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 Japan evening conference #34

イブニング カンファレンス(第34回) 2022.9.30

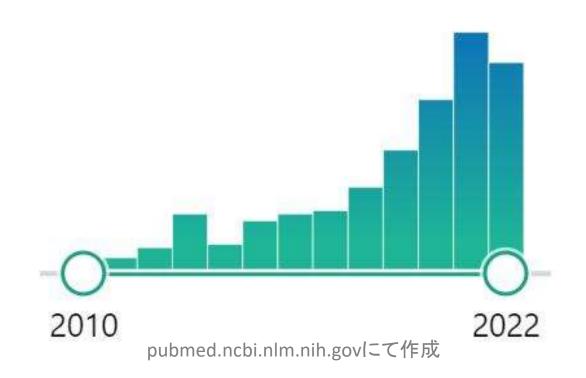


- OHDSI関連論文紹介
- OHDSI シンポジウム(APAC)
- 今月のOHDSI Community Call
- PheValuator 2.0



OHDSI関連論文

Pubmedで"OHDSI or OMOP"を検索



全期間累計:8月280本→9月288本



用語マッピングを半自動で監査する仕組み

> J Biomed Inform. 2022 Aug 25;134:104162. doi: 10.1016/j.jbi.2022.104162. Online ahead of print.

Towards quality improvement of vaccine concept mappings in the OMOP vocabulary with a semi-automated method

```
Rashmie Abeysinghe <sup>1</sup>, Adam Black <sup>2</sup>, Denys Kaduk <sup>2</sup>, Yupeng Li <sup>3</sup>, Christian Reich <sup>4</sup>, Alexander Davydov <sup>2</sup>, Lixia Yao <sup>5</sup>, Licong Cui <sup>6</sup>
```

Affiliations + expand

PMID: 36029954 DOI: 10.1016/j.jbi.2022.104162

The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) provides a unified model to integrate disparate real-world data (RWD) sources. An integral part of the OMOP CDM is the Standardized Vocabularies (henceforth referred to as the OMOP vocabulary), which enables organization and standardization of medical concepts across various clinical domains of the OMOP CDM. For concepts with the same meaning from different source vocabularies, one is designated as the standard concept, while the others are specified as non-standard or source concepts and mapped to the standard one. However, due to the heterogeneity of source vocabularies, there may exist mapping issues such as erroneous mappings and missing mappings in the OMOP vocabulary, which could affect the results of downstream analyses with RWD. In this paper, we focus on quality assurance of vaccine concept mappings in the OMOP vocabulary, which is necessary to accurately harness the power of RWD on vaccines. We introduce a semi-automated lexical approach to audit vaccine mappings in the OMOP vocabulary. We generated two types of vaccine-pairs: mapped and unmapped, where mapped vaccinepairs are pairs of vaccine concepts with a "Maps to" relationship, while unmapped vaccine-pairs are those without a "Maps to" relationship. We represented each vaccine concept name as a set of words, and derived term-difference pairs (i.e., name differences) for mapped and unmapped vaccine-pairs. If the same term-difference pair can be obtained by both mapped and unmapped vaccine-pairs, then this is considered as a potential mapping inconsistency. Applying this approach to the vaccine mappings in OMOP, a total of 2087 potentially mapping inconsistencies were obtained. A randomly selected 200 samples were evaluated by domain experts to identify, validate, and categorize the inconsistencies. Experts identified 95 cases revealing valid mapping issues. The remaining 105 cases were found to be invalid due to the external and/or contextual information used in the mappings that were not reflected in the concept names of vaccines. This indicates that our semi-automated approach shows promise in identifying mapping inconsistencies among vaccine concepts in the OMOP vocabulary.

概要

OMOP CDM は、異なるRWDソースを統合するための統一モデルを提供するものである。OMOP CDM の不可欠な部分は標準化された語彙(以後 OMOP 語彙と呼ぶ)で、これにより OMOP CDM の様々な臨床領域にわたる医療概念の整理と標準化が可能になる。異なるソース語彙から同じ意味を持つ概念に対して、1つを標準概念として指定し、他を非標準(またはソース概念)として標準概念にマッピングする。しかし、ソース語彙の異質性により、OMOP語彙には誤ったマッピングやマッピングの欠落などのマッピングの問題が存在し、RWD による解析結果に影響を与える可能性がある。

本論文では、ワクチンに関する RWD の能力を正確に活用するために必要な、OMOP 語彙におけるワクチン概念のマッピン グの品質保証に焦点をあてる。OMOP語彙のワクチンマッピングを監査するための半自動化された語彙的アプローチを紹介する。

