

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 Meeting #74

2026年1月 イブニング・カンファレンス
2026.1.27



本日の内容

- OHDSI 論文の紹介(1)
- OHDSI 論文の紹介(2)
- OHDSI global/APAC から
- 話題

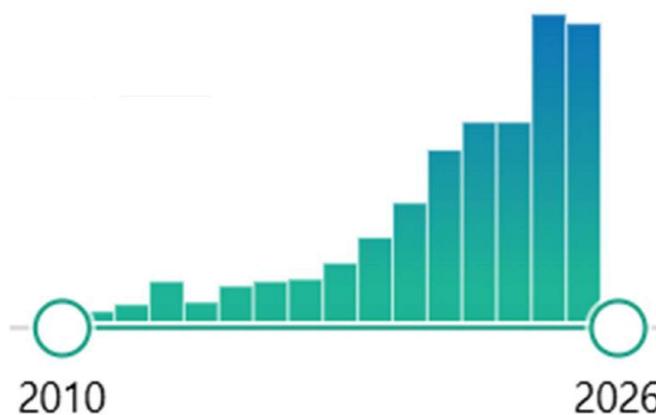


OHDSI 論文の紹介(1)



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: 624報(2025年12月22日) →631報(2026年1月21日)
- 2024年は約126報



1.について
は1月に発
表いただく

(積み残し論文)

1. Aoyagi Y, Baba M, Terao S, Ikeda Y, Nomura K, Sato A. Feasibility of Converting EMR Data to OMOP CDM and Utilizing OHDSI Analysis Tools in Japan. *Stud Health Technol Inform*. 2025年8月;329:1946–7.
2. Bellas L, Català M, Burn E, Guo Y, Du M, Verhamme K, et al. Secular Trends in the Use of Valproate-Containing Medicines in Women of Childbearing Age in Europe: A Multinational DARWIN EU Network Study. *Pharmacoepidemiol Drug Saf*. 2025 Oct;34(10):e70232.
3. Lee KH, Jang S, Kim GJ, Park S, Kim D, Kwon OJ, et al. Large Language Models for Automating Clinical Trial Criteria Conversion to Observational Medical Outcomes Partnership Common Data Model Queries: Validation and Evaluation Study. *JMIR Med Inform*. 2025 Oct;13:e71252.
4. Lim HC, Wong H, Philip R, Van Der Vegt A, Choo KKR, Pole JD, et al. Streamlining electronic medical record data extraction and validation in digital hospitals: A systematic review to identify optimal approaches and methods. *Learn Health Syst*. 2025 Oct;9(4):e70024.
5. Razzaghi H, Dickinson K, Wieand K, Boss S, Weidlich H, Huang Y, et al. A multifaceted approach to advancing data quality and fitness standards in multi-institutional networks. *J Am Med Inform Assoc*. 2025 Oct;ocaf181.
6. Jung G, Lee J, Gho SM, Han Y, Choi B, Cho JW, et al. Predicting oxcarbazepine-induced hyponatremia in adult epilepsy patients: A multicenter machine learning analysis using real-world CDM data. *Seizure*. 2025 Oct;133:167–74.
7. Datzmann T, Lang C, Tesch F, Spoden M, Dröge P, Ehm F, et al. Evaluation of hybrid stroke quality indicators by integrating NIHSS and claims data for improved outcome prediction. *Sci Rep*. 2025 Nov;15(1):38994.
8. Rappoport N, Livne G, Perry Cohen N, Makover N, Eshel-Geva H, Kapach H, et al. Kineret: Israel's Largest Hospital Network Transformed into the OMOP common data model for collaborative research. *PLoS One*. 2025;20(10):e0334848.
9. Floyd SB, Mills A, Woloff J, Lawson C, Hilton C, Oeffinger D, et al. Development and validation of a computable phenotype for adolescent idiopathic scoliosis. *Learn Health Syst*. 2025 Oct;9(4):e70018.
10. Adams MCB, Hurley RW, Bartels K, Perkins ML, Hudson C, Topaloglu U, et al. Extending the Observational Medical Outcomes Partnership (OMOP) Common Data Model for Critical Care Medicine: A Framework for Standardizing Complex ICU Data Using the Society of Critical Care Medicine's Critical Care Data Dictionary (C2D2). *Crit Care Med*. 2025 Nov;
11. Ompad G, Cesta CE, Cohen JM, Leinonen MK, Taipale H, Li H, et al. Current Use of Common Data Models in the Nordic Countries. *Pharmacoepidemiol Drug Saf*. 2025 Nov;34(11):e70242.
12. Guerrero P, Ernebjerg M, Holst T, Weese D, DiBello H, Ibing S, et al. The AIR-MS data platform for artificial intelligence in healthcare. *JAMIA Open*. 2025 Dec;8(6):ooaf145.
13. Finster M, Wenzel M, Taghizadeh E. Common data models and data standards for tabular health data: a systematic review. *BMC Med Inform Decis Mak*. 2025 Nov;25(1):422.
14. Yoo, K. H.; Lee, K. J.; Lee, S. M.; Han, C.; Park, R. W.; Jo, Y. T. Comparative Effectiveness of Selective Serotonin Reuptake Inhibitors versus Serotonin-Norepinephrine Reuptake Inhibitors in the Risk of Diagnostic Conversion from Unipolar Depression to Bipolar Disorder. *Int J Psychiatry Clin Pract* 2025, 1–9. <https://doi.org/10.1080/13651501.2025.2600083>.
15. Torres-Silva, E. A.; Gaviria-Jiménez, J. J.; Guevara-Zambrano, A. M.; Herrera-Almanza, L.; Flórez-Arango, J. [Datos sintéticos de un modelo de datos común para las aplicaciones de inteligencia artificial en salud materna: reporte de experiencia en el contexto colombiano]. *Biomedica* 2025, 45 (Sp. 3), 71–92. <https://doi.org/10.7705/biomedica.7937>.

今回の対象論文:黄色ハイライト



(積み残し、残り)

1. Berger, O.; Menashe, S.; Damti Geva, S.; Yakubov, R.; Ben Yehuda, M.; Peleg, M.; Talisman, R. Lipomas Are Associated with a Higher Prevalence of Metabolic Syndrome Components: A Multicenter Cross-Sectional Study. *Front Endocrinol (Lausanne)* 2025, 16, 1721570. <https://doi.org/10.3389/fendo.2025.1721570>.
2. Kim, G. L.; Yi, Y. H.; Lee, J. G.; Tak, Y. J.; Lee, S. H.; Ra, Y. J.; Choi, B. K.; Lee, S. Y.; Cho, Y. H.; Park, E. J.; Lee, Y.; Choi, J. I.; Lee, S. R.; Kwon, R. J.; Son, S. M. Association Between the Use of DPP4 Inhibitors and Metformin and the Risk of Cancer in Patients with Type 2 Diabetes: A Multicenter Retrospective Cohort Study Using the OMOP CDM Database. *Cancers (Basel)* 2025, 17 (22). <https://doi.org/10.3390/cancers17223620>.

(新出論文)

1. Sadda SR, Wykoff CC, Chowers I, Korobelnik JF, Narayana R, Pappuru RR, et al. Initiation of a global consortium to study the progression of age-related macular degeneration: RIMR AMD consortium report # 1. *Graefes Arch Clin Exp Ophthalmol*. 2025 Dec;
2. Murgia Y, Gazzarata R, Ciampi M, Sicuranza M, Cirillo F, Esposito C, et al. The challenges of national health data ecosystems in feeding the European health data space: the Italian example. *Front Med (Lausanne)*. 2025;12:1644719.
3. Liu J, Pandya S, Coppi A, Young HP, Krumholz HM, Schulz WL, et al. Assessment of the integrity of real-time electronic health record data used in clinical research. *PLoS One*. 2026;21(1):e0340287.
4. Van der Pol H, Kringelbach T, Martin Agudo M, Bratseth Stav G, Fagereng GL, Fiocco M, et al. Procedures of data merging in precision cancer medicine: the PRIME-ROSE project. *Acta Oncol*. 2026 Jan;65:1–8.
5. Gong G, Liu J, Pandya S, Taborda C, Wiesendanger N, Price N, et al. Clinical Trial Patient Matching: A Real-Time, Common Data Model and Artificial Intelligence-Driven System for Semiautomated Patient Prescreening in Cancer. *Clinical Trials*. *JCO Clin Cancer Inform*. 2026 Jan;10:e2500262.
6. Park WY, Sippel Schmidt T, Salvador G, O'Donnell K, Genereaux B, Jeon K, et al. Response to “toward semantic interoperability of imaging and clinical data:

reflections on the DICOM-OMOP integration framework”. *J Am Med Inform Assoc*. 2025 Dec;ocaf216.

7. Cheng W, Yu Z. Toward semantic interoperability of imaging and clinical data: reflections on the DICOM-OMOP integration framework. *J Am Med Inform Assoc*. 2025 Dec;ocaf215.



欧州における出産可能年齢の女性におけるバルプロ酸含有医薬品の使用に関する長期的な傾向: 多国籍DARWIN EUネットワーク研究

[Secular Trends in the Use of Valproate-Containing Medicines in Women of Childbearing Age in Europe: A Multinational DARWIN EU Network Study.]

Lucía Bellas 1,2,3, Martí Català 1, Edward Burn 1 Yuchen Guo1, Mike Du 1, Katia Verhamme 4, Egil Fridgeirsson 4, Talita Duarte-Salles 4,5, Tommi Kauko 6, Eeva Kronqvist 6, James T. Brash 7, Sarah Seager 7, Daniel Prieto-Alhambra 1,4, Annika M. Jödicke 1, Albert Prats-Urbe 1

1. Pharmacoepidemiology-and Device Group, NDORMS, University of Oxford, Oxford, UK
2. Clinical Pharmacology Department, Hospital Universitari Vall d'Hebron, Barcelona, Spain
3. Department of Pharmacology, Therapeutics and Toxicology, Universitat Autònoma de Barcelona, Barcelona, Spain
4. Department of Medical Informatics, Erasmus Medical Center, Rotterdam, the Netherlands
5. Fundació Institut Universitari per a la recerca al'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain
6. Auria Clinical Informatics, ACI VARHA, Turku University Hospital, Turku, Finland
7. IQVIA, Real World Solutions, Brighton, UK



欧州における出産可能年齢の女性におけるバルプロ酸含有医薬品の使用に関する長期的な傾向: 多国籍DARWIN EUネットワーク研究

[Secular Trends in the Use of Valproate-Containing Medicines in Women of Childbearing Age in Europe: A Multinational DARWIN EU Network Study.]

背景

- バルプロ酸(VPA)はてんかんなどの第一選択薬であるが、胎児奇形・神経発達障害のリスクが高いことが確立している
- 欧州では2014年・2018年にEMAがリスク最小化策(RMM)を強化し、妊娠可能年齢女性への使用を原則禁止または厳格制限
- しかし、実臨床では依然として処方が続いている可能性があり、長期的な使用動向の評価が必要とされた

目的

- 2010～2022年における妊娠可能年齢女性(12-55歳)でのVPA使用の年次推移(発生率・有病率)を欧州6か国で評価
- 新規VPA使用者について、年齢、適応、初期用量、治療期間、併存疾患、避妊薬使用状況を記述

方法

- DARWIN EUネットワークを用いた多国籍・分散型データ解析研究
- オランダ、スペイン、英国、ベルギー、ドイツ、フィンランドのEHR/処方データをOMOP CDM1に統一
- 各国で同一解析コードを実行し、集計結果のみを共有



欧州における出産可能年齢の女性におけるバルプロ酸含有医薬品の使用に関する長期的な傾向: 多国籍DARWIN EUネットワーク研究

[Secular Trends in the Use of Valproate-Containing Medicines in Women of Childbearing Age in Europe: A Multinational DARWIN EU Network Study.]

主な結果

- VPAの新規使用(発生率)・使用者割合(有病率)は、2015年以降一貫して減少
 - ・ 特に若年層(<45歳)で減少が顕著。
- 一方で、代替薬(特にプレガバリン、ガバペンチン)の使用は増加
- 新規VPA使用者の中央値年齢は40-43歳と比較的高齢。
- 不安障害・うつ病の併存率が高い
- VPA開始時のホルモン避妊薬使用率は全体的に低水準(国差あり)
- 初期用量は大きく変化していないが、累積使用量は一部国で減少

結論

- 欧州の規制措置後、妊娠可能年齢女性におけるVPA使用は確実に減少
- ただし、代替薬の増加や避妊対策の不十分さなど、新たな安全性課題が示唆される
- DARWIN EUのような標準化ネットワーク解析は、迅速かつ信頼性の高い規制支援エビデンスを提供できる



大規模言語モデルを用いた臨床試験基準から観察研究用医療アウトカムパートナーシップ共通データモデルクエリへの自動変換: 検証と評価研究

[Large Language Models for Automating Clinical Trial Criteria Conversion to Observational Medical Outcomes Partnership Common Data Model Queries: Validation and Evaluation Study.]

Kye Hwa Lee 1, MD, PhD; Sujung Jang 2, BEng; Grace Juyun Kim 3, PharmD, PhD; Sukyoung Park 2, MSc; Doeun Kim 2, MSc; Oh Jin Kwon 3, BSc; Jae-Ho Lee 4, MD, PhD; Young-Hak Kim 5, MD, PhD

1. Department of Information Medicine, Department of Digital Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
2. Department of Biomedical Engineering, AMIST, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
3. Big Data Research Center, Asan Institute for Life Sciences, Asan Medical Center, Seoul, Republic of Korea
4. Department of Emergency Medicine, Department of Information Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
5. Division of Cardiology, Department of Information Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea



大規模言語モデルを用いた臨床試験基準から観察研究用医療アウトカムパートナーシップ共通データモデルクエリへの自動変換: 検証と評価研究

[Large Language Models for Automating Clinical Trial Criteria Conversion to Observational Medical Outcomes Partnership Common Data Model Queries: Validation and Evaluation Study.]

1. 背景・課題

- 臨床試験は被験者リクルート不足により約半数が計画通り進行できない
- RWD(リアルワールドデータ)を用いた実施可能性評価(Feasibility Assessment)が重要
- 課題:
 - 適格基準は自然言語(自由記述)で書かれている
 - OMOP CDMのような構造化データに変換するには専門知識と工数が必要
 - LLMはハルシネーション(存在しない概念ID生成等)を起こしやすい

2. 研究目的

- ClinicalTrials.govの自由記述適格基準を OMOP CDM 対応SQLへ自動変換するエンドツーエンドシステムを構築
- 複数LLMを比較し、ハルシネーションの種類・頻度モデル選択・プロンプト設計の最適解を体系的に評価



大規模言語モデルを用いた臨床試験基準から観察研究用医療アウトカムパートナーシップ共通データモデルクエリへの自動変換: 検証と評価研究

[Large Language Models for Automating Clinical Trial Criteria Conversion to Observational Medical Outcomes Partnership Common Data Model Queries: Validation and Evaluation Study.]

3. システム全体像(3段階パイプライン)

① 前処理(Preprocessing)

分割: 適格基準を論理構造(AND/OR)を保ったまま分解

フィルタリング: 同意取得などDB検索不能な条件を除外

簡略化: 時間表現などを正規化

→トークン数を58.2%削減しつつ医学的意味を保持

② 情報抽出・概念マッピング

疾患、薬剤、検査、数値、否定、時間条件を抽出

GPT-4により SNOMED CT / RxNorm / LOINC など

へマッピング

③ SQL生成

OMOP CDM v5.3準拠SQLを生成

inclusionはINTERSECT、exclusionはEXCEPT等で統合



大規模言語モデルを用いた臨床試験基準から観察研究用医療アウトカムパートナーシップ共通データモデルクエリへの自動変換: 検証と評価研究

[Large Language Models for Automating Clinical Trial Criteria Conversion to Observational Medical Outcomes Partnership Common Data Model Queries: Validation and Evaluation Study.]

4. コンセプトマッピング性能(GPT-4 vs USAGI)

対象: 30試験・357用語

正解率: GPT-4: 48.5%、USAGI: 32.0% (有意差あり)

分野別: 薬剤: 72.7% (高精度)、

検査値: 38.3% (数値・単位解釈が課題)

→ GPT-4は文脈理解に優れるが、完全ではない

大規模言語モデルを用いた臨床試験基準から観察研究用医療アウトカムパートナーシップ共通データモデルクエリへの自動変換：検証と評価研究

[Large Language Models for Automating Clinical Trial Criteria Conversion to Observational Medical Outcomes Partnership Common Data Model Queries: Validation and Evaluation Study.]

5. SQL品質の専門家評価

● 評価観点：

- SQL文法
- CDMスキーマ適合
- 臨床的妥当性
- 7つの著名な臨床試験をOMOPと臨床に詳しい専門家2名が80項目にわたり評価

● 結果：

- 文法・スキーマ：ほぼ満点
- 臨床的文脈理解は不安定(特に時間条件・複雑条件)

6. 実データによるコホート検証(N3C基準と比較)

*N3C = National COVID Cohort Collaborative

適格条件	Jaccard	Overlap coefficient	評価
妊娠	0.00	0	完全失敗
2型糖尿病	0.03	1.0	階層展開不足
1型糖尿病	0.81	1.0	高一致
コントロール不良高血圧	0.18	0.48	部分的成功

→ 概念階層・粒度・単位変換が最大の弱点

Jaccard index (intersection/union) 2つのコホート間の全体的な類似度を対称的に評価する指標

Overlap coefficient (intersection/minimum set size) より小さいコホートがどの程度大きいコホートに含まれているかを示す指標



大規模言語モデルを用いた臨床試験基準から観察研究用医療アウトカムパートナーシップ共通データモデルクエリへの自動変換: 検証と評価研究

[Large Language Models for Automating Clinical Trial Criteria Conversion to Observational Medical Outcomes Partnership Common Data Model Queries: Validation and Evaluation Study.]

