

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



OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

オデッセイ
ジャパン

OHDSI Japan evening conference #28

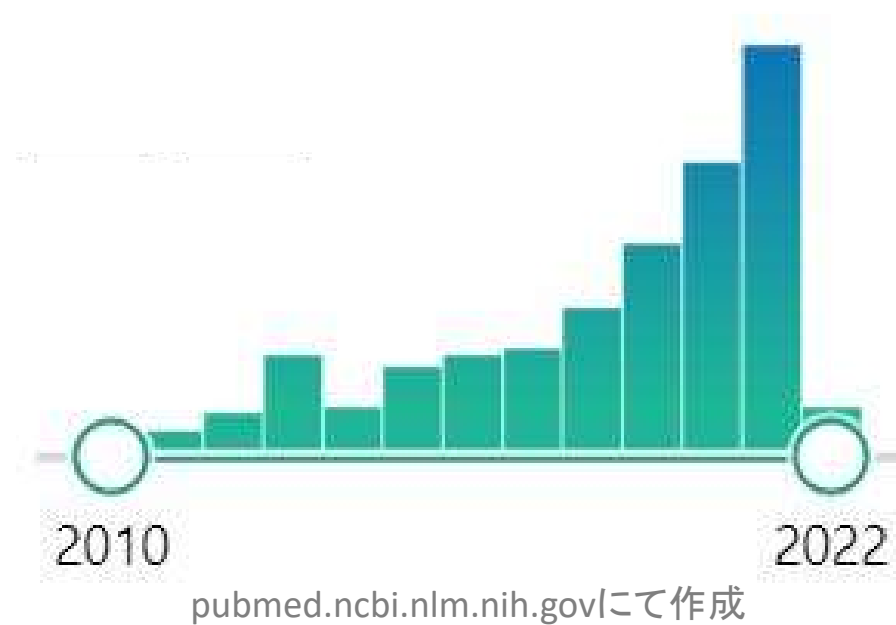
イブニングカンファレンス(第28回)

2022.3.29



OHDSI関連論文

Pubmedで“OHDSI or OMOP”を検索



全期間累計：2月236本→3月244本



認知症外来患者における特徴の性差について

> [BMC Psychiatry](#). 2022 Mar 21;22(1):208. doi: 10.1186/s12888-022-03852-z.

Gender differences in outpatients with dementia from a large psychiatric hospital in China

Jiaojiao Zhou ^{# 1}, Chengwei Guo ^{# 1}, Li Ren ¹, Dandi Zhu ¹, Wenfeng Zhen ¹, Saina Zhang ¹, Qing'e Zhang ²

Affiliations [+ expand](#)

PMID: 35313835 PMCID: PMC8935692 DOI: 10.1186/s12888-022-03852-z

Background: The sociodemographic characteristics and clinical features of dementia patients in psychiatric hospitals have not been thoroughly studied in China. This study aimed to explore the psychiatric outpatient attendance of dementia patients at a psychiatric hospital in China, with particular emphasis on gender differences.

Methods: This retrospective study examined outpatients with dementia from January 2013 to August 2019 using data in the Observational Medical Outcomes Partnership Common Data Model (OMOP-CDM) in Beijing Anding Hospital. Age, sex, number of visits, use of drugs and comorbid conditions were extracted from medical records.

Results: Nine thousand four patients were recruited from a specific outpatient clinic of a hospital in Beijing, and the mean number of visits was 6.92. There were 3,433 (38.13%) male patients and 5,571 (61.87%) female patients. The most common comorbidities were generalized anxiety disorder, nonorganic insomnia, delusional disorder and depressive disorder. The proportion of patients using antidementia was the highest, with the rate of 68.3%, followed by benzodiazepines (48.83%), antipsychotics (45.43%), antidepressants (22.24%) and nonbenzodiazepines (19.96%). Patients with dementia showed a significant gender difference in average age ($t = 6.36, P < 0.0001$). Compared to male patients, female patients had a higher number of visits (7.40 ± 12.90 vs $6.15 \pm 10.50, t = 4.81, P < 0.0001$). There were significant differences in comorbidity composition between male and female patients ($t = 23.09, P < 0.0001$).

Conclusions: Our present findings suggested significant gender differences in the proportion of age, number of visits and comorbidity composition in outpatients with dementia.

背景：精神科病院における認知症患者の社会人口統計学的特徴および臨床的特徴については、中国では十分に研究されていない。本研究は、中国のある精神科病院における認知症患者の精神科外来通院状況について、特に性差に着目して調査することを目的とした。

方法：本研究は、北京安鼎病院のOMOP-CDMのデータを用いて、2013年1月から2019年8月までの認知症外来患者を後方視的に検討したものである。年齢、性別、受診回数、薬剤の使用状況、併存する疾患は診療記録から抽出した。

結果：北京の病院の特定外来から9,004名の患者を募集し、平均受診回数は6.92回であった。男性患者3,433人（38.13%）、女性患者5,571人（61.87%）であった。最も多かった併存疾患は、全般性不安障害、非器質性不眠症、妄想性障害、うつ病性障害であった。抗精神病薬を使用している患者の割合は68.3%と最も高く、ベンゾジアゼピン系（48.83%）、抗精神病薬（45.43%）、抗うつ薬（22.24%）、非ベンゾジアゼピン系（19.96%）の順となりました。認知症患者の平均年齢には有意な性差が認められた（ $t = 6.36, P < 0.0001$ ）。男性患者に比べ、女性患者は受診回数が多かった（ 7.40 ± 12.90 vs $6.15 \pm 10.50, t = 4.81, P < 0.0001$ ）。併存疾患の構成は、男性患者と女性患者で有意差があった（ $t = 23.09, P < 0.0001$ ）。

結論：今回の結果から、認知症外来患者における年齢、受診回数、併存疾患構成の割合に有意な性差があることが示唆された。



> AMIA Annu Symp Proc. 2022 Feb 21;2021:989-998. eCollection 2021.