マッピングされたワクチンペアとマッピングされていないワクチンペアの2種類を生成した。マッピングされたワクチンペアは、「Maps to」の関係を持つワクチン概念のペアであり、マッピングされていないワクチンペアは「Maps to」の関係を持たないペアである。ワクチン概念の名称を単語の集合として表現し、マッピングされたワクチンペアとマッピングされていないワクチンペアの用語差分(名称の差分)のペアを導出しました。このとき、マッピングされたワクチンとマッピングされていないワクチンで同じ用語差分ペアが得られる場合は、マッピングの不一致の可能性があると判断する。このアプローチを OMOP のワクチンマッピングに適用したところ、合計 2087 件のマッピングの不一致の可能性が得られた。

無作為に選んだ 200 サンプルを専門家が評価し、不整合の特定、検証、および分類を行った。専門家の評価では、95 件が有効なマッピングの問題であることが判明した。残りの105件は、マッピングに使用された外部情報および/または文脈情報が、ワクチンの概念名称に反映されていないため、無効であることが判明した。

このことから、我々の半自動化されたアプローチは、OMOP 語彙におけるワクチン概念間のマッピングの不一致を特定する上で有望であることが示された。

OMOP CDM利用時に 匿名性を増すための仕組み



> Stud Health Technol Inform. 2022 Aug 31;298:163-164. doi: 10.3233/SHTI220930.

Preserving Privacy when Querying OMOP CDM Databases

Joao Rafael Almeida 1 2, Joao Paulo Barraca 3, José Luís Oliveira 1

Affiliations - collapse

Affiliations

- 1 DETI/IEETA, University of Aveiro, Portugal.
- 2 Department of Computation, University of A Coruña, Spain.
- 3 IT, DETI, University of Aveiro, Portugal.

PMID: 36073478 DOI: 10.3233/SHTI220930

Anonymisation is currently one of the biggest challenges when sharing sensitive personal information. Its importance depends largely on the application domain, but when dealing with health information, this becomes a more serious issue. A simpler approach to avoid inadequate disclosure is to ensure that all data that can be associated directly with an individual is removed from the original dataset. However, some studies have shown that simple anonymisation procedures can sometimes be reverted using specific patients' characteristics. In this work, we propose a secure architecture to share information from distributed databases without compromising the subjects' privacy. The anonymiser system was validated using the OMOP CDM data schema, which is widely adopted in observational research studies.

概要

匿名化は、現在、機密性の高い個人情報を共有する際の最大の課題の1つである。その重要性はアプリケーションの領域によって大きく異なるが、健康情報を扱う場合、より重大な事項となる。不適切な情報開示を避けるためのより単純なアプローチは、個人と直接関連付けることができるすべてのデータを元のデータセットから削除することを保証することである。しかし、いくつかの研究により、単純な匿名化手続きは、時として特定の患者の特徴を用いて元に戻すことができることが示されている。

本研究では、被験者のプライバシーを侵害することなく、分散型データベースから情報を 共有するための安全なアーキテクチャを提案する。匿名化システムは、観察研究において広 く採用されているOMOP CDMデータスキーマを用いて検証された。

- 今のところ k-匿名性、L-多様性 を使ったもの。
- 仕組みの詳細は不明。Query時に? viewで匿名化する?
- 統計処理結果ではなくデータが共有できることを目指す。 (抽出結果を返せるように?)



OMOP-on-FHIRを使って OMOP CDMをFHIR対応した

> Annu Int Conf IEEE Eng Med Biol Soc. 2022 Jul;2022:4687-4690. doi: 10.1109/EMBC48229.2022.9871865.

Accelerating Multi-site Health Informatics with Streamlined Data Infrastructure using OMOP-on-FHIR

Benoit L Marteau, Yuanda Zhu, Felipe Giuste, Wenqi Shi, Ashley Carpenter, Coleman Hilton, May D Wang

PMID: 36085809 DOI: 10.1109/EMBC48229.2022.9871865

Shriners Children's (SHC) is a hospital system whose mission is to advance the treatment and research of pediatric diseases. SHC success has generated a wealth of clinical data. Unfortunately, barriers to healthcare data access often limit data-driven clinical research. We decreased this burden by allowing access to clinical data via the standardized data access standard called FHIR (Fast Healthcare Interoperability Resources). Specifically, we converted existing data in the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) standard into FHIR data elements using a technology called OMOP-on-FHIR. In addition, we developed two applications leveraging the FHIR data elements to facilitate patient cohort curation to advance research into pediatric musculoskeletal diseases. Our work enables clinicians and clinical researchers to use hundreds of currently available open-sourced FHIR applications. Our successful implementation of OMOP-on-FHIR within a large hospital system will accelerate advancements in pediatric disease treatment and research.

