチングさせた。

### 結果：

- マッチング後、5994人の患者が対象となった。患者の年齢中央値は74歳、追跡期間中央値は4.1年であった。
- ミスマッチ修復(MMR)状態と全生存期間(ハザード比、0.91;95%信頼区間[CI]、0.81-1.03)または無病生存期間(ハザード比、0.89;95% CI、0.78-1.01)との間に有意な関連性は認められなかった。
- 比例ハザード仮説の違反により算出された制限付きの5年無病生存期間中央値では、dMMR群が0.13年(95% CI、0.03~0.23;P = 0.01)の有意な絶対差を示した。

### 結論：

- MMRの状態による全生存期間との有意な関連性は認められなかった。
- 治癒目的の手術を受ける限局性大腸がん患者では、dMMRの状態はpMMRの状態と比較して無病生存期間がわずかに改善されることが分かった。



## 限局性大腸がんにおけるミスマッチ修復状態と外科的治療成績： 全国規模のコホート研究

### Danish Healthcare Registries:

- デンマークの6つの全国データベースから患者レベルのデータ [DCCGデータベース、全国研究用検査結果登録 (RLRR)、デンマーク全国患者登録 (DNPR)、デンマーク全国処方登録 (NPR)] が OMOP CDM に変換され、統合された。
- OMOP CDM には患者の重複は存在せず、データベースの患者データが統合され、DCCG データベースで特定されたユニークな(一意の)患者として追加された。
- OMOP CDM で再発の可能性を推定するため、デンマーク国立病理学レジストリ、デンマークがんレジストリから特定の変数が追加された。DCCG データベースには再発疾患のデータが記録されていないため、再発の推定には Lash 氏らによって開発された検証済みのアルゴリズムが使用された。
- OMOP CDM へのデータの統合は、デンマークの住民全員に個人識別番号を割り当て、レジストリ全体で患者の識別を可能にするデンマークの市民登録システムによって可能となった。DCCG データベースには、2001年以降、CRCと診断された18歳以上のデンマーク国民が登録されている。
- 変数には、患者特性、術前治療、腫瘍の病期、診断日、治療計画、手術などが含まれる。MMRタンパク発現に関するデータは2009年10月1日以降入手可能である。RLRRは2015年より、デンマークの5つの地域すべてから生化学検査結果のデータを収集している。DNPRは1977年以降、入院の種類、入院理由、診断、治療、検査を含むデンマークの病院からのデータを収集している。1995年より、NPRはデンマークの薬局で処方されたすべての処方薬に関する個人レベルの情報を収集している。



# OMOP共通データモデルのためのグラフモデルの進化

## Evolution of a Graph Model for the OMOP Common Data Model

Mengjia Kang, Jose A. Alvarado-Guzman, Luke V. Rasmussen, Justin B. Starren

1. Division of Pulmonary and Critical Care Medicine, Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, United States.
2. Neo4j, Inc., San Mateo, California, United States.
3. Division of Health and Biomedical Informatics, Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States.
4. University of Arizona Health Sciences, Tucson, Arizona, United States

### 研究目的:

- 電子カルテデータ用のグラフデータベースは臨床研究に有用なツールだが、リレーショナルデータベースをグラフデータベースのスキーマに変換する方法は公開されていない。
- 研究機関間で再利用可能なグラフモデルを、OMOP CDM用に開発すること。

### 研究方法:

- OMOPの標準化された臨床および用語集テーブルを、Neo4jグラフデータベース内のプロパティグラフモデルに変換するための2つの異なる戦略を表す4つのモデルを作成し、評価した。
- SCRIPTとCRITICALの2つのコホートを、それぞれ異なるサイズのテストデータセットとして使用し、データベース構築時間、クエリの複雑性、両コホートの実行時間などのデータベースパフォーマンスの観点から、作成されたグラフモデルの2つを比較した。

### 結果:

- 重要な情報を属性ではなくトポロジーとして保存するように最適化されたグラフスキーマを利用することで、データ作成とクエリの両方で大幅な改善が見られた。
- より規模の大きいコホート・CRITICALのグラフデータベースは、134,145人の患者を対象に、合計749,011,396ノード、1,703,560,910エッジで、1時間以内に構築することができた。
- OMOP CDMをグラフ最適化スキーマに変換する初めての汎用的なソリューションで、このモデリング方法は他のOMOP CDM v5.xデータベースにも適用できる。SCRIPTコホートとCRITICALコホートを用いた我々の評価と、現行バージョンと旧バージョンの比較により、コードの簡潔さ、データベース構築、クエリ速度の面で優位性があることが示された。