Extracting Patient-level Social Determinants of Health into the OMOP Common Data Model

Jimmy Phuong^{1 2}, Elizabeth Zampino^{1 2}, Nicholas Dobbins^{1 2}, Juan Espinoza³, Daniella Meeker⁴, Heidi Spratt⁵, Charisse Madlock-Brown⁶, Nicole G Weiskopf⁷, Adam Wilcox¹

Affiliations – collapse

Affiliations

- 1 Division of Biomedical and Health Informatics, UW Medicine, Seattle, Washington.
- 2 University of Washington Medicine Research IT, Seattle, Washington.
- 3 Department of Pediatrics, Children's Hospital Los Angeles, Los Angeles, CA.
- 4 Department of Preventive Medicine, University of Southern California, Los Angeles, California.
- 5 Preventative Medicine and Population Health, University of Texas Medical Branch, Galveston, Texas.
- 6 Dept of Health Informatics and Information Management, University of Tennessee Health Science Center, Memphis, Tennessee.
- 7 Department of Medical Informatics and Clinical Epidemiology, OHSU, Portland, Oregon.

Deficiencies in data sharing capabilities limit Social Determinants of Health (SDoH) analysis as part of COVID-19 research. The National COVID Cohort Collaborative (N3C) is an example of an Electronic Health Record (EHR) database of patients tested for COVID-19 that could benefit from a SDoH elements framework that captures various screening instruments in EHR data warehouse systems. This paper uses the University of Washington Enterprise Data Warehouse (a data contributor to N3C) to demonstrate how SDoH can be represented and managed to be made available within an OMOP common data model. We found that these data varied by type of social determinants data and where it was collected, in the time period that it was collected, and in how it was represented.

データ共有能力の不足により、COVID-19研究の一環である健康の社会的決定要因（SDoH）分析が制限されている。National COVID Cohort Collaborative（N3C）は、COVID-19の検査を受けた患者の電子カルテ（EHR）データベースの一例であり、EHRデータウェアハウスシステムにおける様々なスクリーニング手段を取り込むSDoH要素フレームワークの恩恵を受けることができます。この論文では、ワシントン大学のエンタープライズデータウェアハウス（N3Cのデータ提供者）を使用して、**SDoHをどのように表現し、OMOP共通データモデル内で利用可能になるように管理できるか**を示しています。これらのデータは、社会的決定要因のデータの種類や収集場所、収集期間、表現方法によって異なることがわかりました。

EHR内の調査票結果（EpicClarityFlowsheet）
今回は健康に関する社会的要因に注目



- ・ 構造を分析して取り出し
- ・ 項目と値をOMOP標準にマッピング

OMOP の Observation Table に格納



遺伝子バリエーションデータの有無を問い合わせる
ネットワーク

> Hum Mutat. 2022 Mar 17. doi: 10.1002/humu.24369. Online ahead of print.

Beacon v2 and Beacon Networks: a "lingua franca" for federated data discovery in biomedical genomics, and beyond

Jordi Rambla ^{1 2}, Michael Baudis ³, Roberto Ariosio ¹, Tim Beck ⁴, Lauren A Fromont ¹,
Arcadi Navarro ^{1 5 6 7}, Rahel Paloots ³, Manuel Rueda ¹, Gary Saunders ⁸, Babita Singh ¹,
John Dylan Spalding ⁹, Juha Törnroos ⁹, Claudia Vasallo ¹, Colin D Veal ⁴, Anthony J Brookes ⁴

Affiliations + expand

PMID: 35297548 DOI: 10.1002/humu.24369

Beacon is a basic data discovery protocol issued by the Global Alliance for Genomics and Health (GA4GH). The main goal addressed by version 1 of the Beacon protocol was to test the feasibility of broadly sharing human genomic data, through providing simple "yes" or "no" responses to queries about the presence of a given variant in datasets hosted by Beacon providers. The popularity of this concept has fostered the design of a version 2, that better serves real world requirements and addresses the needs of clinical genomics research and healthcare, as assessed by several contributing projects and organizations. Particularly, rare disease genetics and cancer research will benefit from new case level and genomic variant level requests and the enabling of richer phenotype and clinical queries as well as support for fuzzy searches. Beacon is designed as a "lingua franca" to bridge data collections hosted in software solutions with different and rich interfaces. Beacon version 2 works alongside popular standards like Phenopackets, OMOP or FHIR, allowing implementing consortia to return matches in beacon responses and provide a handover to their preferred data exchange format. The protocol is being explored by other research domains and is being tested in several international projects.

Beaconは、Global Alliance for Genomics and Health (GA4GH)が発行する基本データ探索プロトコルです。Beaconのバージョン1の主な目的は、Beaconプロバイダーがホストするデータセット内の特定のバリエーションの存在に関するクエリーに「はい」または「いいえ」の単純な応答を提供することにより、ヒトゲノムデータを広く共有することの実現可能性をテストすることでした。このコンセプトが好評を博したため、いくつかのプロジェクトや組織によって評価されたように、実世界の要件や臨床ゲノム研究および医療のニーズに対応するバージョン2の設計が促進されました。特に、希少疾患の遺伝学や癌の研究では、症例レベルやゲノムバリエーションレベルのリクエスト、より豊富な表現型や臨床クエリーの実現、ファジー検索のサポートなどの恩恵を受けることができます。Beaconは、異なる豊富なインターフェースを持つソフトウェアソリューションにホストされているデータコレクションを橋渡しする「共通言語」として設計されています。Beaconバージョン2は、Phenopackets、OMOP、FHIRといった一般的な標準規格とともに動作し、実装するコンソーシアムがBeaconレスポンスでマッチングを返し、好みのデータ交換フォーマットへのハンドオーバーを提供できるようにします。このプロトコルは、他の研究領域でも検討されており、いくつかの国際プロジェクトでテストされています。

What's a Beacon?