7. 大規模LLM比較(8LLMモデルx19試験x5種類プロンプト戦略、計760SQL)

- 有効SQL率(文法OK × 非ハルシネーション)
 - Llama3:8b(OSS):75.8% ← 最高
 - GPT-4:45.3%
- 重要な発見
 - モデルサイズ ≠ 実用性能
 - GPT-4は生成力は高いが、プレースホルダや誤概念が多い
 - 小型OSSモデルは安価かつ安定

8. ハルシネーションの内訳

- 最多:
 - ドメイン誤り(疾患を検査テーブルに置く等)
 - プレースホルダ挿入
- 非存在IDは少数だが致命的

9. プロンプト設計の示唆

- Zero-shotが最も安定
- 手順を細かく書きすぎると逆に性能低下
- モデル特性に応じた設計が必須



大規模言語モデルを用いた臨床試験基準から観察研究用医療アウトカムパートナーシップ共通データモデルクエリへの自動変換: 検証と評価研究

[Large Language Models for Automating Clinical Trial Criteria Conversion to Observational Medical Outcomes Partnership Common Data Model Queries: Validation and Evaluation Study.]

10. 結論・示唆

- LLMにより、適格基準SQL化を「数時間 → 数分」に短縮可能
 - ・ ただし:ハルシネーション率21-50%
 - ・ 単独自動化は危険
- 推奨アプローチ:
 - ・ LLM+ルールベース(単位変換、階層展開)
 - ・ 小型OSSモデル+厳密バリデーション#



デジタル病院における電子カルテ(EMR)データ抽出・検証プロセスの最適化—手法とアプローチに関するシステマティックレビュー

[Streamlining electronic medical record data extraction and validation in digital hospitals: A systematic review to identify optimal approaches and methods]

Han Chang Lim 1, Howard Wong 1, Reji Philip 1, Anton Van Der Vegt 1,2,
Kim-Kwang Raymond Choo 3, Jason D. Pole 1, Clair Sullivan 1,4

1. Queensland Digital Health Centre, Centre for Health Services Research, Faculty of Medicine, The University of Queensland, Herston, Queensland, Australia
2. School of Information Technology and Electrical Engineering, Faculty of Engineering, Architecture and Information Technology, The University of Queensland, Brisbane, Australia
3. Department of Information Systems and Cyber Security, The University of Texas at San Antonio, San Antonio, Texas, USA
4. Metro North Hospital and Health Service, Department of Health, Queensland Government, Herston, Queensland, Australia



デジタル病院における電子カルテ(EMR)データ抽出・検証プロセスの最適化—手法とアプローチに関するシステマティックレビュー

[Streamlining electronic medical record data extraction and validation in digital hospitals: A systematic review to identify optimal approaches and methods]

背景と問題意識

- EMRはデータを蓄積しているが、多くの場合「電子化された紙」としてしか使われていない
- データは複雑なテーブル構造に分散しており、二次利用(研究・品質改善)には高度な専門知識と多大な手作業が必要
- 抽出・検証プロセスは、時間がかかる、属人化しやすい、システム改修で壊れやすい
- これらの課題から、共通データモデル(CDM)による標準化が注目されている

OMOP CDM を採用した理由

- 複数のCDM(PCORnet, Sentinel, CDISC等)を比較した先行研究でOMOP CDMが柔軟性・統合性・実装容易性の点で最も優れていた
- 世界80か国以上・3000人超の専門家による国際的標準
- OMOP化により、解析ツールの再利用、組織間データ共有分析の再現性が向上する



デジタル病院における電子カルテ(EMR)データ抽出・検証プロセスの最適化—手法とアプローチに関するシステマティックレビュー

[Streamlining electronic medical record data extraction and validation in digital hospitals: A systematic review to identify optimal approaches and methods]

方法(レビュー設計)

- 対象期間: 2017年6月～2022年12月
- データベース: PubMed, Embase, Web of Science + グレイリテラチャ
- 条件:
 - ・ 病院のEMR/EHRデータ
 - ・ OMOP CDMへの変換を実施
- 最終的に 9本の論文を分析対象とした

主な課題(抽出・検証プロセス)

レビューから、以下の課題が特に頻出した:

1. 標準化されたプロセスの欠如
2. データ構造の不統一
3. 専門人材(ドメイン知識+IT)の不足
4. 用語マッピングの手作業負荷
5. 時間とコストの大きさ
6. 自動化不足
7. システム変更による保守困難性



デジタル病院における電子カルテ(EMR)データ抽出・検証プロセスの最適化—手法とアプローチに関するシステマティックレビュー

[Streamlining electronic medical record data extraction and validation in digital hospitals: A systematic review to identify optimal approaches and methods]

有効とされた手法・ツール

多くの研究で高い成功率を示したのは以下の組み合わせ:

- WhiteRabbit + RabbitInAHat
 - 構造化データの解析・ETL設計
 - 成功率:82~98%
- Usagi
 - 用語(語彙)マッピング
 - 成功率:98%以上
- CLAMP / SOCRATex(NLP)
 - 非構造化テキスト(診療録文章)抽出
 - 成功率:約99%

*CLAMP Clinical Language Annotation, Modeling, and Processing Toolkit, SOCRATex(NLP) Staged Optimization of Curation, Regularization, and Annotation of clinical text

- Achilles / Data Quality Dashboard
 - データ品質(完全性・妥当性・一貫性)の検証

結論

- 単一ツールでは不十分
- 標準化されたCDM(OMOP)を軸に、複数ツールを組み合わせた統合的アプローチが有効
- 将来的には、
 - 実運用での検証
 - 人材育成・コスト評価
 - LLM(大規模言語モデル)活用が重要になる



多施設ネットワークにおけるデータ品質と適合性基準の 向上に向けた多面的なアプローチ

[A multifaceted approach to advancing data quality and fitness standards in multi-institutional networks]

Hanieh Razzaghi, PhD, MPH , Kimberley Dickinson, MS , Kaleigh Wieand, BS , Samuel Boss, BS , Hunter Weidlich, BS , Yungui Huang, PhD , Keith Morse, MD , Sujan Kumar Mutyala, MS , Jyothi Priya Alekapatti Nandagopal, MS , Karthik Viswanathan, MS , Christopher B Forrest, MD, PhD , L Charles Bailey, MD, PhD

研究の背景と目的

- 複数機関にまたがる医療データネットワークでは、「データ品質(Data Quality)」と「研究目的への適合性(Fitness for Use)」を同時に評価できる仕組みが不足している。
- 本研究の目的は、多様な研究目的・利用シーンに対応可能な、包括的データ品質評価システムを構築すること。



多施設ネットワークにおけるデータ品質と適合性基準の向上に向けた多面的なアプローチ

[A multifaceted approach to advancing data quality and fitness standards in multi-institutional networks]

方法(Materials and Methods)

著者らは、既存のデータ品質プログラムの知見やネットワーク研究の実務要件を踏まえ、以下の5つの標準原則を定義し、それに基づくモジュール型DQシステムを設計。

1. 注釈と文書化(Annotation & Documentation)

- チェック内容・結果が明確に説明可能であること

2. 研究準備状況の評価(Research Readiness)

- データが研究に使える状態かを判断できること

3. ネットワーク間での再現性(Reproducibility)

- 複数機関・複数ネットワークで同様に適用できること

4. ユーザーの柔軟性(Flexibility)

- 研究目的や利用者に応じて調整可能であること

5. 解釈可能性(Interpretability)

- 研究者・プロジェクトチームが直感的に理解できること



多施設ネットワークにおけるデータ品質と適合性基準の向上に向けた多面的なアプローチ

[A multifaceted approach to advancing data quality and fitness standards in multi-institutional networks]

結果(Results)

- 10種類のデータ品質チェックタイプを定義
 - 合計 444以上のチェック項目を実装
- 2つの多施設ネットワークに実際に導入・運用
- チェック内容は以下をカバー：
 - ・ データモデルへの構造的整合性
 - ・ 一般的な研究での利用可能性
 - ・ 研究固有ニーズへのカスタマイズ
- 各チェックは相互依存なしで独立して利用可能
- ダッシュボードにより結果を可視化
 - ・ データソース数、マスキングの有無、利用者の関心領域に応じて表示を調整可能
- OMOP CDMを用いるデータにはそのまま適用可能
(他モデルにも容易に拡張可能)



多施設ネットワークにおけるデータ品質と適合性基準の向上に向けた多面的なアプローチ

[A multifaceted approach to advancing data quality and fitness standards in multi-institutional networks]

考察(Discussion)

- 従来研究を発展させ、「ネットワーク全体のデータ改善」+「研究利用」の両立を実現
- 特定研究に特化せず、汎用的・展開可能なDQ評価基盤として設計した点が新規性

結論(Conclusion)

- 本システムは、将来的なデータ品質評価の標準化・機関間連携に不可欠な構成要素を提供
- 臨床データを「意図した用途に適した形」にするための実践的かつ拡張可能なアプローチである



成人てんかん患者におけるオクサカルバゼピン誘発性低ナトリウム血症の予測: 実臨床CDMデータを用いた多施設機械学習解析

[Predicting oxcarbazepine-induced hyponatremia in adult epilepsy patients: A multicenter machine learning analysis using real-world CDM data]

Gucheol Jung a, JaeHyeok Lee a, Sung-Min Gho a, YoungMi Han b, ByungKwan Choi c, Jae Wook Cho d,e, Jiyoung Kim d,f, Gha-hyun Lee d,f

- a. Medical R&D Center, Deepnoid, Inc., Seoul, Republic of Korea
- b. Visual Terminology Co., Ltd., Pusan, Republic of Korea
- c. Department of Neurosurgery, Pusan National University Hospital, Pusan, Republic of Korea
- d. Pusan National University School of Medicine, Research Institute for Convergence of Biomedical Science and Technology, Yangsan, South Korea
- e. Department of Neurology, Pusan National University Yangsan Hospital, 50612 Geumoro 20, Yangsan, South Korea
- f. Department of Neurology, Pusan National University Hospital, Busan, South Korea

1. 研究背景と目的

背景: 抗てんかん薬オキシカルバゼピン(OXC)は低ナトリウム血症を引き起こすことが知られており、重症例では転倒、神経障害、死亡リスクがある。

課題: OXC処方時の低ナトリウム血症リスクを事前に高精度で予測する手法が不足。

目的: 多施設のリアルワールドデータ(OMOP-CDM)を用い、機械学習(ML)でOXC誘発性重症低ナトリウム血症を予測し、主要リスク因子を明らかにする。



成人てんかん患者におけるオクサカルバゼピン誘発性低ナトリウム血症の予測：実臨床CDMデータを用いた多施設機械学習解析

[Predicting oxcarbazepine-induced hyponatremia in adult epilepsy patients: A multicenter machine learning analysis using real-world CDM data]

2. 研究デザイン

- 研究タイプ:後ろ向きコホート研究
- データソース:韓国の三次医療機関2施設(PNUH、PNUYH)
- 対象:18歳以上でOXC投与歴があり、血清Na測定があるてんかん患者
- 重症低Na定義:血清Na ≤ 128 mmol/L(OXC投与日同日)
- 症例数:2,253例(重症低Na 8.4%)

3. 使用データと特徴量

- 人口統計:年齢、性別
- 臨床情報:脳卒中、脳手術、救急受診など
- 薬剤情報:
 - 抗てんかん薬(バルプロ酸、レベチラセタム等)
 - 利尿薬、降圧薬(Ca拮抗薬、 β 遮断薬など)
 - 睡眠薬、抗精神病薬
- 特徴:年齢・OXC用量は連続変数、他は二値化



成人てんかん患者におけるオクサカルバゼピン誘発性低ナトリウム血症の予測: 実臨床CDMデータを用いた多施設機械学習解析

[Predicting oxcarbazepine-induced hyponatremia in adult epilepsy patients: A multicenter machine learning analysis using real-world CDM data]

4. 機械学習モデルと検証

- 実験1: 単施設学習モデル
 - 学習: PNUH
 - 検証: 内部検証(PNUH)、外部検証(PNUYH)
 - モデル: XGBoost
- 実験2: 多施設学習モデル
 - 学習: 2施設統合データ
 - 比較モデル: ロジスティック回帰、ナイーブベイズ、SVM、ランダムフォレスト、XGBoost

5. 予測性能(要点)

- 最良モデル: XGBoost
- AUROC: 約0.83(多施設学習)
- NPV: >0.95(低リスク患者の除外に有用)
- PPV: 0.2前後(稀な事象のため限定的)
- 解釈: 高リスク患者を拾うより、低リスク患者を安全に除外する用途に適する



成人てんかん患者におけるオクサカルバゼピン誘発性低ナトリウム血症の予測: 実臨床CDMデータを用いた多施設機械学習解析

[Predicting oxcarbazepine-induced hyponatremia in adult epilepsy patients: A multicenter machine learning analysis using real-world CDM data]

6. SHAP解析による主要リスク因子

- 一貫して重要だった因子:

- ① 利尿薬併用
- ② バルプロ酸併用
- ③ OXC高用量
- ④ 高齢
- ⑤ 救急受診歴

- その他の寄与因子:

- Ca拮抗薬、 β 遮断薬
- 睡眠薬、抗精神病薬
- 他の抗てんかん薬(レバチラセタム、プレガバリン、ラコサミド)
- 脳卒中既往

7. 臨床的解釈

- バルプロ酸:ADH分泌促進・水排泄低下 → OXCと相乗
- 利尿薬:Na喪失・体液量変動 → 低Naリスク増大
- 高齢:腎機能低下・多剤併用による感受性増大
- 多剤併用(ポリファーマシー)が重要な背景因子



成人てんかん患者におけるオクサカルバゼピン誘発性低ナトリウム血症の予測：実臨床CDMデータを用いた多施設機械学習解析

[Predicting oxcarbazepine-induced hyponatremia in adult epilepsy patients: A multicenter machine learning analysis using real-world CDM data]

8. 臨床応用の意義

- 一律のNaモニタリングではなく、患者別リスク層別化が可能
- 低リスク患者：OXC継続＋最小限モニタリング
- 高リスク患者：電解質頻回チェック、代替治療薬検討
- OMOP-CDM基盤のため多施設展開

9. 限界と今後の課題

- 後ろ向き研究(因果推論に限界)
- PPVが低く、単独アラート用途には調整が必要
- 2施設・単一国データ
- 将来：縦断検査値の導入、より多施設・国際検証、EHR組み込み型CDS評価

10. まとめ(Take Home Message)

- XGBoost+多施設CDMデータでOXC低Naリスクを高精度予測
- SHAPにより臨床的に納得感のあるリスク因子を可視化
- 低Naの“見逃し防止”に強く、個別化治療・安全管理に有用



転帰予測の改善のための NIHSS と請求データを統合した ハイブリッド脳卒中品質指標の評価

[Evaluation of hybrid stroke quality indicators by integrating NIHSS and claims data for improved outcome prediction.]

Thomas Datzmann 1, Caroline Lang 1, Falko Tesch 1, Melissa Spoden 2, Patrik Dröge 2, Franz Ehm 1, Ekkehard Schuler 3, Christos Krogias 4, Christian Günster 2, Jochen Schmitt 1, Christoph Gumbinger 5,7 & Jessica Barlinn6,7

1. Center for Evidence-Based Healthcare, Medical Faculty and University Hospital Carl Gustav Carus, TUD Dresden University of Technology, Dresden, Germany.
2. Wissenschaftliches Institut der AOK (AOK Research Institute), Berlin, Germany.
3. Helios Kliniken GmbH, Berlin, Germany.
4. Department of Neurology, Evangelisches Krankenhaus Herne, Academic Teaching Hospital of the Ruhr University Bochum, Herne, Germany.
5. Department of Neurology, Heidelberg University Hospital, Heidelberg, Germany.
6. Department of Neurology, University Hospital and Faculty of Medicine Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany.
7. Christoph Gumbinger and Jessica Barlinn contributed equally to this work.



転帰予測の改善のための NIHSS と請求データを統合した ハイブリッド脳卒中品質指標の評価

[Evaluation of hybrid stroke quality indicators by integrating NIHSS and claims data for improved outcome prediction.]

研究背景(なぜこの研究が必要か)

- 脳卒中医療の質評価は重要だが、レセプトデータ(請求データ)だけでは限界がある。
- 死亡率などの従来指標は、治療プロセスの質や患者重症度を十分に反映できない。
- 特に、入院時の重症度差(ケースミックス)を考慮しないと、病院比較が歪む可能性がある。

研究目的

- レセプトデータに臨床的重症度指標(NIHSS)を統合した「ハイブリッド品質指標(Hybrid QI)」を開発。
 - *National Institutes of Health Stroke Scale (NIHSS)
- 予後予測精度および病院ランキングへの影響を評価する。



転帰予測の改善のための NIHSS と請求データを統合したハイブリッド脳卒中品質指標の評価

[Evaluation of hybrid stroke quality indicators by integrating NIHSS and claims data for improved outcome prediction.]