### 結論:

- OMOP CDM データベースをグラフデータベースに変換する方法を開発した。最終的なモデルは、コードの簡潔性とクエリーの効率性の両面で、特に複雑なクエリーにおいて、初期のリレーショナル・グラフ変換よりも優れた性能を発揮することが明らかになった。

# CohortDiagnostics: 集団レベルの特性評価を用いた、 観察データソースのネットワーク全体にわたる表現型の評価

## CohortDiagnostics: Phenotype evaluation across a network of observational data sources using population-level characterization

Gowtham A. Rao , Azza Shoaibi , Rupa Makadia, Jill Hardin, Joel Swerdel, James Weaver, Erica A. Voss, Mitchell M. Conover, Stephen Fortin, Anthony G. Sena, Chris Knoll, Nigel Hughes, James P. Gilbert, [...], Patrick B. Ryan 

1. Observational Health Data Analytics, Janssen Research and Development, LLC, Titusville, NJ, USA.
2. OHDSI Collaborators, Observational Health Data Sciences and Informatics (OHDSI), New York, NY, USA.
3. Department of Biostatistics, University of California, Los Angeles, CA, USA.
4. Department of Biomedical Informatics, Columbia University, New York, NY, USA.

### 研究目的:

- オープンソースツールであるCohort Diagnosticsを使用した表現型アルゴリズム(PA)の評価のための新しい枠組みを紹介すること。

### 研究方法:

- 今回用いた方法は、PAによって返された患者コホートを評価するための複数の診断基準に基づく。
- 診断には、発症率の推定、インデックス日のエントリーコードの分類、インデックス日の前後で観察された臨床イベントの全発生率が含まれる。
- 10種類の異なる観察データソースを対象に、全身性エリテマトーデス(SLE)のPA1つとアルツハイマー病(AD)のPA2つを評価することで、この枠組みを検証した。

### 結果:

- CohortDiagnosticsを利用することで、SLEのコホートにおける個人の集団レベルの特性が、この疾患の想定される臨床プロファイルとほぼ一致していることが分かった。
- 具体的には、SLEの発症率は女性の方が常に高く、臨床検査、治療、繰り返される診断などの想定される臨床イベントも観察された。
- ADについては、1つのPAで特定された患者数はかなり少なかったものの、2つのコホート間の臨床特性に顕著な違いが見られなかったことから、特異性は同程度であることが示唆された。
- OMOPデータソースのネットワーク全体で、2つの臨床的疾患を例として、実用的でデータ主導型のPA評価アプローチを提供した。
- Cohort Diagnosticsは、特定のPAによって識別された対象が、研究調査への組み入れ対象と一致することを保証した。

### 結論:

- 大規模な集団レベルの特性に基づく診断は、PAの誤分類エラーに関する洞察を提供することができる。

# HADES Packages

HADES Packages Validation Publications Support Developers

Learn more... Learn more...

Population-level estimation  
Patient-level prediction  
**Characterization**  
Cohort construction and evaluation  
Evidence Quality  
Supporting packages

### Cohort construction and evaluation

- Capr**  
Develop and manipulate complex cohort definitions in R  
[Learn more...](#)
- CirceR**  
An R wrapper for Circe, a library for creating cohort definitions, expressing them as JSON, SQL, or Markdown.  
[Learn more...](#)
- CohortGenerator**  
Instantiating cohorts in a database based on a set of cohort definitions.  
[Learn more...](#)
- PhenotypeLibrary**  
The OHDSI Phenotype Library: a collection of community-maintained pre-defined cohorts.  
[Learn more...](#)
- CohortDiagnostics**  
Generate a wide set of diagnostics to evaluate cohort definitions against databases in the CDM.  
[Learn more...](#)
- PheValuator**  
Semi-automated evaluation of cohorts, producing metrics such as sensitivity, specificity, and positive and negative predictive value.  
[Learn more...](#)
- CohortExplorer**  
Visually explore all individual-level data of patients in a cohort  
[Learn more...](#)
- Keeper**  
Knowledge-Enhanced Electronic Profile Review  
[Learn more...](#)