Beacon is a genetic mutation sharing platform developed by the Global Alliance for Genomics and Health. A beacon is web service that any institution can implement to share genetic data. A beacon answers questions of the form "Do you have information about the following mutation?" and responds with one of "Yes" or "No", among potentially more information.

Query:

GET [https://beacon-network.org/api/responses?](https://beacon-network.org/api/responses?chrom=17&pos=41244981&allele=G&ref=GRCh37&beacon=[amplab,brca-exchange])[chrom=17&pos=41244981&allele=G&ref=GRCh37&beacon=\[amplab,brca-exchange\]](https://beacon-network.org/api/responses?chrom=17&pos=41244981&allele=G&ref=GRCh37&beacon=[amplab,brca-exchange])

```
[
  {
    "beacon": {
      "id": "amplab",
      "name": "AMPLab - 1000 Genomes Project",
      "url": null,
      "organization": "AMPLab, UC Berkeley",
      "description": null,
      "homePage": null,
      "email": null,
      "aggregator": false,
      "visible": true,
      "enabled": true,
      "supportedReferences": [
        "HG38",
        "HG18",
        "HG19"
      ]
    },
    "query": {
      "chromosome": "CHR17",
      "position": 41244981,
      "allele": "G",
      "reference": "HG19"
    },
    "response": true,
    "frequency": null,
    "externalUrl": null,
    "info": null
  },
  {
    "beacon": {
      "id": "brca-exchange",
      "name": "BRCA Exchange",
      "url": null,
      "organization": "BRCA Exchange",
      "description": "Public variant annotations from BRCA Exchange.",
      "homePage": null,
      "email": null,
      "aggregator": false,
      "visible": true,
      "enabled": true,
      "supportedReferences": [
        "HG19"
      ]
    },
    "query": {
      "chromosome": "CHR17",
      "position": 41244981,
      "allele": "G",
      "reference": "HG19"
    },
    "response": true,
    "frequency": null,
    "externalUrl": null,
    "info": {
      "Clinical_significance_citations":
        "PMID:21990134",
      "HGVS_cDNA": "c.2566T>C",
      "Condition_ID_type": "OMIM",
      "Date_last_evaluated": "8/10/15",
      "HGVS_protein": "p.(Tyr856His)",
      "Assertion_method_citation":
        "http://enigmaconsortium.org/documents/ENIGMA_Rules_2015-03-26.pdf",
      "ClinVarAccession": "SCV000244323",
      "alt": "G",
      "Condition_category": "Disease",
      "Comment_on_clinical_significance": "IARC class based on posterior probability from multifactorial likelihood analysis, thresholds for class as per Plon et al. 2008 (PMID: 18951446). Class 1 based on posterior probability = 0.00018. Also class 1 based on frequency >1% in an outbred sample set. Frequency 0.01399 (Asian), derived from 1000 genomes (2012-04-30).",
      "Reference_sequence": "NM_007294.3",
      "Gene_symbol": "BRCA1",
      "URL": "http://hci-exlovd.hci.utah.edu/variants.php?select_db=BRCA1&action=search_all&search_Variant%2FDNA=c.2566T>C",
      "chrom": "17",
      "Abbrev_AA_change": "Y856H",
      "ref": "A",
      "Assertion_method": "ENIGMA BRCA1/2 Classification Criteria (2015)",
      "BIC_Nomenclature": "2685T>C",
      "Clinical_significance": "Benign",
      "Genomic_Coordinate": "chr17:41244982:A>G",
      "Allele_origin": "Germline",
      "position": "41244982",
      "Collection_method": "Curation",
      "Condition_ID_value": "BREAST-OVARIAN CANCER, FAMILIAL, SUSCEPTIBILITY TO, 1; BROVCA1 (604370)"
    }
  }
]
```



アルツハイマー型認知症（AD）患者における
抗認知症薬の処方パターン

➤ Sci Rep. 2022 Mar 15;12(1):4451. doi: 10.1038/s41598-022-08595-1.

**Analysis of treatment pattern of anti-dementia
medications in newly diagnosed Alzheimer's
dementia using OMOP CDM**

JungHyun Byun¹, Dong Yun Lee², Chang-Won Jeong^{3 4}, Yerim Kim⁵, Hak Young Rhee⁶,
Ki Won Moon^{7 8}, Jeongwon Heo^{7 8}, Yoonki Hong^{7 8}, Woo Jin Kim^{7 8}, Seung-Joo Nam^{7 8},
Hoon Sung Choi^{7 8}, Ji In Park^{7 8}, In Kook Chun^{7 9}, So Hyeon Bak^{7 10}, Kyoungyul Lee^{7 11},
Gi Hwan Byeon^{7 12}, Kyoung Lae Kim^{7 13}, Jeong-Ah Kim^{7 13}, Young Joo Park^{7 13},
Jeong Hyun Kim^{7 14}, Eun Ju Lee¹⁵, Sang-Ah Lee¹⁶, Sung Ok Kwon¹⁷, Sang-Won Park^{17 18},
Payam Hosseinzadeh Kasani^{17 18}, Jung-Kyeom Kim^{17 18}, Yeshin Kim^{7 18}, Seongheon Kim^{7 18},
Jae-Won Jang^{19 20 21}