概要

Shriners Children's (SHC)は、小児疾患の治療と研究を進めることを使命とする病院システムである。SHCの成功は、豊富な臨床データを生み出した。残念ながら、医療データへのアクセスには障壁があり、データ主導の臨床研究が制限されることがよくある。そこで、FHIRという標準的なデータアクセス規格を利用して、臨床データにアクセスできるようにすることで、その負担を軽減した。

具体的には、OMOP CDM 規格の既存データを、OMOP-on-FHIRという技術でFHIRのデータ要素に変換した。さらに、小児筋骨格系疾患の研究を促進するために、患者コホートのキュレーションを促進するFHIRデータ要素を活用した2つのアプリケーションを開発した。私たちの研究成果により、臨床医や臨床研究者は、現在オープンソースで提供されている数百のFHIRアプリケーションを利用できるようになった。大規模な病院システム内でのOMOP-on-FHIR の実装に成功したことで、小児疾患の治療と研究の進歩が加速されるだろう。

※シュライナーズ チルドレンズは、北米の非営利医療施設のネットワーク。20数施設

OMOP DBをFHIR RDFとして扱える 仮想ナレッジグラフ(VKG)システム



> J Biomed Inform. 2022 Sep 9;134:104201. doi: 10.1016/j.jbi.2022.104201. Online ahead of print.

FHIR-Ontop-OMOP: Building clinical knowledge graphs in FHIR RDF with the OMOP Common data Model

Guohui Xiao ¹, Emily Pfaff ², Eric Prud'hommeaux ³, David Booth ⁴, Deepak K Sharma ⁵, Nan Huo ⁵, Yue Yu ⁵, Nansu Zong ⁵, Kathryn J Ruddy ⁵, Christopher G Chute ⁶, Guoqian Jiang ⁷

Affiliations - collapse

Affiliations

- 1 University of Bergen, Norway; University of Oslo, Norway; Ontopic S.r.l., Italy. Electronic address: guohui.xiao@uib.no.
- 2 University of North Carolina, Chapel Hill, NC, USA.
- 3 Janeiro Digital, Boston, MA, USA.
- 4 Yosemite Project, Somerville, MA, USA.
- 5 Mayo Clinic, Rochester, MN, USA.
- 6 Johns Hopkins University, Baltimore, MD, USA.
- 7 Mayo Clinic, Rochester, MN, USA. Electronic address: jiang.guoqian@mayo.edu.

Background: Knowledge graphs (KGs) play a key role to enable explainable artificial intelligence (AI) applications in healthcare. Constructing clinical knowledge graphs (CKGs) against heterogeneous electronic health records (EHRs) has been desired by the research and healthcare AI communities. From the standardization perspective, community-based standards such as the Fast Healthcare Interoperability Resources (FHIR) and the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) are increasingly used to represent and standardize EHR data for clinical data analytics, however, the potential of such a standard on building CKG has not been well investigated.

Objective: To develop and evaluate methods and tools that expose the OMOP CDM-based clinical data repositories into virtual clinical KGs that are compliant with FHIR Resource Description Framework (RDF) specification. **Methods:** We developed a system called FHIR-Ontop-OMOP to generate virtual clinical KGs from the OMOP relational databases. We leveraged an OMOP CDM-based Medical Information Mart for Intensive Care (MIMIC-III) data repository to evaluate the FHIR-Ontop-OMOP system in terms of the faithfulness of data transformation and the conformance of the generated CKGs to the FHIR RDF specification.

Results: A beta version of the system has been released. A total of more than 100 data element mappings from 11 OMOP CDM clinical data, health system and vocabulary tables were implemented in the system, covering 11 FHIR resources. The generated virtual CKG from MIMIC-III contains 46,520 instances of FHIR Patient, 716,595 instances of Condition, 1,063,525 instances of Procedure, 24,934,751 instances of MedicationStatement, 365,181,104 instances of Observations, and 4,779,672 instances of CodeableConcept. Patient counts identified by five pairs of SQL (over the MIMIC database) and SPARQL (over the virtual CKG) queries were identical, ensuring the faithfulness of the data transformation. Generated CKG in RDF triples for 100 patients were fully conformant with the FHIR RDF specification. **Conclusion:** The FHIR-Ontop-OMOP system can expose OMOP database as a FHIR-compliant RDF graph. It provides a meaningful use case demonstrating the potentials that can be enabled by the interoperability between FHIR and OMOP CDM. Generated clinical KGs in FHIR RDF provide a semantic foundation to enable explainable AI applications in healthcare.