データと方法

● 対象患者

- 虚血性脳卒中(I63):9,348例
- 出血性脳卒中(I61):1,554例
- NIHSSが利用可能:5,012例(9/15病院)

● 使用データ

- レセプトデータ:年齢、性別、併存症、投薬、介護度など
- 臨床データ:入院時NIHSS(脳卒中重症度スコア)

● 主要アウトカム(すべて二値)

- 30日死亡、90日以内再発(再入院)、180日以内の介護度悪化

● 解析手法

- 機械学習モデル:XGBoost
- NIHSSあり/なしモデルを比較
- 評価指標:ROC-AUC、Brier Score



転帰予測の改善のための NIHSS と請求データを統合したハイブリッド脳卒中品質指標の評価

[Evaluation of hybrid stroke quality indicators by integrating NIHSS and claims data for improved outcome prediction.]

主な結果

① 予測精度の改善

- すべてのアウトカムで、NIHSSを含むモデルの方が高精度。
- 特に30日死亡で顕著(ROC-AUC 最大 +15%)。

② NIHSSの重要性

- NIHSSは、30日死亡:最重要因子、90日再発:最重要因子、180日介護度悪化:2番目に重要
- レセプトデータのみでは「年齢」が最重要因子だったが、→ NIHSSの影響は年齢より大きい。

③ 病院ランキングへの影響

- NIHSSを入れることで、病院の標準化死亡比(SMR)が大きく変動。
- 特に30日死亡では、NIHSSあり/なしでランキング相関が低い
→ 重症度を考慮しない評価は病院評価を誤るリスク



転帰予測の改善のための NIHSS と請求データを統合したハイブリッド脳卒中品質指標の評価

[Evaluation of hybrid stroke quality indicators by integrating NIHSS and claims data for improved outcome prediction.]

解釈・意義

- レセプトデータ単独の質指標は不十分。
- NIHSSを加えることで、リスク調整が改善、公平な病院比較が可能。
- 死亡率のみを質指標とすることへの強い警鐘。

限界

- NIHSSが取得できたのは15病院中9病院のみ。
- データ連携の問題(相互運用性不足)。
- 検査値や治療プロセス指標(Door-to-Needle time等)は未記載。

実務・政策的示唆

- 脳卒中医療の質評価には、重症度(NIHSS)を必須項目として収集・共有すべき。
- ハイブリッドQIは、病院評価の精度向上、資源配分の適正化、医療の質改善・予防戦略に寄与。

まとめ

- NIHSSをレセプトデータに統合すると、予後予測精度が大幅向上、病院ランキングが大きく変わる。
- 重症度を考慮しない死亡率評価は危険。
- 今後の脳卒中医療の質評価には、「ハイブリッド指標」が不可欠。



Kineret:イスラエル最大の病院ネットワーク 共同研究のためのOMOP共通データモデルへの変換

[Kineret: Israel's Largest Hospital Network Transformed into the OMOP common data model for collaborative research.]

Nadav Rappoport 1,2,3, Guy Livne 2, Naama Perry Cohen 2, Nir Makover 2, Hadas Eshel-Geva 2, Hadar Kapach 2, Tomer Hadad 2, Yarin Alon 2, Robyn Rubin 2, Segev Chai 2, Shirell da Villa 2, Ohad Hochman 4

1. Department of Software and Information Systems Engineering, Ben-Gurion University of the Negev, Beer Sheva, Israel
2. The Directorate of Government Medical Centers, State of Israel Ministry of Health, Jerusalem, Israel
3. The Data Science Research Center (DSRC), Ben-Gurion University of the Negev, Beer Sheva, Israel
4. Bnai-Zion Medical Center, Haifa, Israel



Kineret:イスラエル最大の病院ネットワーク 共同研究のためのOMOP共通データモデルへの変換

[Kineret: Israel's Largest Hospital Network Transformed into the OMOP common data model for collaborative research.]

研究背景

- イスラエル保健省(MOH)は、25の政府系医療機関(急性期・精神科・老年医療)を管轄。
- 医療データは施設ごとに異なるEHRや用語体系を用いており、多施設研究・二次利用が困難という課題があった。
- 研究者はデータ抽出をIT部門に依存し、研究開始までに長い時間を要していた。
→ これを解決するため、Kineret構想(国家規模データレイク)を立ち上げ。

研究目的

- 全国の公立病院データを OMOP CDM に統一。
- データ抽出・標準化・共有を効率化。
- 国内外の多施設共同研究を加速する研究基盤を構築すること。



Kineret:イスラエル最大の病院ネットワーク 共同研究のためのOMOP共通データモデルへの変換

[Kineret: Israel's Largest Hospital Network Transformed into the OMOP common data model for collaborative research.]

方法(Methods)

- 採用したデータモデル

- OMOP Common Data Model v5.3
- 国際的に普及(OHDSIネットワーク)
- 観察研究・多施設比較に強い標準モデル

- ETL(Extract-Transform-Load)プロセス

特徴的な点:

- オンプレミスで匿名化
患者IDはハッシュ化、氏名・住所・電話番号などは完全削除
- クラウド(AWS)で変換・統合
AirflowでETLを自動制御、PostgreSQL形式のOMOP CDMに格納

- フルリロード方式

差分更新ではなく、定期的に全件再変換用、語辞書(Athena)の更新や後日修正に対応、毎回新しい疑似IDを生成し、プライバシーを強化



Kineret:イスラエル最大の病院ネットワーク 共同研究のためのOMOP共通データモデルへの変換

[Kineret: Israel's Largest Hospital Network Transformed into the OMOP common data model for collaborative research.]

用語標準化(最大の技術課題)

- ICD-9-CM+イスラエル独自拡張(IC)
- ヘブライ語記載の診断・処置・検査・薬剤
- 各病院独自コードの存在

対応策:

- 言語モデルによる自動マッピング(約85%、精度 >0.9)
- 残り15%は専門家(医師・薬剤師・検査技師)が手動対応約
- 64,000コード → 33,000標準概念 に整理

データ規模と成果(Results)

- 統合状況(2025年時点)
 - 6病院が完全統合、7病院目が統合準備中
- データ規模(OMOPテーブル例)
 - 患者数:約 683万人、診断:約 7,000万件、処置:約 8,300万件、検査(測定):13.5億件、受診:約 5,200万件
- 受診タイプ内訳
 - 外来:約55%、救急:約28%、入院:約16%



Kineret:イスラエル最大の病院ネットワーク 共同研究のためのOMOP共通データモデルへの変換

[Kineret: Israel's Largest Hospital Network Transformed into the OMOP common data model for collaborative research.]

研究利用の仕組み(ワークフロー)

- ATLASを用いた研究支援
 - ・ 研究者はGUIで、コホート定義、実現可能性評価(Feasibility)、特性分析
- 1つのクエリで、全病院を横断検索

IRB・研究実施の流れ

1. ATLASで事前検討
2. IRB申請(単一IRB+他施設承認に簡素化)
3. 承認後、専用の仮想研究環境(VM)を提供
4. データ持ち出し不可、解析のみ可能

意義・インパクト(Discussion)

- 主な効果
 - ・ 研究開始までの時間を大幅短縮
 - ・ IT部門依存から研究者主導へ
 - ・ 国内多施設研究が現実的に
 - ・ 国際研究(OHDSI / EHDEN)への参加が容易に
- 制約・課題
 - ・ OMOPは構造化データ中心で画像、PDF、心電図などは別管理
 - ・ 用語マッピングの人的コスト
 - ・ フルリロード方式による計算負荷



Kineret:イスラエル最大の病院ネットワーク 共同研究のためのOMOP共通データモデルへの変換

[Kineret: Israel's Largest Hospital Network Transformed into the OMOP common data model for collaborative research.]

結論(Conclusion)

- Kineretは国家規模のOMOP CDM実装事例
- 医療データ標準化と研究促進の成功モデル
- 他国・他機関への再現可能なブループリント
- 今後は:NLP(自由記載)機械学習国際分散研究への拡張



思春期特発性側弯症(AIS)に対する計算可能フェノタイプの開発および妥当性検証

[Development and validation of a computable phenotype for adolescent idiopathic scoliosis.]

Sarah B. Floyd 1,2, Ashley Mills 2, Jason Woloff 2, Christian Lawson 2, Coleman Hilton 3, Donna Oeffinger 4, Steven Hwang 5

1. Department of Public Health Sciences, Clemson University, Clemson, South Carolina, USA
2. Department of Research Programs, Shriners Children's Headquarters, Tampa, Florida, USA
3. Department of Data Engineering and Analytics, Shriners Children's Headquarters, Tampa, Florida, USA
4. Shriners Children's, Lexington, Kentucky, USA
5. Shriners Children's, Philadelphia, Pennsylvania, USA

研究背景・目的

- 思春期特発性側弯症(AIS)は小児側弯症の80-90%を占めるが、原因(病因)や治療効果に関するエビデンスが不足している。
- 電子カルテ(EHR)などのリアルワールドデータを用いた大規模観察研究を可能にするため、AIS患者を高精度に同定できる「計算可能フェノタイプ(Computable Phenotype:CP)」の開発と検証を目的とした。



思春期特発性側弯症(AIS)に対する計算可能フェノタイプの開発および妥当性検証

[Development and validation of a computable phenotype for adolescent idiopathic scoliosis.]

研究デザイン

- Shriners Children'sの研究データウェアハウス(RDW)を使用
- OMOP Common Data Model(CDM)に変換されたEHRデータ(2016-2022年)を対象
- AISを同定するための4種類のCPアルゴリズムを構築
- 全国前向き側弯症レジストリ(SSS registry)との照合で性能評価
- 最良のCPをカルテレビュー(人手確認)で妥当性検証

フェノタイプ定義の考え方

- 対象:10-16歳、Cobb角 $>10^\circ$ 、原因不明の側弯
- 含める条件:AISまたは非特異的側弯症の診断コード、必要に応じてX線/CT/MRI実施履歴
- 除外条件:神経筋性・先天性・症候性側弯症、乳児・若年性側弯症、Mehtaキャスト歴(若年発症を示唆)



思春期特発性側弯症(AIS)に対する計算可能フェノタイプの開発および妥当性検証

[Development and validation of a computable phenotype for adolescent idiopathic scoliosis.]

主な結果

- 4つのCPで約1.5万~2.4万人のAIS候補患者を抽出
- 最終採用CP(AIS 2)
 - 条件:AIS診断1回以上 +180日以内の脊椎画像検査(X線/CT/MRI)
 - 感度:92.7%
 - 特異度:81.8%
- カルテレビュー(50例)
 - 真のAIS:36例(72%)
 - 偽陽性:14例(主に脊柱非対称や若年性側弯症)

AIS患者集団の特徴

- 平均年齢:13.6歳
- 女性:73.7%
- 人種:白人が約56%
- 北カリフォルニア、Shreveport、カナダ拠点で症例が多い



思春期特発性側弯症(AIS)に対する計算可能フェノタイプの開発および妥当性検証

[Development and validation of a computable phenotype for adolescent idiopathic scoliosis.]

結論

- 開発したAISの計算可能フェノタイプは実用的な精度を示した
- 将来的には、X線パラメータ(Cobb角など)の直接利用、自然言語処理(NLP)による非構造化データ活用により精度向上が期待される
- 本CPはLearning Health System(学習する医療システム)構築の基盤となる



集中治療医療のための OMOP 共通データモデルの拡張: 米国集中治療医学会 (SCCM) の集中治療データ辞書 (C2D2) を用いた、複雑な ICU データ標準化のためのフレームワーク

[Extending the Observational Medical Outcomes Partnership (OMOP) Common Data Model for Critical Care Medicine:
A Framework for Standardizing Complex ICU Data Using the Society of Critical Care Medicine's Critical Care Data Dictionary (C2D2).]

Meredith C. B. Adams , MD, MS 1, Robert W. Hurley , MD, PhD 2,3, Karsten Bartels, MD, PhD, MBA 4, Matthew L. Perkins, BS 5, Cody Hudson, MS 6, Umit Topaloglu, PhD 7,8, J. Perren Cobb, MD, FCCM 9, Karin Reuter-Rice, PhD, NP 10, Jacqueline C. Stocking, PhD, MBA, MSN 11, Ashish K. Khanna, MD, MS 12,13

1. Departments of Anesthesiology, Artificial Intelligence, Translational Neuroscience, and Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC.
2. Department of Anesthesiology, Translational Neuroscience, and Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC.
3. Pain Outcomes Lab, Wake Forest University Health Sciences, Winston-Salem, NC.
4. Anesthesiology, Psychiatry, and Learning Health Sciences, University of Michigan, Ann Arbor, MI.
5. Department of Comprehensive Cancer Center, Wake Forest University School of Medicine, Winston-Salem, NC.
6. Department of Comprehensive Cancer Center, Wake Forest University School of Medicine, Winston-Salem, NC.
7. Department of Cancer Biology, Wake Forest University School of Medicine, Winston-Salem, NC.
8. National Cancer Institute, Bethesda, MD.
9. Departments of Surgery and of Anesthesiology, Keck School of Medicine of USC, Los Angeles, CA.
10. Departments of Pediatrics and Neurosurgery, Division of Critical Care Medicine, Duke University, School of Nursing, School of Medicine, Durham, NC.
11. Department of Internal Medicine, Division of Pulmonary, Critical Care, and Sleep Medicine, University of California Davis Health, Sacramento, CA.
12. Department of Anesthesiology, Division of Critical Care Medicine, Wake Forest School of Medicine, Atrium Health Wake Forest Baptist Medical Center, Winston-Salem, NC.
13. Outcomes Research Consortium, Houston, TX.



集中治療医療のための OMOP 共通データモデルの拡張: 米国集中治療医学会 (SCCM) の集中治療データ辞書 (C2D2) を用いた、複雑な ICU データ標準化のためのフレームワーク

[Extending the Observational Medical Outcomes Partnership (OMOP) Common Data Model for Critical Care Medicine: A Framework for Standardizing Complex ICU Data Using the Society of Critical Care Medicine's Critical Care Data Dictionary (C2D2).]

背景

- 医療データ標準化は多施設研究・リアルワールドデータ解析に不可欠
- OMOP CDMは観察研究向けの国際標準モデル
- ICUデータは時間的に高頻度・複雑・専門的で、一般的な標準では表現しにくい
- SCCM(米国集中治療医学会)は、専門家合意に基づく Critical Care Data Dictionary(C2D2) を策定

→ 本研究は C2D2をOMOPにマッピング可能かを体系的に

評価

研究目的

- C2D2の各データ要素がOMOP CDMとどの程度互換性を持つかを評価
- ICU特有データを標準化する上での技術的・構造的課題を明確化

集中治療医療のための OMOP 共通データモデルの拡張: 米国集中治療医学会 (SCCM) の集中治療データ辞書 (C2D2) を用いた、複雑な ICU データ標準化のためのフレームワーク

[Extending the Observational Medical Outcomes Partnership (OMOP) Common Data Model for Critical Care Medicine: A Framework for Standardizing Complex ICU Data Using the Society of Critical Care Medicine's Critical Care Data Dictionary (C2D2).]

研究デザイン・方法

- 対象: C2D2 の 226 データ要素
- 方法:
 - 大規模言語モデル (LLM) を用いた自動セマンティックマッチング
 - 3段階評価:
 - 完全一致 (Full match): OMOP に直接対応
 - 部分一致 (Partial match): 修正・拡張が必要
 - 非対応 (No match): 対応概念なし
 - 臨床情報学専門家による二重レビューで検証 #

主な結果 (定量)

- 完全一致: 49.6% (112 項目)
- 部分一致: 46.4% (105 項目)
- 非対応: 4.0% (9 項目)

→ 約 96% の ICU 概念は OMOP に何らかの形で対応可能

LLM マッチング性能

- 適切な OMOP 候補を提示できた割合: 96%
- 精度・再現率 (代表値):
 - Precision: 59.5% (厳格条件) / 97.2% (許容条件)
 - Recall: 87.0%
- AUROC: 0.786

→ 人手作業を大幅に削減できる実用レベル

集中治療医療のための OMOP 共通データモデルの拡張: 米国集中治療医学会 (SCCM) の集中治療データ辞書 (C2D2) を用いた、複雑な ICU データ標準化のためのフレームワーク

[Extending the Observational Medical Outcomes Partnership (OMOP) Common Data Model for Critical Care Medicine: A Framework for Standardizing Complex ICU Data Using the Society of Critical Care Medicine's Critical Care Data Dictionary (C2D2).]