Affiliations [+](#) expand

PMID: 35292697 PMCID: PMC8924152 DOI: 10.1038/s41598-022-08595-1

Anti-dementia medications are widely prescribed to patients with Alzheimer's dementia (AD) in South Korea. This study investigated the pattern of medical management in newly diagnosed patients with AD using a standardized data format-the Observational Medical Outcome Partnership Common Data Model from five hospitals. We examined the anti-dementia treatment patterns from datasets that comprise > 5 million patients during 2009-2019. The medication utility information was analyzed with respect to treatment trends and persistence across 11 years. Among the 8653 patients with newly diagnosed AD, donepezil was the most commonly prescribed anti-dementia medication (4218; 48.75%), followed by memantine (1565; 18.09%), rivastigmine (1777; 8.98%), and galantamine (494; 5.71%). The rising prescription trend during observation period was found only with donepezil. The treatment pathways for the three cholinesterase inhibitors combined with N-methyl-D-aspartate receptor antagonist were different according to the drugs (19.6%; donepezil; 28.1%; rivastigmine, and 17.2%; galantamine). A 12-month persistence analysis showed values of approximately 50% for donepezil and memantine and approximately 40% for rivastigmine and galantamine. There were differences in the prescribing pattern and persistence among anti-dementia medications from database using the Observational Medical Outcome Partnership Common Data Model on the Federated E-health Big Data for Evidence Renovation Network platform in Korea.

韓国では、アルツハイマー型認知症（AD）患者に抗認知症薬が広く処方されている。本研究では、標準化されたデータフォーマット（OMOP CDM）を用いて、**新たにADと診断された患者の医療管理パターンを5つの病院から調査した**。2009年から2019年の間に500万人以上の患者からなるデータセットから、抗認知症治療のパターンを検討した。薬物有用性情報は、11年間の治療傾向および持続性に関して分析した。新たにADと診断された8653人の患者のうち、最もよく処方された抗認知症薬はドネペジル（4218、48.75%）、次いでメマンチン（1565、18.09%）、リバスチグミン（1777、8.98%）、ガラントミン（494、5.71%）であった。なお、観察期間中に処方数が増加したのは、ドネペジルのみであった。コリンエステラーゼ阻害剤3剤とN-methyl-D-aspartate受容体拮抗剤との併用療法は、薬剤によって治療経路が異なっていた（ドネペジル19.6%、リバスチグミン28.1%、ガラントミン17.2%）。12ヶ月の持続性解析では、ドネペジルとメマンチンで約50%、リバスチグミンとガラントミンで約40%という値が示された。韓国のFederated E-health Big Data for Evidence Renovation Network(FEEDER-NET)プラットフォーム上のOMOP CDMを用いたデータベースから、抗認知症薬の処方パターンと継続性に差異があることが示された。



成人てんかん患者における
抗けいれん薬治療パスの特徴調査

➤ [Epilepsy Behav. 2022 Mar 8;129:108630. doi: 10.1016/j.yebeh.2022.108630. Online ahead of print.](#)

Patient characteristics and antiseizure medication pathways in newly diagnosed epilepsy: Feasibility and pilot results using the common data model in a single-center electronic medical record database

Matthew Spotnitz¹, Anna Ostropolets¹, Victor G Castano², Karthik Natarajan¹,
Genna J Waldman³, Michael Argenziano², Ruth Ottman⁴, George Hripcsak¹, Hyunmi Choi³,
Brett E Youngerman⁵

Affiliations + expand

PMID: 35276502 DOI: [10.1016/j.yebeh.2022.108630](#)

Introduction: Efforts to characterize variability in epilepsy treatment pathways are limited by the large number of possible antiseizure medication (ASM) regimens and sequences, heterogeneity of patients, and challenges of measuring confounding variables and outcomes across institutions. The Observational Health Data Science and Informatics (OHDSI) collaborative is an international data network representing over 1 billion patient records using common data standards. However, few studies have applied OHDSI's Common Data Model (CDM) to the population with epilepsy and none have validated relevant concepts. The goals of this study were to demonstrate the feasibility of characterizing adult patients with epilepsy and ASM treatment pathways using the CDM in an electronic health record (EHR)-derived database.

Methods: We validated a phenotype algorithm for epilepsy in adults using the CDM in an EHR-derived database (2001-2020) against source records and a prospectively maintained database of patients with confirmed epilepsy. We obtained the frequency of all antecedent conditions and procedures for patients meeting the epilepsy phenotype criteria and characterized ASM exposure sequences over time and by age and sex.

Results: The phenotype algorithm identified epilepsy with 73.0-85.0% positive predictive value and 86.3% sensitivity. Many patients had neurologic conditions and diagnoses antecedent to meeting epilepsy criteria. Levetiracetam incrementally replaced phenytoin as the most common first-line agent, but significant heterogeneity remained, particularly in second-line and subsequent agents. Drug sequences included up to 8 unique ingredients and a total of 1,235 unique pathways were observed.

Conclusions: Despite the availability of additional ASMs in the last 2 decades and accumulated guidelines and evidence, ASM use varies significantly in practice, particularly for second-line and subsequent agents. Multi-center OHDSI studies have the potential to better characterize the full extent of variability and support observational comparative effectiveness research, but additional work is needed to validate covariates and outcomes.