概要

背景: 知識グラフ(KG)は、ヘルスケアにおける説明可能な人工知能(AI)アプリケーションを実現するために重要な役割を担っている。異種電子医療記録(EHR)に対して臨床知識グラフ(CKG)を構築することは、研究コミュニティやヘルスケアAIコミュニティで望まれてきたことである。標準化の観点からは、FHIR や OMOP CDM などのコミュニティベースの標準が、臨床データ分析のためのEHRデータの表現と標準化にますます使用されているが、CKG構築におけるこうした標準の可能性は十分に検討されていない。

目的: OMOP CDMに基づく臨床データリポジトリをFHIR Resource Description Framework (RDF) 仕様に準拠した仮想臨床KGに公開するための方法とツールを開発し、評価する。

方法: FHIR-Ontop-OMOP と呼ばれるシステムを開発し、OMOP リレーショナルデータベースから仮想臨床 KG を生成した。OMOP CDMベースのMedical Information Mart for Intensive Care(MIMIC-III) データリポジトリを活用し、データ変換の忠実性と生成されたCKGのFHIR RDF仕様への適合性の観点からFHIR-Ontop-OMOPシステムを評価した。

結果: システムのベータ版がリリースされた。OMOP CDMの臨床データ、医療システム、語彙の11テーブルから、合計100以上のデータ要素マッピングがシステムに実装され、11のFHIRリソースをカバーした。MIMIC-IIIから生成された仮想CKGは、FHIR Patientの46,520インスタンス、Conditionの716,595インスタンス、Procedureの1,063,525インスタンス、MedicationStatementの24,934,751インスタンス、Observationsの365,181,104インスタンスおよび CodeableConcept の 4,779, 672インスタンスで構成されている。5組のSQL(MIMICデータベース上)とSPARQL(仮想CKG上)のクエリによって識別された患者数は同一であり、データ変換の忠実性を保証している。100人の患者について生成されたRDFトリプルのCKGは、FHIR RDF仕様に完全に適合していた。

結論: FHIR-Ontop-OMOPシステムは、OMOPデータベースをFHIRに準拠したRDFグラフとして公開することができる。FHIR と OMOP CDM の相互運用性によって実現される可能性を示す有意義なユースケースを提供するものである。FHIR RDFで生成された臨床KGは、医療における説明可能なAIアプリケーションを可能にするセマンティックな基盤を提供する。



薬剤耐性てんかんのPhenotyping開発

> Epilepsia. 2022 Sep 14. doi: 10.1111/epi.17409. Online ahead of print.

Identification of patients with drug resistant epilepsy in electronic medical record data using the Observational Medical Outcomes Partnership Common Data Model

```
Victor G Castano <sup>1</sup>, Matthew Spotnitz <sup>2</sup>, Genna J Waldman <sup>3</sup>, Evan F Joiner <sup>1</sup>, Hyunmi Choi <sup>3</sup>, Anna Ostropolets <sup>2</sup>, Karthik Natarajan <sup>2</sup>, Guy M McKhann <sup>1</sup>, Ruth Ottman <sup>3</sup> <sup>4</sup> <sup>5</sup> <sup>6</sup>, Alfred I Neugut <sup>5</sup> <sup>7</sup> <sup>8</sup>, George Hripcsak <sup>2</sup>, Brett E Youngerman <sup>1</sup>

Affiliations + expand Columbia University, New York, New York, USA.
```

PMID: 36106377 DOI: 10.1111/epi.17409

Objective: Over one-third of appropriately treated patients with epilepsy have continued seizures despite two or more medication trials, meeting criteria for drug resistant epilepsy (DRE). Accurate and reliable identification of patients with DRE in observational data would enable large-scale, real-world comparative effectiveness research and improve access to specialized epilepsy care. In the present study, we aim to develop and compare the performance of computable phenotypes for DRE using the Observational Medical Outcomes Partnership (OMOP) Common Data Model.

Methods: We randomly sampled 600 patients from our academic medical center's electronic health record (EHR)-derived OMOP database meeting previously validated criteria for epilepsy (1/2015-8/2021). Two reviewers manually classified patients as having DRE, drug responsive epilepsy, undefined drug responsiveness, or not epilepsy as of the last EHR encounter in the study period based on consensus definitions. Demographic characteristics and codes for diagnoses, antiseizure medications (ASM), and procedures were tested for association with DRE. Algorithms combining permutations of these factors were applied to calculate sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for DRE. The F1-score was used to compare overall performance.