主な非互換・課題領域

1. 重症度スコア
 - APACHE II / SOFA / PRISM などの複合スコア
 - OMOPは「計算済みスコア」を前提としていない。
2. 時間的精度
 - 「ICU入室後24時間の最大MAP」など
 - OMOPは時間窓指定が弱い
3. 年齢特異性
 - 小児専用スコア (PRISM など) の年齢区別が保持できない
4. 人工呼吸器・臓器サポート
 - 詳細設定・方向性 (high/low) の表現不足

解釈 (著者の主張)

- OMOP 拡張は技術的には可能
- 課題の多くは「概念が近いが少し足りない」レベル
- 問題は技術よりも、どこまで標準に寄せるか、誰が拡張を維持するか (OHDSI 連携)



集中治療医療のための OMOP 共通データモデルの拡張: 米国集中治療医学会 (SCCM) の集中治療データ辞書 (C2D2) を用いた、複雑な ICU データ標準化のためのフレームワーク

[Extending the Observational Medical Outcomes Partnership (OMOP) Common Data Model for Critical Care Medicine: A Framework for Standardizing Complex ICU Data Using the Society of Critical Care Medicine's Critical Care Data Dictionary (C2D2).]

戦略的選択肢(提案)

- OMOP完全統合: 研究連携重視(時間・コスト大)
- C2D2 v2.0開発: ICU特化を維持
- ハイブリッド型:
 - 使える部分(約50%)は即OMOP
 - 残りは拡張・独自管理

本研究の意義(まとめ)

- ICUデータ標準化は「不可能」ではない
 - 約半分は即座にOMOP利用可能
 - 残りも段階的拡張で対応可能
 - LLMはデータ標準化の強力な実務ツール
- 今後は技術ではなくコミュニティ合意と戦略判断が鍵



北欧諸国における共通データモデルの現状利用

[Current Use of Common Data Models in the Nordic Countries.]

Gerard Ompad, Carolyn E. Cesta, Jacqueline M. Cohen, Maarit K. Leinonen, Heidi Taipale, Huiqi Li, Lárus S. Guðmundsson, Mika Gissler, Maurizio Sessa

研究目的(Objectives)

- 北欧諸国(デンマーク、フィンランド、アイスランド、ノルウェー、スウェーデン)において、どのようなCDMが導入されているか、どのように活用されているか、その利点と限界は何かを体系的に把握することを目的とした。

研究方法(Methods)

- 対象: Nordic PharmacoEpidemiological Network (NorPEN)に所属する研究グループ
- 方法: 質問票調査(アンケート)
- 回答内容: 使用しているCDMの種類、国際・国内共同研究での利用状況、CDMの利点・制約に関する認識
- 解析: 集計データに基づく記述統計(個別回答者の匿名性を保持)



北欧諸国における共通データモデルの現状利用

[Current Use of Common Data Models in the Nordic Countries.]

主な結果(Results)

- 回答状況
 - 25研究グループ中 22グループが回答(回答率88%)
 - 13グループ(59%)が1つ以上のCDMを使用
 - 全ての北欧諸国でCDM利用が確認された
- 利用されているCDM
 - OMOP CDM
 - 最も利用頻度が高い、主に学術研究機関で使用、国際共同研究との親和性が高い
 - 北欧独自・プロジェクト特化型CDM
 - Nordic CDM、NorPreSS、ConcePTION など、母子連結データ(母-子データ)を扱う研究で特に有用

利点

- 医療データの標準化が進み、国・機関を跨いだ比較が容易
- 特定研究目的に合わせて設計されたCDMは高い実用性を持つ
- 北欧の研究グループがグローバルな研究ネットワークに参加するための基盤として機能

限界・課題

- 多くのCDMは妊娠・周産期特有のデータ構造への対応が不十分
- CDM間の相互運用性(interoperability)が限定的
- 柔軟な拡張性に欠ける場合がある



北欧諸国における共通データモデルの現状利用

[Current Use of Common Data Models in the Nordic Countries.]

結論(Conclusions)

- 北欧諸国におけるCDM導入状況には国・研究グループ間で大きなばらつきがある
- OMOP CDMは国際共同研究に強みを持つ
- 一方、Nordic CDMやNorPreSSのような地域特化型CDMは、北欧特有の研究ニーズに適合している
- 今後は、妊娠・母子データへの対応強化、CDM間の相互運用性向上が、北欧内および国際研究におけるCDM活用の鍵となる



医療分野における人工知能のための AIR MSデータプラットフォーム

[The AIR-MS data platform for artificial intelligence in healthcare.]

Pablo Guerrero , PhD 1,2, Morten Ernebjerg , PhD 1, Thomas Holst , MSc 1, David Weese , PhD 1, Herve DiBello 3,4, Susanne Ibing , PhD 2,3,4, Linea Schmidt , MSc 2, Ryan Ungaro , MD 5, Bernhard Renard , PhD 2,3,4, Christoph Lippert , PhD 2,3,4, Eugenia Alleva , MD 3,4, Timothy David Quinn , PhD 6, Patricia Kovatch , BS 3,4, Esther-Maria Antao , PhD 2, Elmien Heyneke , PhD 2, Aadil Rasheed , MSc 2, Stefan Kalabakov , MSc 2,3,4, Bert Arnrich , PhD 2,3,4, Alexander Charney , MD, PhD 3,4,7, Lothar H. Wieler , PhD 2,3,4, Girish Nadkarni , MD, MPH3,4,7

1. D4L Data4Life gGmbH, Potsdam, Brandenburg 14482, Germany
2. Hasso Plattner Institute, Digital Engineering Faculty, University of Potsdam, Potsdam, Brandenburg 14482, Germany
3. Hasso Plattner Institute for Digital Health at Mount Sinai, Icahn School of Medicine at Mount Sinai New York, NY 10029-5674, United States
4. Windreich Department of Artificial Intelligence and Human Health, Icahn School of Medicine at Mount Sinai, New York, NY 10029-5674, United States
5. Division of Gastroenterology, Icahn School of Medicine at Mount Sinai, New York, NY 10029-5674, United States
6. Icahn School of Medicine at Mount Sinai, New York, NY 10029-5674, United States
7. The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, United States



医療分野における人工知能のための AIR MSデータプラットフォーム

[The AIR-MS data platform for artificial intelligence in healthcare.]

背景と課題

- 電子カルテ(EHR)、臨床ノート、医療画像、ゲノムなど医療データは急増
- AI活用には「大規模・多様・長期縦断データ」が不可欠
- 既存の課題：
 - ・ データ形式がバラバラ(構造化／非構造化)
 - ・ 横断検索や即時分析が困難
 - ・ プライバシー・規制対応が重い

AIR-MSプラットフォームの全体像

- 基本コンセプト：
「多様な臨床データ」+「高速クエリ」+「AI計算基盤」を1つの安全な研究環境に統合

主な特徴：

- OMOP-CDM準拠の共通データモデル
 - ・ インメモリ型DB(SAP HANA)による高速分析
 - ・ HPC(Minerva)による大規模AI学習
 - ・ 米国・ドイツの国際共同研究対応



医療分野における人工知能のための AIR MSデータプラットフォーム

[The AIR-MS data platform for artificial intelligence in healthcare.]

データ構成

1 構造化データ(OMOP-CDM)

- 患者数:1,200万人以上
- 診断、検査、処方、処置、来院履歴
- 臨床ノート全文(自然言語)

2 非構造化データ(メタデータ連携)

- 病理画像(デジタルスライド 約250万件)
- 放射線画像(DICOM)
- 心電図(ECG)
- ゲノム・バイオバンク情報

※ 生データは別ストレージ、メタデータをDBで統合

データ統合と品質管理

1 データ統合

- Epic EHR → Caboodle → OMOP-CDMへETL
- 約200万コード中、157万以上を標準語彙にマッピング

2 匿名化(De-identification)

- HIPAA準拠
- IDハッシュ化+日付シフト(SANT方式)
- 自由記載テキストは匿名化困難なため除外

3 データ品質

- 独自ルール+OHDSI Data Quality Dashboard
- 完全性・一貫性・変換忠実性を継続監視



医療分野における人工知能のための AIR MSデータプラットフォーム

[The AIR-MS data platform for artificial intelligence in healthcare.]

計算・アプリ基盤

1 計算環境

- Minerva HPC (High Performance Computing):
約25,000コア
GPUノードあり
2ペタフロップス級

2 AppTier(分析アプリ)

- Data2Evidence: 構造化データ探索
- Visian: 医療画像アノテーション
- TURTLE: 自然言語 → SQL 変換(LLM活用)

パフォーマンス検証

- OHDSI標準コホート11種で性能評価
- 9/11クエリが10秒以内に完了
- 処理時間は「データ量」より「クエリ複雑性」に依存
- 臨床ノート全文検索: 通常1秒未満#



医療分野における人工知能のための AIR MSデータプラットフォーム

[The AIR-MS data platform for artificial intelligence in healthcare.]

研究ユースケース

1 クローン病(Crohn's disease)表現型解析

- 構造化EHR+臨床ノートを統合
- ルールベース+LLM(GPT-4)による情報抽出
- 発症年齢・合併症抽出の精度向上

2 月経困難症と心疾患リスク

- 臨床ノート解析で月経困難症を同定
- 約56,800人・最大20年追跡の大規模コホート
- 若年女性における虚血性心疾患リスク因子を新規発見

3 PreCareML(心血管リスク予測)

- 5年以内の心血管イベント予測
- フェデレーテッドラーニングで国際共同学習
- 外部検証での汎化性能を評価

意義と他プラットフォームとの差別化

● AIR-MSの強み:

- 単一医療機関として最大級のOMOP-CDM DB
- マルチモーダル(EHR+画像+テキスト)統合
- インメモリDBによる対話的分析
- 国際・学際的AI研究を前提設計

まとめ(メッセージ)

- AI医療研究には「データ×計算×ガバナンス」の統合が不可欠
- AIR-MSはその実装例として高い完成度
- 他施設・他国への展開モデルとして参考価値が高い



形式の医療データのための共通データモデルとデータ標準： 系統的レビュー

[Common data models and data standards for tabular health data: a systematic review.]

Melissa Finster 1, Markus Wenzel 1,2 and Elham Taghizadeh 1

1. Department, Fraunhofer Institute for Digital Medicine MEVIS, Max-Von-Laue-Str. 2, 28359 Bremen, Germany
2. Department, Constructor University, Campus Ring 1, 28359 Bremen, Germany

研究背景(Background)

- 医療データの二次利用は、患者中心の医療、医薬品安全性評価、パンデミック対応などに不可欠。
- 施設・国境を越えたデータ共有には Common Data Model(CDM) と データ標準 が重要。
- 実際には、データ形式・用語・範囲の違いや、プライバシー保護が大きな障壁。
- FAIR原則(Findable, Accessible, Interoperable, Reusable)を満たすための基盤としてCDMとデータ標準が注目されている。

研究目的

- 医療データの調和(ハーモナイゼーション)に用いられる代表的な CDM と データ標準 を体系的に比較・評価する。
- それぞれの強み・弱みを明らかにし、単一モデルではなく「相互変換・相互運用」の重要性を示す。



#形式の医療データのための共通データモデルとデータ標準： 系統的レビュー

[Common data models and data standards for tabular health data: a systematic review.]

対象としたモデル・標準

1 CDM(主にデータ保存・分析用)

- Sentinel CDM (SCDM)
- PCORnet CDM
- i2b2 CDM
- OMOP CDM

2 データ標準(主にデータ取得・交換用)

- HL7 Version 2
- CDA (Clinical Document Architecture)
- FHIR (Fast Healthcare Interoperability Resources)
- openEHR

評価方法(Methods)

- Scopusを用いた体系的文献レビュー(2018–2024)。
- 医学分野に限定し、FAIR、相互運用性、CDM、データ標準に関連する論文を抽出。
- 既存レビュー論文およびEMAガイドラインを基に評価基準を設定。



#形式の医療データのための共通データモデルとデータ標準： 系統的レビュー

[Common data models and data standards for tabular health data: a systematic review.]

評価基準(5カテゴリ)

① 適合性(Suitability)

- 多様なデータソース(EHR、請求データ、検査、自由記述など)への対応力。

② 普及度(Popularity)

- 実運用での利用実績、ネットワークの広がり、ETL資産の豊富さ。

③ 適応性(Adaptability)

- 拡張のしやすさ、継続的な更新・保守体制。

④ 相互運用性(Interoperability)

- 標準用語(SNOMED CT、LOINC等)の利用、ガバナンス、データ品質検証。

(評価基準(5カテゴリ))

⑤ サポート(Support)

- 使いやすさ、ツール群、バージョン管理。



#形式の医療データのための共通データモデルとデータ標準: 系統的レビュー

[Common data models and data standards for tabular health data: a systematic review.]

主な結果(Results)

1 CDMの総合評価

- OMOP CDM: 総合評価が最も高い(幅広いデータ対応、強力なツールとネットワーク)
- PCORnet CDM: OMOPに次ぐ評価、ガバナンスが強い
- i2b2 CDM: 柔軟性・拡張性に優れる
- SCDM: 医薬品安全性に特化、米国中心

2 データ標準の総合評価

- FHIR: すべての評価カテゴリで最高スコア
- openEHR: FHIRに匹敵、複雑データや意思決定支援に強み
- CDA / HL7 v2: 歴史的に重要だが、近年はFHIRに劣後

考察(Discussion)

- 単一の「万能CDM／標準」は存在しない。
- OMOP CDM は研究用途・国際ネットワークに強いが、コンセプトに厳密に従うことで、複雑な関係性を有するデータのマッピングにおいて、精度が制限され、情報が失われる可能性がある。
- i2b2 はローカル拡張に柔軟だが、ガバナンスは弱め。
- FHIR はデータ交換・API・セキュリティ面で最も成熟。
- openEHR はデータ表現と意味を分離する設計で、意味的相互運用性に優れる。



#形式の医療データのための共通データモデルとデータ標準： 系統的レビュー

[Common data models and data standards for tabular health data: a systematic review.]

結論(Conclusion)

- グローバルに単一のCDMや標準へ統一することは現実的でない。
- 重要なのは CDM・データ標準間の変換(ETL、クエリ変換、レイヤ統合)を可能にすること。
- 複数モデルを併用し、FAIR原則を満たすエコシステムを構築することが、国際的医療研究の鍵。



単極性うつ病から双極性障害への診断移行リスクにおける選択的セロトニン再取り込み阻害薬とセロトニン・ノルアドレナリン再取り込み阻害薬の比較有効性

[Comparative Effectiveness of Selective Serotonin Reuptake Inhibitors versus Serotonin-Norepinephrine Reuptake Inhibitors in the Risk of Diagnostic Conversion from Unipolar Depression to Bipolar Disorder.]

Ka Hee Yoo, Kyung Joo Lee, Sang Min Lee, Changwoo Han, Rae Woong Park & Young Tak Jo

研究背景

- 抗うつ薬の使用、とくにSNRI(セロトニン・ノルアドレナリン再取り込み阻害薬)とSSRI(選択的セロトニン再取り込み阻害薬)の違いが、単極性うつ病から双極性障害へ診断移行(diagnostic conversion)されるリスクに影響するかどうかについては、これまで議論が続いてきた。
- 本研究は、この診断転換リスクとSSRI・SNRI使用との関連を明らかにすることを目的とした。

方法

- データソース:韓国版 OMOP-Common Data Model (CDM)
- 対象集団:
 - ・ 解析対象群:SNRIを処方された患者
 - ・ 比較対象群:SSRIを処方された患者
- 主要アウトカム:抗うつ薬開始から6か月以降に新たに双極性障害と診断された場合
- 解析方法:
 - ・ 傾向スコア(Propensity Score)調整
 - ・ 1:1および1:2のマッチング解析
 - ・ 分散ネットワーク解析(多施設データ統合)



単極性うつ病から双極性障害への診断移行リスクにおける選択的セロトニン再取り込み阻害薬とセロトニン・ノルアドレナリン再取り込み阻害薬の比較有効性

[Comparative Effectiveness of Selective Serotonin Reuptake Inhibitors versus Serotonin-Norepinephrine Reuptake Inhibitors in the Risk of Diagnostic Conversion from Unipolar Depression to Bipolar Disorder.]

結果

- 傾向スコア調整後、SSRI使用者とSNRI使用者の間で、双極性障害への診断移行リスクに有意差は認められなかった。
- 分散ネットワーク解析の結果：
 - 1:1マッチング：
ハザード比(HR) = 1.28
95%信頼区間: 0.90-1.82
→ 統計学的に有意ではない
 - 1:2マッチング：
HR = 1.16
95%信頼区間: 0.88-1.53
→ 統計学的に有意ではない

結論

- SSRIとSNRIの使用によって、単極性うつ病から双極性障害へ診断が移行するリスクに有意な差は認められなかった。

臨床的・実務的な示唆

- 抗うつ薬選択(SSRIかSNRIか)そのものが、双極性障害への診断移行リスクを大きく左右する可能性は低い。
- 薬剤クラスよりも、患者背景、既往歴、症状経過といった要因の評価が重要である可能性が示唆される。



単極性うつ病から双極性障害への診断転換リスクにおける選択的セロトニン再取り込み阻害薬とセロトニン・ノルアドレナリン再取り込み阻害薬の比較有効性

[Comparative Effectiveness of Selective Serotonin Reuptake Inhibitors versus Serotonin-Norepinephrine Reuptake Inhibitors in the Risk of Diagnostic Conversion from Unipolar Depression to Bipolar Disorder.]