はじめに: てんかんの治療経路のばらつきを明らかにするための努力は、可能な抗けいれん薬 (ASM) レジメンとシーケンスの数が多く、患者の不均一性、施設間での交絡変数と転帰の測定という課題によって制限されている。OHDSI共同研究は、共通のデータ標準を用いた10億以上の患者記録を代表する国際的なデータネットワークです。しかし、OHDSIの共通データモデル (CDM) をてんかん患者群に適用した研究はほとんどなく、関連する概念の検証も行われていない。本研究の目的は、電子カルテ (EHR) 由来のデータベースにおいて、**CDMを用いて成人てんかん患者およびASM治療パスの特徴を明らかにできるかを実証**することである。

方法: EHRデータベース (2001~2020年) において、CDMを用いた成人てんかんの表現型アルゴリズムを、原資料および前向きに維持されているてんかん確定患者のデータベースと比較し、検証した。てんかんの表現型基準を満たす患者のすべての先行疾患および処置の頻度を求め、経時的および年齢・性別ごとのASM曝露シーケンスの特徴を明らかにした。

結果: 表現型アルゴリズムは、73.0-85.0%の陽性適中率と86.3%の感度でてんかんを同定した。多くの患者は、てんかんの基準を満たす前に、神経学的状態や診断を受けていた。第一選択薬は、フェニトインに代わりレベチラセタムが主流となったが、特に第二選択薬以降には大きな不均質性が見られた。薬物配列には最大8種類の成分が含まれ、合計1,235種類のパスウェイが観察された。

結論: 過去20年間に追加のASMが利用可能となり、ガイドラインやエビデンスが蓄積されたにもかかわらず、ASMの使用は、特に二次治療薬およびそれ以降の薬剤において、実際には大きく異なっている。多施設共同OHDSI試験は、ばらつきの全容をより正確に把握し、観察的比較効果研究を支援する可能性があるが、共変量と転帰を検証するための追加作業が必要である。



看護記録をOMOP化し、
転倒リスク予測モデルを作成

> [JMIR Med Inform. 2022 Mar 11;10\(3\):e35104. doi: 10.2196/35104.](#)

Patient-Level Fall Risk Prediction Using the Observational Medical Outcomes Partnership's Common Data Model: Pilot Feasibility Study

Hyesil Jung ¹, Sooyoung Yoo ¹, Seok Kim ¹, Eunjeong Heo ¹, Borham Kim ¹, Ho-Young Lee ¹, Hee Hwang ²

Affiliations – collapse

Affiliations

- ¹ Office of eHealth Research and Business, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea.
- ² Kakao Healthcare Company-In-Company, Seongnam-si, Republic of Korea.

PMID: 35275076 DOI: [10.2196/35104](#)

Background: Falls in acute care settings threaten patients' safety. Researchers have been developing fall risk prediction models and exploring risk factors to provide evidence-based fall prevention practices; however, such efforts are hindered by insufficient samples, limited covariates, and a lack of standardized methodologies that aid study replication.

Objective: The objectives of this study were to (1) convert fall-related electronic health record data into the standardized Observational Medical Outcome Partnership's (OMOP) common data model format and (2) develop models that predict fall risk during 2 time periods.

Methods: As a pilot feasibility test, we converted fall-related electronic health record data (nursing notes, fall risk assessment sheet, patient acuity assessment sheet, and clinical observation sheet) into standardized OMOP common data model format using an extraction, transformation, and load process. We developed fall risk prediction models for 2 time periods (within 7 days of admission and during the entire hospital stay) using 2 algorithms (least absolute shrinkage and selection operator logistic regression and random forest).

Results: In total, 6277 nursing statements, 747,049,486 clinical observation sheet records, 1,554,775 fall risk scores, and 5,685,011 patient acuity scores were converted into OMOP common data model format. All our models (area under the receiver operating characteristic curve 0.692-0.726) performed better than the Hendrich II Fall Risk Model. Patient acuity score, fall history, age ≥ 60 years, movement disorder, and central nervous system agents were the most important predictors in the logistic regression models.

Conclusions: To enhance model performance further, we are currently converting all nursing records into the OMOP common data model data format, which will then be included in the models. Thus, in the near future, the performance of fall risk prediction models could be improved through the application of abundant nursing records and external validation.

背景: 急性期医療における転倒は、患者の安全を脅かす。研究者は、エビデンスに基づく転倒予防の実践を提供するために、**転倒リスク予測モデル**を開発し、危険因子を探求してきた。しかし、このような取り組みは、不十分なサンプル、限られた共変量、研究の再現を助ける標準的な方法論の欠如によって妨げられるものである。

目的: 本研究の目的は、(1) 転倒関連の電子カルテデータを、OMOP CDMに変換し、(2) 2つの期間中の転倒リスクを予測するモデルを開発することであった。

方法: 試験的なフィジビリティテストとして、転倒関連の電子カルテデータ（看護記録、転倒リスク評価表、患者急性期評価表、臨床観察表）を、抽出・変換・ロード(ETL)プロセスを用いて、標準化されたOMOP CDMに変換した。2つのアルゴリズム（最小絶対値縮小・選択演算子ロジスティック回帰、ランダムフォレスト）を用いて、2つの期間（入院7日以内、入院期間中）について転倒リスク予測モデルを作成した。

結果: 合計6277件の看護記述、747,049,486件の臨床観察票記録、1,554,775件の転倒リスクスコア、5,685,011件の患者急性期スコアがOMOP CDMに変換された。すべてのモデルはHendrich II転倒リスクモデルより良好な結果を示した（ROC-AUC 0.692-0.726）。ロジスティック回帰モデルでは、患者視カスコア、転倒歴、年齢 ≥ 60 歳、運動障害、中枢神経系薬剤が最も重要な予測因子であった。

結論: モデル性能をさらに向上させるため、現在、すべての看護記録をOMOP CDMのデータ形式に変換中であり、これをモデルに含める予定である。このように、近い将来、豊富な看護記録の適用と外部検証により、転倒リスク予測モデルの性能は向上する可能性がある。



> J Clin Transl Sci. 2021 Dec 9;6(1):e10. doi: 10.1017/cts.2021.880. eCollection 2022.