### Evidence Quality

- Achilles**  
Generate descriptive statistics on an entire OMOP CDM database  
[Learn more...](#)
- DataQualityDashboard**  
Expose and evaluate observational data quality.  
[Learn more...](#)
- EmpiricalCalibration**  
Use negative control exposure-outcome pairs to profile and calibrate a particular analysis design.  
[Learn more...](#)
- MethodEvaluation**  
Use real data and established reference sets as well as simulations

<https://ohdsi.github.io/Hades/packages.html>



# FHIR PIT:対象者レベルの臨床研究をサポートする 地理空間および時空間データの統合パイプライン

## FHIR PIT: a geospatial and spatiotemporal data integration pipeline to support subject-level clinical research

Karamarie Fecho<sup>1,2\*</sup>, Juan J. Garcia<sup>3†</sup>, Hong Yi<sup>1†</sup>, Griffin Roupe<sup>1,3</sup> and Ashok Krishnamurthy<sup>1,3</sup>

1. Renaissance Computing Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.
2. Copperline Professional Solutions, LLC, Pittsboro, NC, USA.
3. Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.
4. Renaissance Computing Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

### 背景:

- 電子健康記録(EHR)データや臨床データを使用して臨床研究を行う際には、大気汚染物質などの環境曝露や社会経済指標を考慮することが重要であるとますます認識されるようになっている。
- 研究をサポートする公共の地理空間および時空間データのソースは数多く存在するが、ファイルフォーマットや時空間解像度の不整合、ファイルサイズが大きいことによる計算上の問題、患者または被験者の時空間データ用のツールの不足などにより、データの利用は困難となっている。

FHIR PIT(HL7® Fast Healthcare Interoperability Resources Patient data Integration Tool)



### 結果:

- FHIR®形式のEHRデータを使用し、空間的・時間的解像度やファイル形式が異なる環境曝露データと患者・被験者レベルのデータを統合する、オープンソースのモジュール式データ統合ソフトウェアパイプラインとして、FHIR PITを開発した。
- FHIR-PITを用いて、喘息/肺疾患の患者、原発性纖毛運動不全/稀な肺疾患の患者の2つのコホートから環境曝露データと統合された患者レベルのEHRデータを含む「統合機能テーブル」を生成した。
- データは代表的な大気汚染物質(粒子状物質とオゾン)への曝露と、呼吸器系の問題による救急外来または入院の年間発生件数との関係を調査するために照会された。
- 呼吸器系の問題による入院は、粒子状物質とオゾンの曝露レベルが比較的高い患者に多く見られ、喘息患者よりも原発性纖毛運動不全症患者の方が全体的に入院件数が多かった。

### 結論:

- FHIR PIT v1.0 の主要リリース、FHIR PIT の使用に関する技術的なデモユースケースと臨床応用例を提示し、喘息/原発性纖毛運動異常症に関連する環境曝露と健康結果の研究をサポートすることを示した。
- FHIR以外のCDMへのツールの適用については、PCORnet、i2b2、OMOP CDMからFHIRへのマッピングを行うオープンソース変換ツールを提供している。



# Global/APACの動き



# 1月の OHDSI Global/APAC

## ● Global Community Call テーマ

- Jan. 21 Clinical Guideline Review, Session I
- Jan. 28 Clinical Guideline Review, Session II

## ● APAC Call テーマ

- Jan. 16 APAC 2025 Kickoff (Mui), Newcomers Session (Steven, Keiko, Jiwon)



Steven Yong, CEO of NovoHeal



Jiwon Um



# Save The Dates!

OHDSI Europe Symposium - Save-the-date!



OHDSI BELGIUM



Save-the-date

5-7 July 2025

Location

Old Prison - Hasselt  
University  
Martelarenlaan  
Hasselt - BELGIUM



**Save the Date**  
31/07 e 01/08

Evento OHDSI  
LATAM

SALVADOR  
**Bahia**  
» **BRASIL** «





# Workgroup OKRs

Each year, workgroup representatives join a February community call to present the mission, objectives and key results for their respective groups. These 2-4 minute presentations are recorded and posted on the Workgroups homepage on OHDSI.org.