限界と今後の課題

- 個々の抗うつ薬ごとの差異は検討されていない
- 多剤併用(ポリファーマシー)の影響は未評価
- 後ろ向き研究であり、前向き研究が今後必要



母体健康における人工知能アプリケーションのための共通データモデルの要約データ:コロンビアにおける経験報告

[Datos sintéticos de un modelo de datos común para las aplicaciones de inteligencia artificial en salud materna: reporte de experiencia en el contexto colombiano]

Ever Augusto Torres-Silva 1 , Juan José Gaviria-Jiménez 2 , Ana María Guevara-Zambrano 3 ,
Laura Herrera-Almanza 4 , José Flórez-Arango 5

1. Futuro, Netux S.A.S., Medellín, Colombia
2. Facultad de Medicina, Universidad de Antioquia, Medellín, Colombia
3. Ginecología y Obstetricia, Hospital Pablo Tobón Uribe, Medellín, Colombia
4. Investigación, Clínica Universitaria Bolivariana, Medellín, Colombia
5. Population Health Sciences, Weill Cornell Medicine, New York, USA

背景

- 母体保健分野では、電子カルテデータの不足、標準化の欠如、アクセス制限、プライバシー問題により、AI活用が困難。
- 特に妊婦データは国際的にも不足しており、低頻度だが重大な事象(妊産婦死亡など)のモデル化が難しい。
- 合成データ(Synthetic Data)は、これらの課題を補う有望な手段。

研究目的

- コロンビアの妊娠・周産期ケアプロセスを再現する完全合成データ生成モデルを構築。
- 生成データを OMOP Common Data Model(CDM)に変換し、母体保健AIアプリケーションへの活用と国際的な相互運用性を高める。



母体健康における人工知能アプリケーションのための共通データモデルの要約データ:コロンビアにおける経験報告

[Datos sintéticos de un modelo de datos común para las aplicaciones de inteligencia artificial en salud materna: reporte de experiencia en el contexto colombiano]

方法

- 研究デザイン: ケーススタディ(経験報告)
- ツール: Synthea™(オープンソースの合成患者生成ツール)
- 開発プロセス(3段階)
 - コロンビアの妊婦診療フロー設計(専門医レビュー)
 - Synthea™でのモデル構築・状態遷移設計
 - 合成データ生成+OMOP CDM v5.4へのETL変換
- 用語標準化
 - 診断: SNOMED CT、ICD-10
 - 処置: コロンビア国内コード体系

モデル内容(臨床的カバー範囲)

- 妊娠期別の主なアウトカム
 - 第1三半期: 流産、異所性妊娠
 - 第2三半期: 妊娠高血圧症、子癇前症、子癇
 - 第3三半期: 前置胎盤、常位胎盤早期剥離
 - 産後: 産後出血
- ワクチン接種(破傷風、インフルエンザ、百日咳)も含む
- 各事象の発生確率は文献または専門医判断で設定



母体健康における人工知能アプリケーションのための共通データモデルの要約データ:コロンビアにおける経験報告

[Datos sintéticos de un modelo de datos común para las aplicaciones de inteligencia artificial en salud materna: reporte de experiencia en el contexto colombiano]

結果

- 10,637人の妊婦合成データセットを生成
- 地理分布は国勢調査(DANE 2018)に基づき、自治体単位で再現
- 臨床アウトカムの頻度は、国内外文献の報告値と概ね整合
- 既存Synthea妊娠モデルより
 - 状態数:179 → 156(簡素化)
 - 臨床アウトカム数:増加(母体・胎児リスクを重視)

主な限界

- 妊産婦死亡はモデル上含めたが、実際には生成されなかった
- 喫煙・飲酒など生活習慣因子は未考慮
- 確率が固定的で、データの多様性が限定的
- 実データとの定量的検証が困難(国内データアクセス制限)



母体健康における人工知能アプリケーションのための共通データモデルの要約データ:コロンビアにおける経験報告

[Datos sintéticos de un modelo de datos común para las aplicaciones de inteligencia artificial en salud materna: reporte de experiencia en el contexto colombiano]

結論・意義

- コロンビアにおける妊娠ケアを対象とした合成データ生成は実現可能。
- OMOP CDM対応により、AI研究・国際比較・相互運用性への道を開く先駆的取り組み。
- 実データ不足を補完し、母体保健AI開発を加速させる基盤となる。



脂肪腫は代謝症候群の構成要素の高頻度と関連している： 多施設横断研究

[Lipomas Are Associated with a Higher Prevalence of Metabolic Syndrome Components: A Multicenter Cross-Sectional Study.]

Ori Berger 1, Shaked Menashe 2, Shiri Damti Geva 3, Reychel Yakubov 4, Maor Ben Yehuda 4, Mor Peleg 4 and Ran Talisman 1

1. Department of Plastic Surgery, Barzilai Medical Center, Ashkelon, Israel
2. Department of Plastic Surgery, Shamir Medical Center, Rishon LeZion, Israel
3. Department of Otolaryngology and Head and Neck Surgery, Galilee Medical Center, Nahariya, Israel
4. Department of Information Systems, University of Haifa, Haifa, Israel

研究背景

- 脂肪腫(lipoma)は最も頻度の高い良性脂肪性腫瘍で、これまで主に美容的・偶発的所見と考えられてきた。しかし、肥満や糖代謝異常を有する成人に多く認められることから、全身的な代謝異常の皮膚所見ではないかという仮説がある。

目的

- 脂肪腫を有する成人が、一般人口と比較してメタボリックシンドローム構成要素(肥満、脂質異常症、高血圧、2型糖尿病)を高頻度に有するかを、大規模・多施設データで検証することである。



脂肪腫は代謝症候群の構成要素の高頻度と関連している： 多施設横断研究

[Lipomas Are Associated with a Higher Prevalence of Metabolic Syndrome Components: A Multicenter Cross-Sectional Study.]

研究デザイン・方法

- 研究デザイン: 多施設・後ろ向き・横断研究
- データソース: イスラエル国内3病院(2000-2022年)の電子カルテ(The Kinneret Israel Health Data Lake)
- 対象: 21歳以上で脂肪腫(ICD-9: 214)と診断された患者
- 解析人数: 7,868人(平均年齢53歳、女性53.6%)
- 比較対象: 2023年イスラエル全国健康調査(KAP調査)の年齢・性別別有病率
- 評価項目: 肥満、脂質異常症、高血圧、2型糖尿病
- 診断コード・薬剤・検査値を用いた感度重視の定義

結果

- 脂肪腫患者では、4つすべての代謝異常が一般人口より有意に高頻度
- 特に顕著だったのは:
 - ①脂質異常症(最も差が大きい)、②高血圧、③2型糖尿病、④肥満(差は小さいが一貫して高い)
- 若年層(35歳未満)でも差が顕著で、早期から代謝異常が集積
- 3項目以上が同時に存在する「メタボ様クラスター」は、35歳以降で急増、中年期～高齢期で特に顕著



脂肪腫は代謝症候群の構成要素の高頻度と関連している： 多施設横断研究

[Lipomas Are Associated with a Higher Prevalence of Metabolic Syndrome Components: A Multicenter Cross-Sectional Study.]

考察・解釈

- 脂肪腫は単なる局所的腫瘍ではなく、全身的な脂肪組織機能異常・慢性炎症・ミトコンドリア異常の可視的マーカーである可能性
- メタボリックシンドロームに伴う：脂肪細胞肥大、慢性炎症、インスリン抵抗性、ミトコンドリア機能障害が、脂肪腫形成を促進する生物学的妥当性が示唆される

臨床的意義

- 脂肪腫を見つけたら代謝スクリーニングを考慮すべき
 - ・ 血圧
 - ・ BMI / 腹囲
 - ・ 脂質血糖・HbA1c
 - ・ 特に、35歳以上、多発性脂肪腫、他の代謝異常を伴う場合には重要性が高い

限界

- 横断研究のため因果関係は不明
- 真の「脂肪腫なし対照群」が存在しない
- 病院ベースデータのため重症例に偏る可能性



OHDSI 論文の紹介(2)



メトホルミンとDPP4阻害薬の使用はがんリスクの有意な低下と関連する

PMID 41300987

> [Cancers \(Basel\)](#). 2025 Nov 10;17(22):3620. doi: 10.3390/cancers17223620.

Association Between the Use of DPP4 Inhibitors and Metformin and the Risk of Cancer in Patients with Type 2 Diabetes: A Multicenter Retrospective Cohort Study Using the OMOP CDM Database

Gyu Lee Kim^{1 2}, Yu Hyeon Yi^{1 2}, Jeong Gyu Lee^{1 2}, Young Jin Tak^{1 2}, Seung Hun Lee^{1 2}, Young Jin Ra^{1 2}, Byung Kwan Choi³, Sang Yeoup Lee^{2 4}, Young Hye Cho^{2 5}, Eun Ju Park^{2 5}, Youngin Lee^{2 5}, Jung In Choi^{2 5}, Sae Rom Lee^{2 5}, Ryuk Jun Kwon^{2 5}, Soo Min Son^{2 5}

- 1 Department of Family Medicine and Medical Research Institute, Pusan National University Hospital, Busan 49241, Republic of Korea.
- 2 Department of Family Medicine, School of Medicine, Pusan National University, Yangsan 50612, Republic of Korea.
- 3 Department of Neurosurgery, Pusan National University Hospital, Busan 49241, Republic of Korea.
- 4 Family Medicine Clinic, Obesity, Metabolism and Nutrition Center, Pusan National University Yangsan Hospital, Yangsan 50612, Republic of Korea.
- 5 Department of Family Medicine and Biomedical Research Institute, Pusan National University Yangsan Hospital, Yangsan 50612, Republic of Korea.

1. 研究の背景と目的

- 2型糖尿病 (T2DM) は、膵臓がん、肝臓がん、乳がん、大腸がんなど複数のがんリスク増加と関連している。
- メトホルミンにはがん予防効果が示唆されてきたが、先行研究の結果は必ずしも一貫していない。
- DPP4阻害薬 (DPP4is) のがんリスクへの影響については、増加を示唆する報告と関連なしとする報告があり、議論が続いている。
- 本研究の目的は、標準化されたデータ構造を用い、大規模な多施設共同研究によって DPP4阻害薬および/またはメトホルミンの使用とがん発生リスクの関連を評価することである。

2. 研究方法

- 韓国の11病院の電子健康記録 (EHR) を用いた多施設共同レトロスペクティブコホート研究である。
- すべてのデータベースは OMOP CDM に標準化され、OHDSI ネットワークのツールを用いて解析された。
- 19歳以上の2型糖尿病患者を対象とし 2群を比較した。
 - DPP4is/Met群: DPP4阻害薬またはメトホルミンのみを処方された群。
 - その他のGLD群: インスリン、スルホニル尿素薬 (SU) など、他種の血糖降下薬のみを処方された群。
- 傾向スコア (PS) を用いた1:1マッチングを行い、年齢、性別、併存疾患 (チャールソン指数) などの背景因子を調整した。

3. 主な結果

- DPP4is/Met群は、その他のGLD群と比較して、がん発生リスクが有意に低かった (ハザード比 [HR] 0.54; 95% 信頼区間 [CI] 0.41-0.69)。
- この保護的な関連はすべての病院において一貫しており、施設間の異質性は認められなかった。
- 薬剤を180日以上継続して使用した患者に限定した解析でも、がんリスクの有意な低下が認められた (HR 0.73; 95% CI 0.63-0.85)。
- がん種別の分布において、DPP4is/Met群では乳がんや前立腺がんの割合が低く、その他のGLD群では下部消化器がんの割合が高い傾向にあった。

4. 考察および結論

- 大規模多施設研究の結果、メトホルミンとDPP4阻害薬の使用はがんリスクの有意な低下と関連していることが示された。
- 本知見は、メトホルミンの潜在的ながん保護効果を支持し、DPP4阻害薬の中立ないし有益な安全性を裏付けるものである。
- ただし、DPP4is/Met群が糖尿病の初期段階で処方されやすい傾向にあることから、患者背景の良好さが結果に寄与した可能性 (健康な利用者効果) も考慮すべきである。
- 研究の限界として、喫煙や食事といった未測定の交絡因子の存在、管理用コードに基づく診断の誤分類、一部施設での追跡期間の短さが挙げられた。

■丁寧な考察がなされている

1. 薬剤による直接的ながん抑制メカニズム

- **メトホルミンの効果:** インスリン感受性の向上やAMPK-mTORシグナル経路の活性化を通じて、腫瘍の増殖を抑制する機序が先行研究と一致している。
- **DPP4阻害薬の安全性と可能性:** 過去に懸念された膵臓がん等のリスク増加は認められず、むしろ免疫調節作用などを通じた抗腫瘍効果の可能性が示唆されている。

2. 比較対象による影響(薬剤の相対的リスク)

- 比較対象となった「その他の血糖降下薬」群では、インスリンやスルホニル尿素(SU)薬が多く使用されていた。
- これらの薬剤は、先行研究においてがんリスクの増加との関連が指摘されているため、相対的にDPP4is/Met群のリスクが低く算出された可能性がある。

3. 「健康な利用者効果(Healthy User Effect)」の検討

- 考察のなかで最も重要な議論の一つが、このバイアスである。
- DPP4isやメトホルミンは糖尿病の比較的初期段階で処方されることが多い。
- 対照的に、インスリン等を使用する患者は、糖尿病が進行しており、併存疾患も多く、高齢である傾向がある。
- つまり、観察されたリスク低下は薬剤の薬理作用だけでなく、「もともと健康状態が良い患者がその薬を飲んでいた」という背景が影響している可能性がある。

4. 結論と今後の展望

- 本研究は、標準化された大規模データ(OMOP CDM)を用いることで、施設間の差を超えた一貫した傾向を示した点に大きな価値がある。
- 今後は、特定のがん種に絞った解析や、累積投与量・投与期間の影響をより詳細に検討する必要がある。



分権的な医療データをプライバシーを守りつつ安全に再利用する『連邦型データエコシステム』のイタリアでの構築戦略

> [Front Med \(Lausanne\)](#). 2025 Dec 8:12:1644719. doi: 10.3389/fmed.2025.1644719. eCollection 2025. PMID 41438153

The challenges of national health data ecosystems in feeding the European health data space: the Italian example

Ylenia Murgia¹, Roberta Gazzarata^{2 3 4}, Mario Ciampi^{4 5}, Mario Sicuranza^{4 5}, Franco Cirillo⁶, Christian Esposito⁶, Norbert Maggi^{1 2}, Gabriella Balestra^{7 8}, Lucia Sacchi^{8 9 10}, Mauro Giacomini^{1 8}

¹Department of Informatics, Bioengineering, Robotics and System Engineering (DIBRIS), University of Genoa, Genoa, Italy.

²Healthropy, Savona, Italy.

³HL7 Europe, Brussels, Belgium.

⁴HL7 Italy, Rome, Italy.

⁵Institute for High Performance Computing and Networking, National Research Council of Italy (ICAR-CNR), Naples, Italy.

⁶Department of Computer Science, University of Salerno, Fisciano, Italy.

⁷PoliToBIOMedLab, Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy.

⁸Italian Scientific Society of Biomedical Informatics (SIBIM), Pavia, Italy.

⁹Department of Industrial and Information Engineering, University of Pavia, Pavia, Italy.

¹⁰OHDSI Italian National Node, Italy.