Ensuring a safe(r) harbor: Excising personally identifiable information from structured electronic health record data

Emily R Pfaff ¹, Melissa A Haendel ², Kristin Kostka ³, Adam Lee ⁴, Emily Niehaus ⁵,
Matvey B Palchuk ⁶, Kellie Walters ⁴, Christopher G Chute ⁷

Affiliations – collapse

Affiliations

- 1 Department of Medicine, UNC Chapel Hill School of Medicine, Chapel Hill, North Carolina, USA.
- 2 University of Colorado Anschutz Medical Campus, Aurora, Colorado, USA.
- 3 The OHDSI Center at the Roux Institute, Northeastern University, Portland, Maine, USA.
- 4 TraCS Institute, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA.
- 5 Palantir Technologies, Denver, Colorado, USA.
- 6 TriNetX LLC, Cambridge, Massachusetts, USA.
- 7 Schools of Medicine, Public Health, and Nursing, Johns Hopkins University, Baltimore, Maryland, USA.

Recent findings have shown that the continued expansion of the scope and scale of data collected in electronic health records are making the protection of personally identifiable information (PII) more challenging and may inadvertently put our institutions and patients at risk if not addressed. As clinical terminologies expand to include new terms that may capture PII (e.g., Patient First Name, Patient Phone Number), institutions may start using them in clinical data capture (and in some cases, they already have). Once in use, PII-containing values associated with these terms may find their way into laboratory or observation data tables via extract-transform-load jobs intended to process structured data, putting institutions at risk of unintended disclosure. Here we aim to inform the informatics community of these findings, as well as put out a call to action for remediation by the community.

最近の知見では、電子カルテで収集されるデータの範囲と規模が拡大し続けているため、個人特定可能情報（PII）の保護が難しくなっており、対応しなければ私たちの施設と患者を不用意に危険にさらす可能性があることが示されています。臨床用語が拡大し、PIIを捕捉する可能性のある新しい用語（例：患者のファーストネーム、患者の電話番号）が含まれるようになると、医療機関は臨床データの収集にそれらを使い始めるかもしれません（場合によっては、すでに使用していることもあります）。いったん使用されると、これらの用語に関連するPIIを含む値が、構造化データの処理を目的とした抽出-変換-読み込み（ETL）作業によって、検査や観察データのテーブルに入り込み、意図しない開示のリスクにさらされる可能性がある。ここでは、これらの発見を情報科学コミュニティに知らせるとともに、コミュニティによる是正のための行動を呼びかけることを目的としています。

構造化データにも注意が必要

LOINCの個人情報項目例

- 42077-8 患者の自宅の電話
- 45401-7 患者の郵便番号
- 56799-0 患者の住所
- LOINCコードがついているものは
検査結果だ、と思い込んではいけない。
- フィルタリングが必要。

複数のアプローチがありえる

- ブラックリスト
→LOINCやOHDSIの協力を得て作成公開した。
メンテも行なう。
- ホワイトリスト
安全度が高い。
含まれないものは個別に判断。
- パターンによる判別
項目名や値
例：電話番号は不可、単なる数値はOK



最近のGlobal OHDSI話題から

DARWIN EU Coordination Centre

- EMAにより設立される
- Erasmus University Medical Center Rotterdam と契約(2022/2)
- OHDSI Europe のcoordinating center も同じMedical Centerにある。

Erasmus University Medical Center Rotterdam



google mapより

Data Analysis and Real World Interrogation Network

- ワクチンを含むヒト用医薬品の使用、安全性、有効性に関するエビデンスを、EU全域のリアルワールドの医療データベースからタイムリーに提供する。
- これにより、EMAおよび欧州医薬品規制ネットワーク内の各国所轄官庁は、医薬品のライフサイクルを通じて必要なときにいつでもこれらのデータを利用することができるようになる。
- 関連プロジェクトと連携
 - European medicines regulatory network by EMA
 - European Health Data Space (EHDS) by EC
- タイムライン
 - 2022年 設立&パイロットモード
 - 2023年 開発と利用開始
 - 2024年 フル稼働