**Please choose a date to sign up for a February date; once a date has at least 10 workgroups, it will be closed.**



## Already Signed Up:

- Oncology
- Rare Disease
- Common Data Model
- Steering
- CDM Survey Subgroup
- Latin America
- Clinical Trials
- GIS - Geographic Information System
- Health Systems Interest Group
- Eye Care and Vision Research
- Transplant
- Themis
- Medical Devices





# 2024 OHDSI APAC Achievements

## Research

**Build research expertise and collaboration amongst the different chapters through publication**

### Milestones

- Initiated new APAC study in collaboration with Singapore HSA
- Showcased initial study results at the 2024 APAC Symposium

## Training

**Create an APAC training program to expand reach to the general community**

### Milestones

- Hosted 4 in-person trainings in Japan, Thailand, Taiwan and Singapore
- Initiated and successfully completed first community-wide ETL project

## Communication

**Create collaboration activities that encourage collaborative generation and dissemination of evidence that promotes better health decisions and better care**

### Milestones

- Hosted 2024 APAC symposium
- Distributed 4 quarterly newsletters
- Hosted 12 community calls and 13 scientific forums



# 2025 OHDSI Focus Areas

Guideline-driven  
Evidence  
Generation

Evidence-driven  
Data  
standardization

Evidence-driven  
Open Source  
Development

Evidence-driven  
Collaborative  
Education

## Dry January:

Guideline review to determine evidence needs where RWE could potentially contribute

## Analysis April:

Prepare protocol and analysis specification to initiate network execution

## Phenotype February:

Develop/evaluate cohorts needed to support filling the evidence gaps

## Meta-analysis May:

Collaborative interpretation of results from across network

## March to Data Fitness:

Evidence network to determine which partners are appropriate to generate which evidence

## Journey to June:

Mid-year reflection on evidence generation process and progress

## Spread-the-Word Second Half: Focus on Evidence Dissemination

July: OHDSI EU

August:

September:

October: OHDSI Global (tbd)

November:

December: OHDSI APAC





# 2025 OHDSI Focus Areas

Guideline-driven  
Evidence  
Generation

Evidence-driven  
Data  
standardization

Evidence-driven  
Open Source  
Development

Evidence-driven  
Collaborative  
Education

## 2025 Priorities:

### 1. Evidence Network engagement

- Data partner organizations with source data converted to OMOP CDM v5.4 are encouraged to become part of OHDSI Evidence Network
- We will conduct Evidence Network evaluations of 'fitness-for-use' based on evidence needs identified by the community

### 2. OHDSI Standardized Vocabularies community contributions

- OHDSI Vocabulary team has defined 2025 roadmap for February and August releases
- We are seeking community contributions to expand vocabulary content (concepts) and improve mapping (relationships), and to improve our own processes for incorporating community contributions





# 2025 OHDSI Focus Areas

Guideline-driven  
Evidence  
Generation

Evidence-driven  
Data  
standardization

Evidence-driven  
Open Source  
Development

Evidence-driven  
Collaborative  
Education

## 2025 Priorities:

1. Harden Strategus and associated HADES packages to enable improved user installation and execution experience
2. Improve OHDSI packages based on the user needs and experience from Guideline-driven Evidence Generation study teams
3. Redesign ATLAS with focus on improving the experience of cohort design and evaluation



# 2025 OHDSI Focus Areas

Guideline-driven  
Evidence  
Generation

Evidence-driven  
Data  
standardization

Evidence-driven  
Open Source  
Development

Evidence-driven  
Collaborative  
Education

## 2025 Priorities:

1. Refresh 'Book of OHDSI' to align with current practice and increase emphasis on evidence generation and dissemination for evidence *generators*
2. Lean into JACC Partnership to promote best practice to evidence *consumers*



# 2025 OHDSI APAC Goals

## Research

Guideline-driven  
Evidence  
Generation

Evidence-driven  
Data  
standardization

Build research expertise and collaboration amongst the different chapters through publication

### Milestones

- Continue and complete 2024 APAC Study (GLP-1 ALI)
- Select and conduct 2-3 new network studies led by APAC
- Strengthen APAC presence in OHDSI Evidence Network

## Training

Evidence-driven  
Data  
standardization

Evidence-driven  
Collaborative  
Education

Create an APAC training program to expand reach to the general community and provide more opportunities for hands-on learning

### Milestones

- Formalize an OHDSI/OMOP curriculum in APAC
- Conduct a community-wide project on ETL and/or vocabulary mapping
- Support update and localization of Book of OHDSI

## Communication

Guideline-driven  
Evidence  
Generation

Create collaboration activities that encourage collaborative generation and dissemination of evidence that promotes better health decisions and better care

### Milestones

- Host APAC symposium and issue certificates for participation
- Distribute quarterly newsletters
- Host monthly community calls and scientific forums



# LLMを使ったボキャブラリマッピング