欧州健康データスペース (EHDS) の概要

- 目的: EHDSは、個人の健康データの管理(一次利用)と、研究やイノベーションのためのデータ活用(二次利用)を安全かつ標準化された方法で促進する欧州連合(EU)の取り組みである。
- 実施状況: EHDS規則は2025年3月26日に発効し、2034年3月までに段階的に完全実施される予定である。
- 主要カテゴリ: 患者サマリー、電子処方箋、画像診断レポート、臨床検査結果、退院レポートなどが優先的なデータ共有対象となっている。

イタリアの保健医療システムの現状と課題

- 分権化された構造: イタリアの国立保健サービス(NHS)は高度に分権化されており、20の地域保健システム(RHS)がそれぞれ独自のガバナンスと管理戦略を持っている。
- 情報の断片化: 地域ごとに異なるインフラやポリシー運用がなされているため、全国的なデータの相互運用性や標準化が大きな障壁となっている。
- 電子健康記録(FSE): イタリアではFSE(Fascicolo Sanitario Elettronico)としてインフラ整備が進められており、現在は次世代版の「FSE 2.0」への移行と、データ中心のエコシステム(EDS)の構築が進行中である。

採用されている主な技術標準

イタリアの医療情報基盤では、以下の標準やコード体系が重要な役割を果たしている。

- HL7 (V2, V3, FHIR): データ交換プロトコルおよび構造化形式として採用。
- IHE XDS: 地域や施設を越えたドキュメントの登録と検索に使用。
- LOINC: 臨床検査結果や観察項目の意味的調和に使用。
- AIC-ATC: 医薬品の識別と分類に使用。
- OMOP CDM: 観察医療データの二次利用を可能にするための共通データモデル。

DHEAL-COM(ディアル・コム)プロジェクト

技術的アプローチ

- プロジェクトの狙い: 地域のデジタルヘルスを支援するためのナショナルハブを構築し、EDS(医療データエコシステム)と連携するモデルを目指している。

データの調和戦略

- ローカルノード: 各拠点のデータはHL7 FHIR形式に変換される。
- 中央リポジトリ: 二次利用を目的として、匿名化されたデータがOMOP CDMへと変換され、広範なデータエコシステムへの統合が図られる。
- フェデレーション(連邦型)アーキテクチャ: 識別可能な個人記録は各拠点のファイアウォール内に保持し、中央リポジトリはメタデータのみを管理して、生の個人データに触れずにクエリを実行できる仕組みを採用している。

セキュリティ、プライバシー、教育

- GDPRへの対応: 健康データはGDPR第9条(1)により原則処理が禁止されているが、医療目的や公共の利益などの例外(法的根拠)に基づいて処理される。
- イタリア当局(Garante)の姿勢: イタリアのデータ保護当局は保守的な立場をとっており、特にデータの二次利用や再識別リスク、欧州域外へのデータ移転に対して厳格な解釈を維持している。
- 専門スキルの習得: HL7 Italyなどは、FHIRやHL7の基礎コースを提供し、デジタル変革を支える専門人材の育成に注力している。

イタリアの分権型モデルにおける情報の断片化を克服するための、具体的な法的および技術的戦略は以下の通り。

1. 法的・規制的枠組みの強化

- FSE 2.0改革: 2023年と2024年の大臣令により、電子健康記録(FSE)の進化版である「FSE 2.0」が定義された。これは、単なる文書の共有から「データ中心」のエコシステムへの転換を法的・技術的に強制するものである。
- 医療データエコシステム(EDS)の導入: 全国規模で相互運用可能な環境を構築するため、保健データエコシステム(EDS)という概念が導入された。これにより、医療従事者、行政機関、および市民の間でシームレスにデータを共有できるガバナンス枠組みが確立された。
- 国家復興・回復計画(NRRP): パンデミックを契機に、医療のデジタル化に対して約156億ユーロという多額の予算が割り当てられ、デジタルインフラの完成と標準化の推進が法的に裏付けられた。

2. 技術的インフラの高度化

- INI(全国相互運用インフラ): 各地域が独自に管理するFSE ITシステムを接続するための、中央ハブとして機能する。INI自体は臨床データを保存せず、連邦型のクエリおよびルーティングサービスとして、全国どこからでも患者データにアクセスできる仕組みを提供している。
- Gatewayコンポーネント: FSE 2.0で導入された新しいアーキテクチャ要素であり、医療従事者や市民の間でデータをシームレスに共有するためのゲートウェイとして機能する。
- フェデレーション(連邦型)モデルの採用: データを中央に集約するのではなく、ローカルノード(各施設)にデータを残したまま、メタデータのみを中央で管理してクエリを実行する「分散型アーキテクチャ」が採用されている。

3. 意味的・構造的相互運用性の確保

- HL7 FHIRへの移行: 従来の文書ベースの形式(HL7 CDA)から、より粒度の細かいデータ交換が可能なHL7 FHIRモデルの採用が推奨されている。
- 標準コード体系の統合: 臨床観察項目にはLOINC、医薬品にはAIC-ATC、診断にはICDといった国際標準や国内標準コードの採用が進められ、地域間のセマンティックな差異を埋める努力がなされている。
- OMOP CDMIによる二次利用: DHEAL-COMなどのプロジェクトでは、研究や分析(二次利用)のためにデータをOMOP Common Data Modelに変換する戦略がとられている。これにより、異なる病院のデータであっても共通の言語で分析が可能になる。

4. プライバシー保護と研究の両立

- 匿名化と合成データの活用: イタリアのデータ保護当局(Garante)の厳格な規制に対応するため、再識別リスクを最小限に抑える「合成データ(Synthetic Data)」の生成や、高度な匿名化技術の導入が検討されている。
- バイオエシックス専門家との連携: テリトリアルなプロジェクトにおいて、臨床的な有用性を維持しつつGDPRに準拠するための、7段階のフェデレーションクエリ・ワークフローなどの安全策が講じられている。
- イタリアの事例は、強い地域自治を持つ国が、どのようにして欧州全体の健康データスペース(EHDS)という超国家的なシステムに統合しようとしているかの先駆的なモデルとなっている。



ディアル・コム

総予算: 3,000万ユーロ(約55億円)

「1,500万ユーロ + 1,500万ユーロ」の2段階(4年 + 4年の計8年間)

1. プロジェクトの主な目的

DHEAL-COMは、イタリアの補完的国家計画(CNP)から資金提供を受けており、以下の達成を目指している。

- **コミュニティ医療の強化:** テレメディシンやデジタルツールを活用し、地域に基づいた新しい医療サービスモデルを構築する。
- **データ駆動型イノベーション:** 収集された健康データを研究やリスク分析(AIによる層別化など)に再利用できる環境を整える。
- **欧州標準への準拠:** 欧州健康データスペース(EHDS)の要件に合わせ、国境を越えたデータ共有が可能なインフラを整備する。

2. 技術的な「二段構え」の戦略

データが施設ごとに異なる形式で保存されているという「断片化」を解決するため、DHEAL-COMは高度な調和戦略を採用している。

1. **ローカル調和(FHIR):** 各拠点のデータは、まずデータ交換のための国際標準規格である **HL7 FHIR** 形式に変換される。
2. **中央ハブでの統合(OMOP CDM):** 研究や分析での再利用を容易にするため、匿名化されたFHIRデータはさらに **OMOP CDM** へと変換される。

3. プライバシーを守る「分散型」アーキテクチャ

イタリアの厳しいデータ保護規制(GDPRおよび国内当局の解釈)に対応するため、データは「中央集権」ではなく「分散」して管理される。

- **ローカルノード:** 患者の個人を特定できる生のデータは、各医療機関のファイアウォール内に残される。
- **中央リポジトリ:** プロジェクト資料や匿名化された集計データ、そして各ノードがどのようなデータを持っているかを示す「メタデータ」のみを管理する。
- **フェデレーション・クエリ:** 中央からクエリ(検索命令)を送ると、各ノードが自施設内で計算を行い、**統計的な結果(匿名化された出力)のみ**を中央に返す仕組みである。

4. 期待される成果

- **市民・専門家向けのプラットフォーム:** 開発されたデジタルツールやデータにアクセスできるオープンなウェブプラットフォーム。
- **オープンラボ:** 革新的なヘルスケアソリューションをテストするための高度なハードウェア・ソフトウェアを備えた施設。
- **スキルの向上:** 医療従事者や技術者が標準規格を使いこなせるようにするための教育・トレーニングの促進。

DHEAL-COM が採用している、個人情報保護しながら分散されたデータを横断的に検索する「7段階のクエリ実行ワークフロー」について解説する。このワークフローは、中央ハブが各施設の生のデータに直接アクセスすることなく、統計的な集計結果のみを安全に回収するための厳格な手順である。

フェデレーション・クエリの7段階ワークフロー

1.クエリの提出と検証

- ユーザーが検索条件(例:「メトホルミンを服用している2型糖尿病の成人患者数」など)を提出する。
- 中央ノードは、ユーザーの身元、権限、および要求された出力形式(集計データのみか等)が適切であることを検証する。

2.実行計画と施設能力の評価

- 中央ノードは、各施設が持つメタデータ(対応しているFHIRリソース、用語体系、データの鮮度など)を照会し、どの施設にクエリを送信すべきか計画を立てる。

3.クエリパッケージの安全な配布

- 認証されたセキュアな通信路(VPN等)を通じて、クエリを各施設へ送信する。
- この際、施設ごとのプライバシーポリシー(例:最小セルサイズを10以上とする等)や期限付きの署名トークンが同封される。個人識別子は一切含まれない。

4.ローカルでの実行

- 各施設(ローカルノード)のファイアウォールの内側で、標準化されたデータ(FHIRまたはOMOP形式)に対してクエリが実行される。

5.プライバシー強化と集計出力の生成

- 出力前に、各施設でプライバシールールを適用する。
- 例えば、該当者数が極端に少ない場合に値を秘匿(サブレッション)したり、統計的なノイズを加えたりして、個人の特性を不可能にした「集計結果」のみを生成する。

6.横断的集計と結果の編纂

- 中央ノードが各施設から送られてきた非識別的な集計結果を収集する。
- データの整合性をチェックした上で、全体の統計値としてまとめ、最終的なレポートを作成する。

7.包括的なログ記録と監査

- 誰が、いつ、どの施設に対して、どのようなポリシーで検索を行ったかをすべて記録する。
- このログは改ざん不能な形で保存されるが、ログ自体にも生の個人情報は一切含まれない。



RIMR AMDコンソーシアム報告第1報

› Graefes Arch Clin Exp Ophthalmol. 2025 Dec 29. doi: 10.1007/s00417-025-07081-4.

PMID 41460325

Online ahead of print.

Initiation of a global consortium to study the progression of age-related macular degeneration: RIMR AMD consortium report # 1

Srinivas R Sadda^{1 2 3}, Charles C Wyckoff⁴, Itay Chowers⁵, Jean-Francois Korobelnik^{6 7}, Raja Narayana⁸, Rajeev R Pappuru⁸, Frank G Holz⁹, Robyn Guymmer¹⁰, Chui Ming Gemmy Cheung¹¹, David S Boyer¹², Michael Ip^{13 14 15}, Heiko G Niessen¹⁶, Nancy Holenkamp¹⁷, Mary Durbin¹⁸, Stephanie Magazzeni¹⁹, Anne-Marie Cairns²⁰, Carlos Ciller²¹, Natasa Jovic²¹, Joseph Blair²¹, Sandro De Zanet²¹, Aaron Y Lee^{13 22}

¹RIMR AMD Consortium, Pasadena, CA, USA. ssadda@doheny.org.

²Doheny Eye Institute, 150 North Orange Grove Blvd, Pasadena, CA, 91103, USA. ssadda@doheny.org.

³Department of Ophthalmology, University of California - Los Angeles, Los Angeles, CA, USA. ssadda@doheny.org.

⁴Retina Consultants of Texas, Houston, TX, USA.

⁵Hadassah Medical Center, Faculty of Medicine, the Hebrew University of Jerusalem, Jerusalem, Israel.

⁶CHU Bordeaux, Service d'ophtalmologie, France.

⁷University of Bordeaux, INSERM, BPH, UMR1219, Bordeaux, F-33000, France.

⁸Anant Bajaj Retina Institute, L.V. Prasad Eye Institute, Hyderabad, India.

⁹Department of Ophthalmology, University of Bonn, Bonn, Germany.

¹⁰Center for Eye Research Australia, Royal Victorian Eye and Ear Hospital, The University of Melbourne, Melbourne, Australia.

¹¹Singapore National Eye Center, Singapore, Singapore.

¹²Retina Vitreous Associates, Los Angeles, CA, USA.

¹³RIMR AMD Consortium, Pasadena, CA, USA.

¹⁴Doheny Eye Institute, 150 North Orange Grove Blvd, Pasadena, CA, 91103, USA.

¹⁵Department of Ophthalmology, University of California - Los Angeles, Los Angeles, CA, USA.

¹⁶Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany.

研究の概要と目的

- 加齢黄斑変性 (AMD) の進行をより深く理解し、治療薬開発を促進するため、世界規模の共同コンソーシアム「RIMR AMD Consortium」が設立された。
- 主な目的は、多施設から縦断的なマルチモーダル画像データを集約・解析し、新しいバイオマーカーの定義や、より精細な疾患重症度スケールの構築を行うことである。

組織構成と運営

- 2023年にカリフォルニア州で設立された非営利団体であり、アカデミア (画像・臨床データの提供)、バイオ医薬品企業 (資金提供)、画像技術企業 (ソフトウェア開発支援) のパートナーシップによって構成されている。
- データ提供者が所有権を保持したまま、クラウド上で解析を行うモデルを採用しており、プライバシー保護や規制 (GDPR等)、データの独占といった課題に対処している。

収集データの現状

- 現時点で、世界4大陸の7つのコホート (米国、ドイツ、フランス、イスラエル、インド、オーストラリア) から、5,000人以上の被験者、10万件を超える光干渉断層計 (OCT) ボリュームデータが収集されている。
- データには正常な加齢から末期AMDまでの幅広いステージが含まれており、一部の被験者では15年以上の長期追跡データも存在する。

技術的基盤と解析手法

- 収集された画像データはDICOM形式に変換・統一され、関連する臨床メタデータは**OMOP共通データモデル (CDM)** にマッピングされている。
- 解析はセキュアなクラウド環境 (AWS) 内で行われ、人工知能 (AI) ツールを用いた自動解析と、専門家による検証 (データの約1%) を組み合わせた体制が整えられている。

今後の展望

- 地域や民族によって異なるAMDの臨床的特徴 (ドルーゼンの形態や萎縮の進行速度など) の解明が期待されている。
- 今後は、現在データの少ない東アジア、アフリカ、中南米などの地域からの参加を募り、データの多様性をさらに高める計画である。



欧州における精密がん医療（PCM）のデータ統合プロジェクト 「PRIME-ROSE」の目的、手法、および現在の進捗状況について

› [Acta Oncol.](#) 2026 Jan 6;65:1-8. doi: 10.2340/1651-226X.2026.44889.

PMID 41496458

Procedures of data merging in precision cancer medicine: the PRIME-ROSE project

Henk Van der Pol ¹, Tina Kringelbach ², Maria Martin Agudo ³, Gabriel Bratseth Stav ³,
Gro Live Fagereng ³, Marta Fiocco ⁴, Ragnhild Sørum Falk ⁵, Victoria Homer ⁶,
Soemeya Haj Mohammad ⁷, Hans Timmer ⁷, Loic Verlingue ⁸, Åslaug Helland ⁹,
Kristoffer Rohrberg ¹⁰, Ulrik Lassen ¹⁰, Sarah Halford ¹¹, Katriina Jalkanen ¹², Tanja Juslin ¹²,
Matthew G Krebs ¹³, Julio Oliveira ¹⁴, Edita Baltruskeviciene ¹⁵, Kristiina Ojamaa ¹⁶, Kjetil Taskén ⁹,
Hans Gelderblom ¹⁷

¹Department of Medical Oncology, Leiden University Medical Centre, Leiden, The Netherlands; Mathematical Institute, Leiden University, Leiden, The Netherlands.

²Department of Oncology, Copenhagen University Hospital - Rigshospitalet, Copenhagen, Denmark.

³Institute for Cancer Research and Division of Cancer Medicine, Oslo University Hospital, Oslo, Norway.

⁴Mathematical Institute, Leiden University, Leiden, The Netherlands; Princess Maxima Center, Utrecht, The Netherlands; Department of Biomedical Data Science, Leiden University Medical Center, Leiden, The Netherlands.

⁵Oslo Centre for Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway.

⁶Cancer Research (UK) Clinical Trials Unit, University of Birmingham, Birmingham, United Kingdom.

⁷Department of Medical Oncology, Leiden University Medical Centre, Leiden, The Netherlands.

⁸Centre Léon Bérard, Lyon, France.

⁹Institute for Cancer Research and Division of Cancer Medicine, Oslo University Hospital, Oslo, Norway; Institute of Clinical Medicine, University of Oslo, Oslo, Norway.

¹⁰Department of Oncology, Copenhagen University Hospital - Rigshospitalet, Copenhagen, Denmark; Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark.

¹¹Cancer Research UK, London, United Kingdom.

¹²Helsinki University Hospital Comprehensive Cancer Center, Helsinki, Finland.

研究の背景と目的

- 精密がん医療(PCM)の臨床試験では、腫瘍の分子学的特性によってサブタイプが細分化されるため、各試験における対象患者数が少なくなることが大きな課題である。
- PRIME-ROSEプロジェクトは、欧州全域の「DRUP型臨床試験(DLCT)」からデータを集約することで、症例蓄積の遅いコホートを加速させることを目的としている。

組織構成とデータ共有体制

- 現在、欧州全域で11のDLCT(臨床試験)が進行中または開始予定であり、これらがネットワークを形成している。
- データ共有契約(DSA)に基づき、中央サーバーでデータを共有・管理する体制を構築している。
- 各試験の代表者が毎月会議を行い、症例登録の進捗確認やコホートの継続・閉鎖について意思決定を行っている。

データ統合の標準化手法

- 異なる試験間のデータを統合するため、OMOP CDMを採用し、標準化を図っている。
- エンドポイントの整合化: 多くの試験で、無増悪生存期間(PFS)、全生存期間(OS)、16週時点での病勢コントロール率(DCR)などが共通の評価項目として設定されている。
- 用語の調和(ハーモナイゼーション): 腫瘍タイプやバイオマーカーの定義を統一されたリストに基づいて整理し、臨床アウトカムデータの形式を標準化している。

現在の進捗と成果

- 2024年第2四半期からデータの共有・統合が開始され、2025年夏時点で20種類以上の治療法、300以上のコホート、1,000人以上の患者データをモニタリングしている。
- 中間解析の結果、少なくとも20のコホートが次の段階へ進展しており、治療効果の欠如により閉鎖された統合コホートはまだ存在しない。
- 現在、4つのコホートが症例登録を完了し、最終的な解析段階にある。

結論と今後の展望

- 欧州諸国間でのデータ共有は実現可能であり、希少な腫瘍とバイオマーカーの組み合わせに対する薬剤の有効性を迅速に評価する上で極めて有効である。
- プロトコルの差異(ECOGパフォーマンスステータスの基準など)によるデータの異質性には注意が必要だが、OMOP CDMの活用により将来的なデータ統合はさらに効率化される見込みである。



リアルタイムEHRデータの臨床研究における完全性評価

➤ [PLoS One](#). 2026 Jan 9;21(1):e0340287. doi: 10.1371/journal.pone.0340287. eCollection 2026.