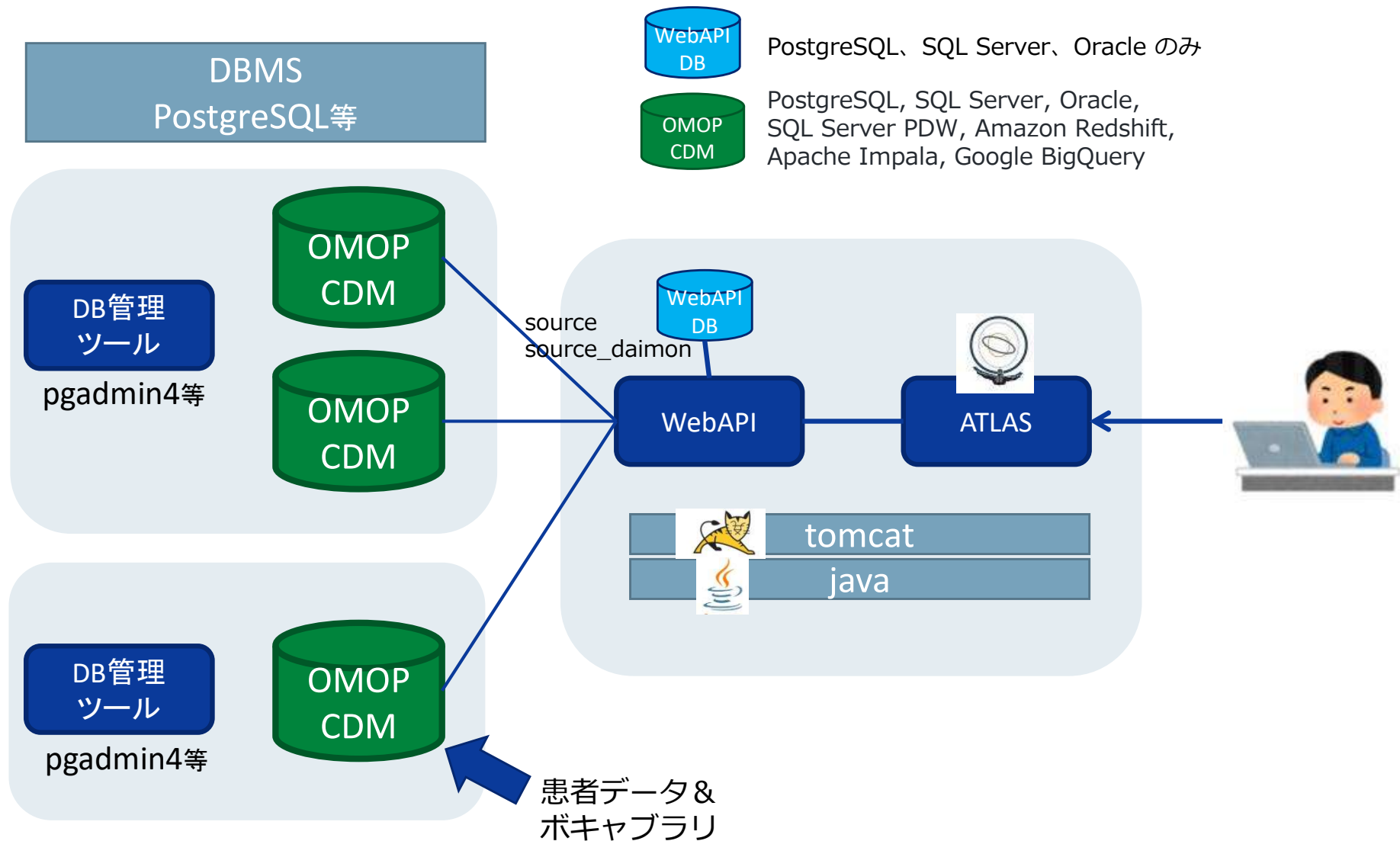


Inside OMOP #1 全体構成

- データベース
CDMデータベース、DB管理ツール
- データツール
Tomcat、WebAPI、ATLAS
- R系
R、R Studio、ACHELLES、shiny、DQD、HADES
- ETL設計系
WhiteRabbit、Rabbit-in-a-Hat、Usagi



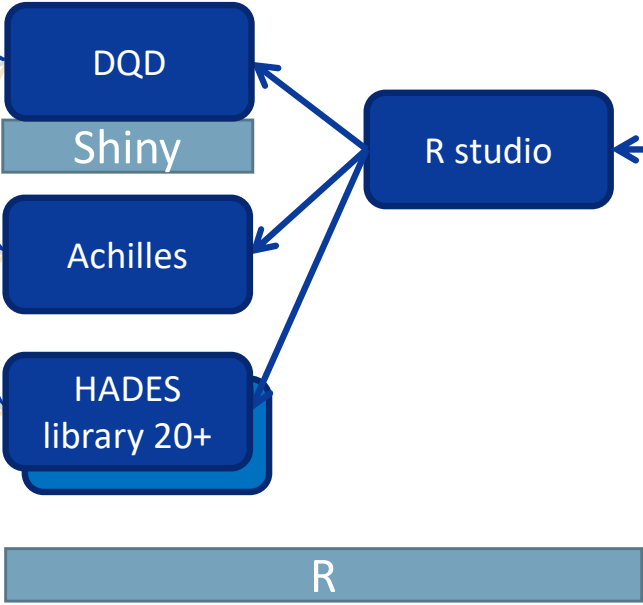
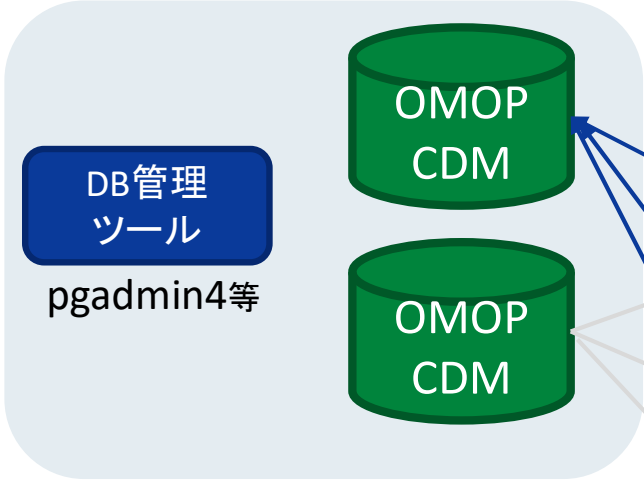
データベース&データツール





R系

DBMS
PostgreSQL等



HADES HEALTH ANALYTICS DATA-TO-EVIDENCE SUITE				
CohortMethod New-user cohort studies using large-scale regression for propensity and outcome models. Learn more...	SelfControlledCaseSeries Self-Controlled Case Series analysis using few or many predictors, includes splines for age and seasonality. Learn more...	Cyclops Highly efficient implementation of regularized logistic, Poisson and Cox regression. Learn more...	DatabaseConnector Connect directly to a wide range of database platforms, including SQL Server, Oracle, and PostgreSQL. Learn more...	SqlRender Generate SQL on the fly for the various SQL dialects. Learn more...
SelfControlledCohort A self-controlled cohort design, where time preceding exposure is used as control. Learn more...	EvidenceSynthesis Routines for combining causal effect estimates and study diagnostics across multiple data sites in a distributed study. Learn more...	ParallelLogger Support for parallel computation with logging to console, disk, or e-mail. Learn more...	FeatureExtraction Automatically extract large sets of features for user-specified cohorts using data in the CDM. Learn more...	Andromeda Storing very large data objects on a local drive, while still making it possible to manipulate the data in an efficient manner. Learn more...
PatientLevelPrediction Build and evaluate predictive models for user-specified outcomes, using a wide array of machine learning algorithms. Learn more...	EmpiricalCalibration Use negative control exposure-outcome pairs to profile and calibrate a particular analysis design. Learn more...	BigKnn A large scale k-nearest neighbor classifier using the Lucene search engine. Learn more...	ROhdsWebApi Interact with OHDSI WebAPI web services. Learn more...	OhdsiSharing Securely sharing (large) files between OHDSI collaborators. Learn more...
MethodEvaluation Use real data and established reference sets as well as simulations injected in real data to evaluate the performance of methods. Learn more...	CohortDiagnostics Generate a wide set of diagnostics to evaluate cohort definitions against databases in the CDM. Learn more...	Hydra Hydrating package skeletons into executable R study packages based on specifications in JSON format. Learn more...	Economia A standard CDM dataset for testing and demonstration purposes that runs on an embedded SQLite database. Learn more...	CirceR An R wrapper for Circe, a library for creating cohort definitions, expressing them as JSON, SQL, or Markdown. Learn more...

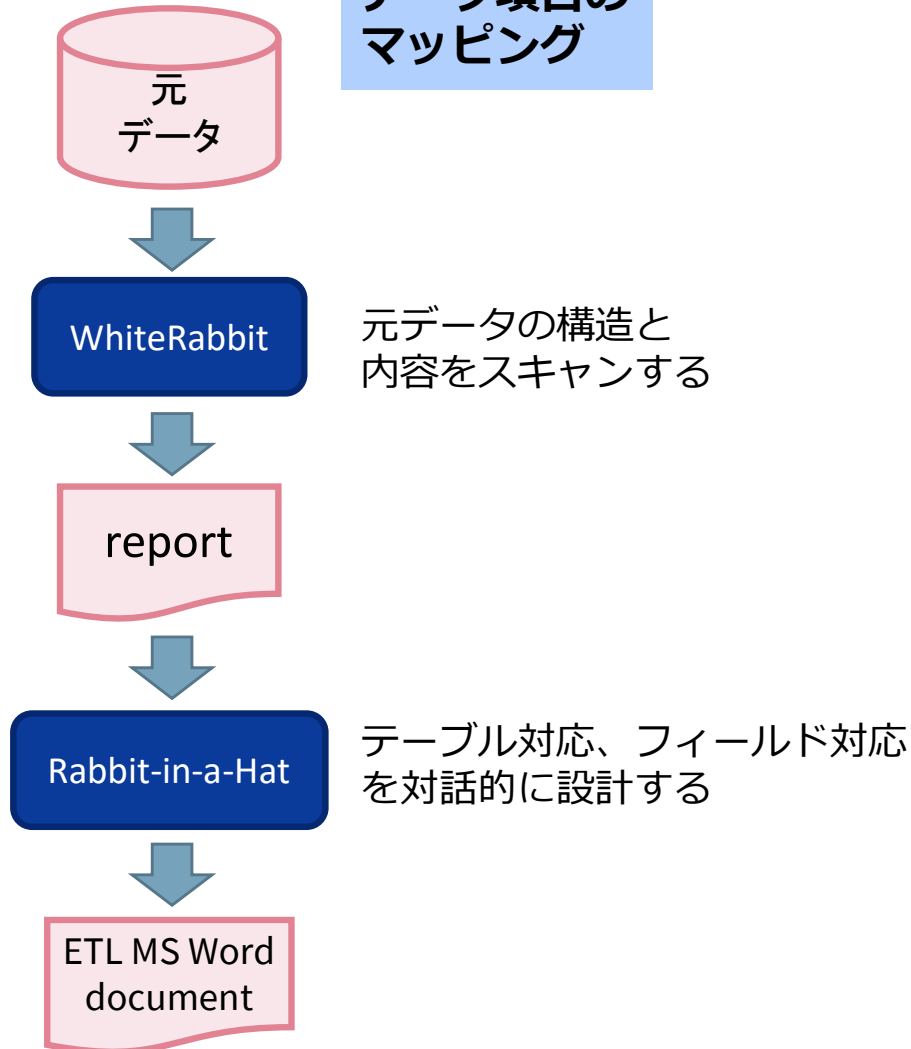


ETL設計系



javaで書かれたdesktopツール

データ項目のマッピング



データコードのマッピング