PMID 41511976

Assessment of the integrity of real-time electronic health record data used in clinical research

Jessica Liu ¹, Sameer Pandya ², Andreas Coppi ^{3 4}, H Patrick Young ², Harlan M Krumholz ^{3 4 5},
Wade L Schulz ², Guannan Gong ¹

- ¹ Yale Cancer Center, Yale School of Medicine, New Haven, Connecticut, United States of America.
- ² Department of Laboratory Medicine, Yale University School of Medicine, New Haven, Connecticut, United States of America.
- ³ Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, Connecticut, United States of America.
- ⁴ Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, Connecticut, United States of America.
- ⁵ Department of Health Policy and Management, Yale School of Public Health, New Haven, Connecticut, United States of America.

研究の目的と背景

- リアルタイムのEHRデータは二次利用(研究、運営、臨床ケア)において大きな可能性を持つが、データの品質と安定性の確保が課題となっている。
- 従来のデータ品質評価は遡及的なデータセットを対象としたものが多く、リアルタイムデータの「研究への即応性(research readiness)」を体系的に調査した研究は少ない。
- 本研究は、リアルタイムEHRデータの安定性と完全性を評価し、データがいつ「分析可能」な状態になるかを特定することを目的としている。

研究手法

- 分析プラットフォーム: イェール・ニューヘブーン・ヘルス(YNHH)の臨床データウェアハウスのデータを、OMOP CDMに変換して分析を行った。
- データセット: 2024年7月~11月の週単位スナップショット(29回分)と、2025年4月~5月の日単位スナップショット(22回分)を使用した。
- 検証: 分析コードの妥当性を確認するため、MIMIC-IIIから派生させた合成データセットを用いてベンチマークパイプラインの検証を行った。
- 評価指標(ベンチマーク):
 - BM-1(臨床アクション): 患者の追加、削除、統合、ID変更などのデータベースレベルの変化
 - BM-2(人口統計学的変数): 生年月日、性別、人種、民族の更新
 - BM-3(退院情報): 退院時間および退院ステータスの安定性

主な結果

- 臨床アクションとIDの更新: 連続するスナップショット間で、新規患者の追加やIDの変更、データの統合といった臨床アクションによる更新が頻繁に発生していた。
- 人口統計データの不安定性: 人口統計情報の変更は、特に人種(race)と民族(ethnicity)のフィールドで最も頻繁に発生した。これは、緊急受診時の不完全なデータ入力や、後の受診時の修正、または管理者による標準化作業を反映している可能性がある。
- 退院情報の安定化時間: 退院時間と退院ステータスの値は、イベント発生直後は不安定であり、安定した状態に達するまでに通常4~7日を要することが明らかになった。
- バッチ処理の観察: 外来患者のデータにおいて、特定の日に退院ステータスの更新が急増する現象が観察され、月末の調整ワークフローや一括更新の存在が示唆された。

結論と提言

- 本研究は、リアルタイムEHRデータの完全性を評価し、分析可能なタイミングを特定するための自動ベンチマークパイプラインの実現可能性を示した。
- リアルタイムデータは静的なデータ品質ツールでは捉えられない動的な変化を伴うため、二次利用にあたってはデータの安定化期間(4~7日など)を考慮することが極めて重要である。
- この自動化フレームワークは、他の医療システムにも適用可能であり、高品質なリアルタイム研究や疾患監視、臨床試験の準備を支援する根拠に基づいた指針を提供する。



Correspondence: MI-CDMに関するレターと回答

論文21
論文22

> [J Am Med Inform Assoc. 2025 Dec 5;ocaf215. doi: 10.1093/jamia/ocaf215. Online ahead of print.](#)

Toward semantic interoperability of imaging and clinical data: reflections on the DICOM-OMOP integration framework

PMID 41453141

[Weihao Cheng](#)^{1, 2}, [Zekai Yu](#)¹

¹School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou, Zhejiang Province, 310018, PR China.

²School of Communication Engineering, Hangzhou Dianzi University, Hangzhou, Zhejiang Province, 310018, PR China.

> [J Am Med Inform Assoc. 2025 Dec 24;ocaf216. doi: 10.1093/jamia/ocaf216. Online ahead of print.](#)

Response to "toward semantic interoperability of imaging and clinical data: reflections on the DICOM-OMOP integration framework"

PMID 41453072

[Woo Yeon Park](#)¹, [Teri Sippel Schmidt](#)¹, [Gabriel Salvador](#)¹, [Kevin O'Donnell](#)², [Brad Genereaux](#)^{1, 3}, [Kyulee Jeon](#)^{4, 5}, [Seng Chan You](#)^{4, 5}, [Blake E Dewey](#)^{1, 6}, [Paul Nagy](#)¹

¹Biomedical Informatics and Data Science, Johns Hopkins University, Baltimore, MD, 21205, United States.

²Canon Medical Research, Vernon Hills, IL, 60061, United States.

³NVIDIA Corporation, Santa Clara, CA, 95051, United States.

⁴Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, 03722, Republic of Korea.

⁵Institute for Innovation in Digital Healthcare, Yonsei University Health System, Seoul, 03722, Republic of Korea.

⁶Department of Neurology, School of Medicine, Johns Hopkins University, Baltimore, MD, 21205, United States.

■ MI-CDMについては下記参照

MI-Common Data Model: Extending Observational Medical Outcomes Partnership-Common Data Model (OMOP-CDM) for Registering Medical Imaging Metadata and Subsequent Curation Processes

PMID: 38061012



がん治験スクリーニングシステム CPTM

> JCO Clin Cancer Inform. 2026 Jan;10:e2500262. doi: 10.1200/CCI-25-00262. Epub 2026 Jan 9.

Clinical Trial Patient Matching: A Real-Time, Common Data Model and Artificial Intelligence-Driven System for Semiautomated Patient Prescreening in Cancer Clinical Trials

PMID 41512229

Guannan Gong^{1 2}, Jessica Liu¹, Sameer Pandya³, Cristian Taborda^{1 2}, Nathalie Wiesendanger¹, Nate Price⁴, Will Byron⁴, Andreas Coppi^{5 6}, Patrick Young³, Christina Wiess^{1 7}, Haley Dunning^{1 7}, Courtney Barganier^{1 7}, Rachel Brodeur^{1 7}, Neal Fischbach^{1 2 7}, Patricia LoRusso^{1 2}, Lajos Pusztai^{1 2}, So Yeon Kim^{1 2}, Mariya Rozenblit^{1 2}, Michael Cecchini^{1 2}, Anne Mongiu^{1 2}, Lourdes Mendez^{1 2}, Edward Kaftan¹, Charles Torre Jr⁴, Harlan Krumholz^{6 8}, Ian Krop^{1 2 7}, Wade Schulz³, Maryam Lustberg^{1 2}, Pamela L Kunz^{1 2}

¹Yale Cancer Center, Yale School of Medicine, New Haven, CT.

²Department of Medicine, Division of Medical Oncology, Yale School of Medicine, New Haven, CT.

³Department of Laboratory Medicine, Yale School of Medicine, New Haven, CT.

⁴Digital and Technology Solutions, Yale New Haven Hospital, New Haven, CT.

⁵Department of Medicine, Division of Cardiovascular Medicine, Yale School of Medicine, New Haven, CT.

⁶Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, CT.

⁷Clinical Trials Office, Yale Cancer Center, Yale University, New Haven, CT.

⁸Department of Health Policy and Management, Yale School of Public Health, New Haven, CT.

Abstractのさらに要約

がん治験の患者登録率は5~7%と極めて低く、手作業によるマッチングの停滞が大きな課題である。既存のAIシステムもデータ標準化の不備により汎用性に欠けていた。本研究では、電子カルテの構造化・非構造化データをOMOP共通データモデルで標準化し、ルールベースと自然言語処理(NLP)を統合した半自動マッチングツール「CTPM」を開発・検証した。

大腸がん治験での検証では、感度100%、精度88~94%という高い性能を示した。これにより、チャートレビューの負荷は10分の1に激減し、1件あたりのスクリーニング時間も41%短縮された。2022年9月からの運用では、29の治験にわたり約9.8万人を精査し、実際に117名の登録を実現している。

本ツールは、研究チームを膨大なカルテ確認作業から解放し、真に適格な候補者への対応に注力させることを可能にする。OMOP基盤の採用により施設間での拡張性も高く、治験アクセスの向上に寄与する有力な解決策である。



Global/APACの動き



12月の OHDSI Global/APAC

●Global Community Call テーマ

Jan 13 Where Can We Go Together in 2026

Jan. 20 Building Connections for Future Collaborations

●APAC Call テーマ <https://ohdsi.org/apac/>

Jan 15 Agenda:

- Selected Studies from OHDSI Taiwan
- 2026 APAC Study Proposal

●Europe Community Call テーマ

Oct. 9 Meet the New National Nodes

Nov. 13 Patient-Reported Outcome Measures (PROMs)

Dec. 11 Vocabularies in Europe

Jan. 8 Update on European Projects Using OMOP



Our Journey

*Where The OHDSI Community Has Been
And Where We Are Going*
2025 edition

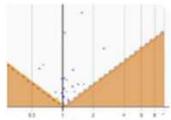
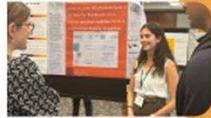


OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

<https://www.ohdsi.org/wp-content/uploads/2025/12/OurJourney-2025.pdf>



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



#JoinTheJourney

TABLE OF CONTENTS

I. Welcome To The Community.....	2
II. Mission, Values, And What We Do.....	5
How OHDSI Works.....	7
Columbia University as Coordinating Center.....	8
III. Collaborators.....	9
Map of Collaborators.....	10
Workgroups.....	12
Regional Chapters.....	14
OHDSI Around The World.....	15
European National Nodes.....	16
DARWIN EU®.....	18
EHDEN Foundation.....	20
Sponsorship Opportunities.....	21
Organizations Involved with OHDSI.....	22
Community Testimonials.....	24
Titan Awards.....	26
IV. Events & Activities.....	29
Symposia Around The World.....	30
2025 Global Symposium.....	34
Community Calls.....	40
Studyathons and Other Events.....	42
Guideline-Driven Evidence.....	44
Phenotype Phebruary/DevCon.....	46
V. Educational Resources.....	47
Tutorials.....	48
EHDEN Academy.....	49
The Book of OHDSI.....	50
OHDSI Summer School at Columbia University.....	52
VI. Data Standards.....	53
OMOP Common Data Model.....	54
OMOP CDM Data Sources.....	56
OHDSI Evidence Network.....	58
OHDSI Standardized Vocabularies.....	60
Themis Conventions.....	64
OMOP and FHIR.....	65
VII. Open-Source Software.....	66
HADES Packages.....	67
Package Statuses/Maintainers.....	71
ATLAS.....	72
Strategus.....	73
Open-Source Impact on Evidence.....	74
VIII. Methods Research.....	75
The LEGEND Project.....	76
Causal Effect Estimation.....	78
Patient-Level Prediction.....	80
Generative AI and Foundational Models.....	82
IX. Publications.....	83
Collaborations Graph.....	84
OHDSI Publications (2010 - September 2025).....	86
X. Join The Journey.....	129
OHDSI in a Word.....	130
Closing Letter.....	132
How Can You Join The Journey?.....	Inside Back Cover



Package Name	Version	Author	License
HADES	1.0.0	OHDSI	MIT
ATLAS	1.0.0	OHDSI	MIT
Strategus	1.0.0	OHDSI	MIT

Publication Title	Year	Journal
OHDSI Publications (2010 - September 2025)	2010-2025	Various

Europe National Nodes

Over the last two years, OHDSI Europe developed National Nodes, which are collections of research institutions within a member country. The Nodes include a broad range of member groups, ranging from research institutes, pharmaceutical and IT companies, and SMEs. They are generally led by academic institutions and are inclusive and open to any organization that wants to share experiences with OHDSI-related work, such as mapping data to the OMOP Common Data Model or performing network studies.

As of September 2025, there are 17 National Nodes, and two more nations have shared their intent to form their own node.



Europe National Nodes include:

- 17 nations
- 145 data holders
- more than 800 members
- ... and counting

Node..... Lead(s)

Belgium	Liesbet Peeters, Annelies Verbiest, Ilse Vermeulen
Denmark	Ismail Gögenur, Martin Hoyer Rose, Andreas Weinberger Rosen
Estonia.....	Raivo Kolde, Sulev Reisberg
Finland.....	Eric Fey, Gustav Klingstedt
Germany.....	Ines Reinecke, Michele Zoch
Greece.....	Anastasia Farmaki, Pantelis Natsiavas, Grigoris Papapostolou
Hungary.....	Zsolt Bagyura, Ágota Mészáros
Ireland.....	Aedin Culhane, Mark Lawler, Catherine Mahoney
Israel.....	Chen Yanover
Italy.....	Lucia Sacchi, Matteo Gabetta
Luxembourg.....	Claudine Backes, Andreas Kremer, Maria Quaranta
Netherlands.....	Renske Los, Aniek Markus
Norway.....	Espen Enerly, Siri Larønningen
Portugal.....	Patricia Couceiro, Carmen Nogueira
Spain.....	Miguel Angel Mayer, Talita Duarte Salles
Switzerland.....	Olga Endrich, Karen Triep
United Kingdom.....	Dani Prieto-Alhambra

coming soon Austria, Sweden

If you are working with OMOP data in a country that does not yet have a national node, get your OHDSI friends together, find your common ground, define objectives and form a node! More information can be found here:

www.ohdsi-europe.org/index.php/national-nodes

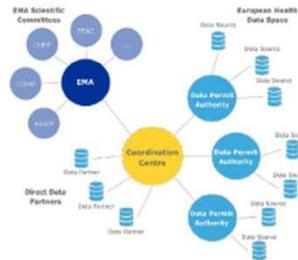
DARWIN EU®

The European Medicines Agency (EMA) and the European Medicines Regulatory Network established a coordination centre at Erasmus University Medical Center Rotterdam to provide timely and reliable evidence on the use, safety and effectiveness of medicines for human use, including vaccines, from real world healthcare databases across the European Union (EU). This capability is called the Data Analysis and Real World Interrogation Network (DARWIN EU®).

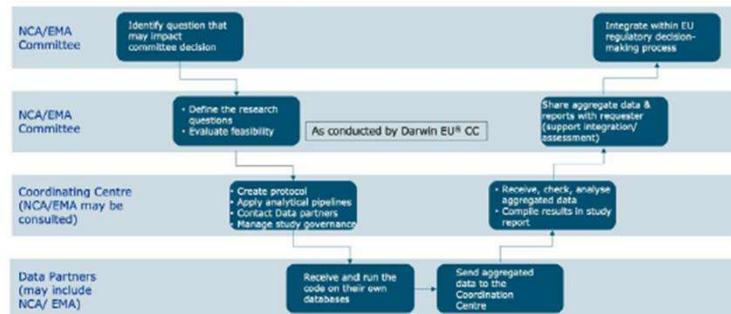
The DARWIN EU® data network currently includes 30 data partners across 16 nations — comprising more than 181 million patients — that have been mapped to the OMOP common data model. Multiple OHDSI Titan Award winners can be found leading the DARWIN EU® coordination centre, and many OHDSI tools, methods and practices are present in completed and ongoing studies.

DARWIN EU Objectives

- 1) To establish and maintain a continually enlarging network of accessible observational data sources
- 2) To execute all steps of high quality non-interventional studies with the network
- 3) To make the study results available to the EU Regulatory network to support decision-making



What is the DARWIN EU® process for conducting studies?



Types of Studies

Category	Description
Off-The-Shelf Studies	These are mainly characterisation questions that can be executed with a generic protocol. This includes studies on disease epidemiology, for example the estimation of the prevalence or incidence of health outcomes in defined time periods and population groups, or drug utilization studies at the population or patient level.
Complex Studies	These are studies requiring development or customisation of specific study designs, protocols, analytics and phenotypes. This includes studies on the safety and effectiveness of medicines and vaccines.
Routine Repeated Analyses	Routine analyses based on Off-The-Shelf or Complex Studies (see above), which are repeated with a pre-specified regularity (e.g. yearly)
Very Complex Studies	Studies which cannot rely only on electronic health care databases, or which require complex and/or novel methodological work

Number of studies

	PHASE I Establishment – 1st year	PHASE II Establishment – 2nd year	PHASE III Operation – 1st year	Operation 2nd year	Operation 3rd year
Phases	Phase I	Phase II	Phase III	Option 1	Option 2
Routine Repeated analysis	-	-	4	35-50	35-50
Off the shelf studies	3	14	23	35-50	35-50
Complex Studies	1	4	10	15-20	15-20
Very Complex Studies	0	0	0	1	1



darwin-eu.org

DARWIN EU® Leadership



Peter Rijnbeek
Erasmus MC
Executive Director, Technology Pillar Lead



Dani Prieto-Alhambra
Erasmus MC, Oxford University
Deputy Director, Development Pillar Lead



Katia Verhamme
Erasmus MC
Deputy Director, Study Operations Pillar Lead



Carlos Diaz
Synapse Research Management Partners
Management Pillar Lead



Maxim Moinat
Erasmus MC
Network Operations Pillar Lead

EHDEN Foundation

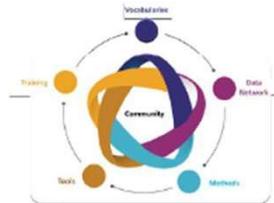
EHDEN — the European Health Data & Evidence Network — was launched in 2019 to address the challenges of generating insights and evidence from real-world clinical data at scale throughout Europe. Building on OHDSI tools and practices, EHDEN developed a public-private consortium consisting of 25 organizations that successfully implemented its mission: *to provide a new paradigm for the discovery and analysis of health data in Europe, by building a large scale, sustainable federated network of data sources standardised to the OMOP common data model.*

EHDEN leaders recognized the need to continue this work, so they created the EHDEN Foundation, which set a mission to operationalize a new paradigm for the discovery and analysis of health data, building on a large-scale federated network of data sources standardized to the OMOP CDM.

EHDEN Foundation Vision: The EHDEN

Foundation aspires to be the trusted key actor in Europe to facilitate and accelerate the generation of high-quality real-world evidence to improve healthcare of patients.

To implement its vision, the EHDEN Foundation will maintain the EHDEN Catalogue and EHDEN Academy, create research opportunities for multiple stakeholders, execute federated studies and support methodological and technical developments to further expand and improve the Data Partner Network in strong collaboration with the OHDSI community.



Learn More: www.ehden.eu

Key Research Activities for the EHDEN Foundation

Advisory	Studies	Research Programmes	Training
Data Landscaping	Disease Epidemiology	Neuroscience	Study-a-thon
Feasibility	Drug Utilisation	Oncology	Online Training
Study Design	Drug Safety/Effectiveness	Other	On-site Training

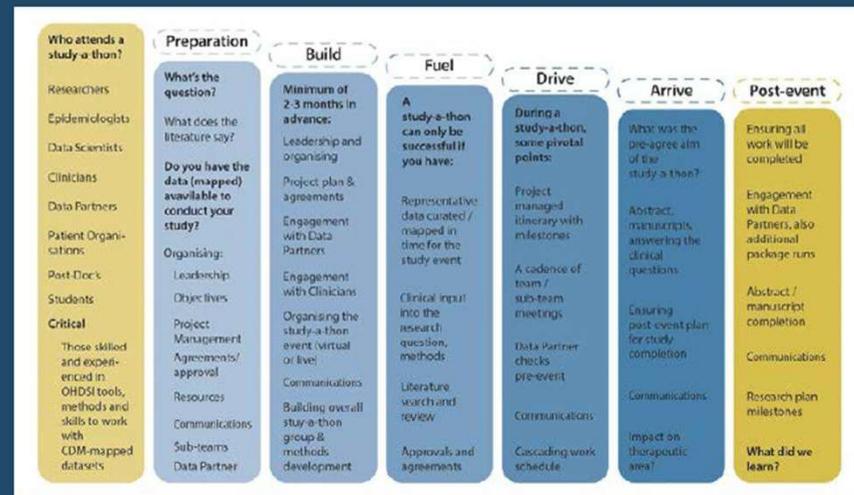
OHDSI Studyathons & Other Events

How does OHDSI go about *empowering a community to collaboratively generate the evidence that promotes better health decisions and better care?*

We do it by innovating on what it means to do collaborative research.

The premise of the studyathon is simple: bring together a diverse group of researchers aligned on a common question and focus together on collaboratively designing research protocols, executing analyses across databases, and interpreting results over an intense but fun-filled few days.

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



Outline flow of a study-a-thon. Graphic was shared in "Evaluating a novel approach to stimulate open science collaborations: a case series of "study-a-thon" events within the OHDSI and European IMI communities" • Jamia Open, Volume 5, Issue 4, December 2022, ooac100, <https://doi.org/10.1093/jamiaopen/ooac100>.

<https://academic.oup.com/jamiaopen/article/5/4/ooac100/6832031>

EDUCATIONAL RESOURCES

Join Us at the 2026 OHDSI Summer School at Columbia University

The Columbia OHDSI Summer School (June 22-26, 2026) offers health professionals, researchers, and industry practitioners an immersive, hands-on introduction to working with real-world health data and generating real-world evidence (RWE). Participants will learn how to transform electronic health records and claims data into the OMOP Common Data Model to support collaborative, distributed research.

Program Highlights

- Explore three analytic use cases:
 - *Clinical characterization* – describing disease natural history and treatment patterns
 - *Population-level estimation* – assessing drug safety and comparative effectiveness
 - *Patient-level prediction* – applying machine learning for early detection and precision medicine
- Work through the full RWE study lifecycle: study design, use of OHDSI open-source tools (ATLAS, HADES), and execution across real-world datasets
- Blend of foundational lectures, interactive exercises, and faculty-led group work
- Dedicated time to develop your own study ideas with mentoring and feedback

Audience

Ideal for clinicians, data scientists, statisticians, epidemiologists, informaticians, health policy researchers, and professionals from academia, healthcare, industry, and government. No programming experience required.

Testimonials

"The OHDSI Summer School exceeded all expectations. The course offered a deep, hands-on dive into real-world data methodology, led by world-class instructors and supported by an incredibly open, collaborative community. Working hands-on with a team to explore a real research question brought the full OHDSI workflow to life. I highly recommend this course to anyone working with health data, especially clinical scientists and data scientists eager to strengthen their skills in transparent, high-quality RWD research."

- David Bard, Professor of Pediatrics and Director, Biomedical & Behavioral Methodology Core, University of Oklahoma

"I particularly appreciated the OHDSI system in conducting the observational study to generate real-world evidence, as well as the methodology that handled hundreds of covariates, and the application of a negative control outcome to validate the research results. Through this summer school course, I gained a deeper understanding of OHDSI and its potential to generate real-world evidence."

- Long-Sheng Chen, Professor, National Taipei University of Technology, Taiwan

Meet Our Faculty



George Hriposak



Patriok Ryan



Anna Ostropolets



Karthik Natarajan

OHDSI Evidence Network

In this section, you learned about the OMOP Common Data Model, our foundational tool which standardizes patient-level data. You saw the reach of OMOP, which connects nearly one billion patients across six different continents in a way no other community can match.

However, if we aren't using this data to make an impact, we are wasting an incredible resource. We are building the OHDSI Evidence Network to give researchers faster access to real-world data. This is an ongoing priority in OHDSI, and we are excited about the early progress. We share a responsibility to empower network studies so we can generate reliable real-world evidence.

Data partners from across 15 nations and six continents have already joined the Evidence Network, and we are always looking to add more. You can learn more at our website:

ohdsi.github.io/EvidenceNetwork.



IQVIA
Contact us: iqvia@ohdsi.org

Country	Source Name	Patient Count
Australia	Australia EMR	2.7M
Belgium	IQVIA Belgium LPD	1.1M
France	France LPD	17.4M
France	IQVIA France DA	6.2M
Germany	IQVIA Germany DA	40.8M
Spain	LPD Spain	2.7M
United Kingdom	UK IMRD EHIS	5.1M
United States	US Hospital	313.1M
United States	US Open Claims	829.8M
United States	US PharMetrics Plus	170.2M

Johnson & Johnson
Contact us: jjj@ohdsi.org

Country	Source Name	Patient Count
Japan	JMDC	17.6M
United States	Merative MDCR	11.3M
United States	Merative CCAE	172.2M
United States	Merative MDCD	36.1M
United States	Premier	338.4M
United States	Optum ClinFormatics	99.3M
United States	Optum EHR	114.4M

Boehringer Ingelheim
Contact us: bi@ohdsi.org

Country	Source Name	Patient Count
Japan	JMDC	22.6M
United States	Optum ClinFormatics	85.1M
United States	Optum Market Clarity	90M

University of New Mexico
Contact us: unm@ohdsi.org

Country	Source Name	Patient Count
United States	Merative MDCR	11.2M
United States	Merative CCAE	170.7M

Ajou University
Contact us: ajou@ohdsi.org

Country	Source Name	Patient Count
Korea	Ajou University School of Medicine	2.7M

Clinical Hospital Center Zvezdara
Contact us: chcz@ohdsi.org

Country	Source Name	Patient Count
Serbia	CHCZ Zvezdara	618K

Columbia University
Contact us: culmc@ohdsi.org

Country	Source Name	Patient Count
United States	Columbia University Irving Medical Center Data	7M

Emory University
Contact us: emory@ohdsi.org

Country	Source Name	Patient Count
United States	Emory University	6.5M

GUSTO Singapore Cohort
Contact us: gusto@ohdsi.org

Country	Source Name	Patient Count
Singapore	GUSTO Singapore Cohort	2.6K

Join The Evidence Network!

We are proud of the data partners who have joined the Evidence Network, but we are excited to grow. If you want to join this group dedicated to improving healthcare globally, please reach out to evidencenetwork@ohdsi.org

HealthPartners Institute
Contact us: hpi@ohdsi.org

Country	Source Name	Patient Count
United States	HealthPartners Data	3.2M

Johns Hopkins University
Contact us: jhm@ohdsi.org

Country	Source Name	Patient Count
United States	Johns Hopkins Medicine	2.2M

Lancashire Teaching Hospitals NHS Foundation Trust
Contact us: lsc@ohdsi.org

Country	Source Name	Patient Count
United Kingdom	IDRL	1.5M

Papageorgiou General Hospital
Contact us: csteph@ohdsi.org

Country	Source Name	Patient Count
Greece	Papageorgiou General Hospital	1.4M

Penn State Health
Contact us: PSU@ohdsi.org

Country	Source Name	Patient Count
United States	Penn State Health	8.7M

Precision Data
Contact us: precisiondata@ohdsi.org

Country	Source Name	Patient Count
Brazil	DATASUS	8.7M

Semmelweis University
Contact us: semmelweis@ohdsi.org

Country	Source Name	Patient Count
Hungary	Semmelweis University Clinical Database	1.9M

Seoul National University Bundang Hospital
Contact us: snubh@ohdsi.org

Country	Source Name	Patient Count
Korea	Seoul National University Bundang Hospital	2.1M

Seoul National University Hospital
Contact us: snuh@ohdsi.org

Country	Source Name	Patient Count
Korea	Seoul National University Hospital	2.1M

SMG-SNU Boramae Medical Center
Contact us: boramae@ohdsi.org

Country	Source Name	Patient Count
Korea	SMG-SNU Boramae Medical Center	1M

Stanford University
Contact us: stanford@ohdsi.org

Country	Source Name	Patient Count
United States	Stanford Research Repository	3.8M

Taipei Medical University
Contact us: tmu@ohdsi.org

Country	Source Name	Patient Count
Taiwan	TMU Clinical Research Database	3.6M

Tufts University
Contact us: tmu@ohdsi.org

Country	Source Name	Patient Count
United States	Tufts Research Data	3.9M

University of Colorado Anschutz Medical Campus
Contact us: cuanschutz@ohdsi.org

Country	Source Name	Patient Count
United States	UCHealth	4.8M

University of Massachusetts Chan Medical Center
Contact us: umassmed@ohdsi.org

Country	Source Name	Patient Count
United States	UMass Medical	3.4M

University of Southern California
Contact us: usc@ohdsi.org

Country	Source Name	Patient Count
United States	USC Keck Medical	883K

University of Texas Southwestern
Contact us: utsw@ohdsi.org

Country	Source Name	Patient Count
United States	UT Southwestern	5.5M

Veteran's Affairs
Contact us: va@ohdsi.org

Country	Source Name	Patient Count
United States	VINCI	26.5M

Yonsei University Hospital
Contact us: yuhs@ohdsi.org

Country	Source Name	Patient Count
Korea	Yonsei University Hospital	6.4M

OMOP and FHIR

When the OMOP CDM and FHIR (a healthcare data exchange standard) work together, they improve healthcare data sharing and research. OMOP organizes large amounts of patient data from different sources, while FHIR ensures secure and standardized data exchange between systems. Combining the two allows researchers and healthcare providers to access and use patient data more efficiently, improving precision medicine, personalized treatments, and overall patient care. It also enhances collaboration across hospitals, research institutions, and healthcare technologies, leading to better outcomes and innovations in medical care.

2025 Update

In September 2025, the OHDSI and HL7 communities reached a historic milestone with the successful ballot of the FHIR to OMOP Implementation Guide (IG) v1.0, marking a foundational, formally vetted standard for transforming healthcare data between these two critical frameworks.

Why This Matters

This IG bridges two dominant open-source standards, creating a force multiplier for global health research and real-world evidence generation. With 71% of countries actively using FHIR for national healthcare initiatives and over 950 million patient records stored in OMOP format, standardized transformations between the advances OHDSI's RW evidence generation goals. The IG provides standardization of core-canonical EHR data element transformations to OMOP, creating more uniform transformations from FHIR and reducing redundant ETL efforts.

What Was Accomplished

After 2+ years of collaboration involving 100+ participants from 80+ organizations, in partnership with and sponsored by the Vulcan FHIR accelerator, the working group delivered:

Pragmatic Guidance: Best practices distilled from proven implementations; Detailed concept mapping principles with decision logic; Three common mapping patterns for any transformation scenario

Technical Artifacts: OMOP CDM logical models expressed in FHIR; Machine-readable Structure Maps for automated transformations; Validation packages tested through Connectathons; Establishment of Echidna FHIR Terminology Server for OHDSI vocabularies



The IG has just completed its initial Informative Ballot round! Already the FHIR-to-OMOP IG is increasingly being cited as an authoritative resource across the healthcare data community. This recognition validates our team's efforts and the broader community working to bridge clinical care and research data standards.

As we move forward, we're actively incorporating feedback from HL7 members and community stakeholders to refine and strengthen the guide. If you're interested in shaping this important work, we'd love to hear from you — links to comment and resources can be found under the Support tab in the IG. Your insights help us improve the IG before it moves forward as a Standard for Trial Use (STU) in an upcoming HL7 ballot.

"The FHIR to OMOP IG proves that when communities collaborate, we accelerate the journey from data to evidence to better health outcomes."



- Davera Gabriel
FHIR to OMOP IG Co-Author

"Together, we're not just mapping data formats — we're building the infrastructure for tomorrow's medical breakthroughs."



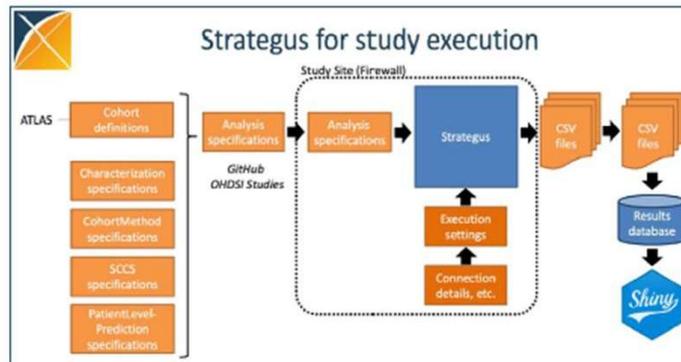
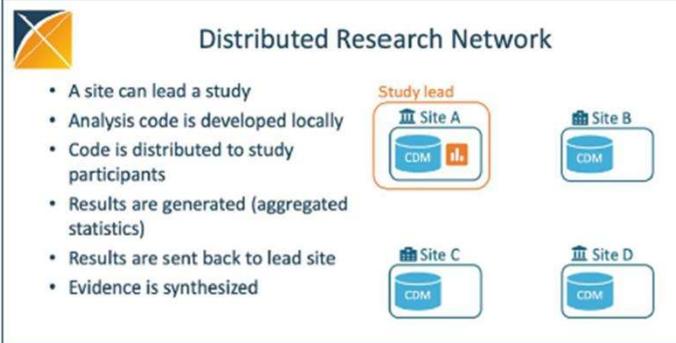
- Jean Duteau
FHIR to OMOP IG Co-Author

- **達成された成果**
- Vulcan FHIRアクセラレーターの支援のもと、80以上の組織から100名以上が参加した2年以上にわたる協働により、本ワーキンググループは以下を実現した。
- **実践的ガイダンス**：実証済み実装から抽出したベストプラクティス、意思決定ロジックを含む詳細なコンセプトマッピング原則、あらゆる変換シナリオに対応する3つの共通マッピングパターン
- **技術成果物**：FHIRで表現されたOMOP CDM論理モデル、自動変換のための機械可読なStructureMap、コネクタソンで検証されたバリデーションパッケージ、OHDSI語彙向けのEchidna FHIR Terminology Serverの構築



Strategus

The OHDSI Strategus package is a tool that helps researchers run large-scale health data studies more easily and consistently. Instead of needing to piece together different programs or workflows, Strategus provides a standardized way to design, execute, and share analyses across many healthcare databases. This makes it faster and more reliable to generate evidence about treatments, diseases, and outcomes.



- OHDSIの **Strategus** パッケージは、大規模な医療データ研究を、より容易かつ一貫性をもって実施するためのツールである。個別に複数のプログラムやワークフローを組み合わせる必要はなく、Strategusを用いることで、複数の医療データベースにまたがる解析を、標準化された方法で設計・実行・共有できる。これにより、治療、疾患、アウトカムに関するエビデンスを、より迅速かつ信頼性高く創出することが可能となる。