Results: Among 412 patients with source-record confirmed epilepsy, 62 (15.0%) had DRE, 163 (39.6%) drug responsive epilepsy, 124 (30.0%) undefined drug responsiveness, and 63 (15.3%) insufficient records. The best performing phenotype for DRE in terms of the F1-score was the presence of ≥1 intractable epilepsy code and ≥2 unique non-gabapentinoid ASM exposures each with ≥90-day drug era (sensitivity 0.661, specificity 0.937, PPV 0.594, NPV 0.952, F1-score 0.626). Several phenotypes achieved higher sensitivity at the expense of specificity and vice versa.

Significance: OMOP algorithms can identify DRE in EHR-derived data with varying tradeoffs between sensitivity and specificity. These computable phenotypes can be applied across the largest international network of standardized clinical databases for further validation, reproducible observational research, and improving access to appropriate care.

概要

目的: 適切な治療を受けたてんかん患者の3分の1以上が、2種類以上の薬物治療を受けたにもかかわらず発作が持続し、薬剤耐性てんかん(DRE)の基準を満たすと言われている。観察データから正確かつ信頼性の高いDRE患者を特定できれば、大規模で実環境に近い比較効果研究が可能になり、てんかん専門医療へのアクセスも改善される。本研究では、OMOP共通データモデルを用いて、DREの計算可能なPhenotypingを開発し、その性能を比較することを目的としている。

方法: 私たちの学術医療センターの電子カルテ(EHR)由来のOMOPデータベースから、事前に検証されたてんかんの基準を満たす600人の患者を無作為に抽出した(2015年1月~2021年8月)。2名のレビュアーが、コンセンサス定義に基づき、調査期間中の最後のEHRエンカウンターの時点で、患者をDRE、薬剤反応性てんかん、薬剤反応性未確定、てんかんではないに手入力で分類した。人口統計学的特徴、診断名、抗てんかん薬(ASM)、処置のコードについて、DREとの関連性を検証した。これらの因子の並べ換えを組み合わせたアルゴリズムを適用し、DREの感度、特異度、陽性適中率(PPV)、陰性適中率(NPV)を算出した。F1スコアは総合的な性能の比較に使用された。

結果: ソースレコードで確認されたてんかん患者412名のうち、62名(15.0%)がDRE、163名(39.6%)が薬剤反応性てんかん、124名(30.0%)が薬剤反応性未確定、63名(15.3%)が記録不十分であった。F1スコアの観点から、DREの最も優れたPhenotypingは、≥1つの難治性てんかんコードと≥2つの非ガバペンチノイドASM曝露がそれぞれ≥90日の薬剤期間を持つことであった(感度0.661、特異度0.937、PPV0.594、NPV0.952、F1スコア0.626)。いくつかのPhenotypingは、特異度を犠牲にして高い感度を達成し、逆もまた然りであった。

意義: OMOPアルゴリズムは、感度と特異度のトレードオフを変えながら、EHR由来のデータからDREを同定することができる。これらの計算可能なPhenotypingは、標準化された臨床データベースの最大の国際的ネットワークに適用することで、さらなる検証、再現性のある観察研究、適切な治療へのアクセスの改善を図ることができる。



実臨床でのiGlarLixiの効果

> Diabetes Obes Metab. 2022 Aug 19. doi: 10.1111/dom.14844. Online ahead of print.

iGlarLixi versus basal plus Rapid-Acting insulin in adults with type 2 diabetes advancing from basal insulin therapy: The SoliSimplify Real-World study

Rory J McCrimmon ¹, Alice Y Y Cheng ², Gagik Galstyan ³, Khier Djaballah ⁴, Xuan Li ⁵, Mathieu Coudert ⁴, Juan P Frias ⁶

Affiliations - collapse

Affiliations

- 1 Division of Systems Medicine, School of Medicine, University of Dundee, Dundee, UK.
- 2 Department of Medicine, University of Toronto, Toronto, Ontario, Canada.
- 3 Diabetic Foot Department, Endocrinology Research Center, Moscow, Russia.
- 4 Sanofi, Paris, France.
- 5 Sanofi, Bridgewater, New Jersey, USA.
- 6 Velocity Clinical Research, Los Angeles, California, USA.

PMID: 36123617 DOI: 10.1111/dom.14844

Aim: For people with suboptimally controlled type 2 diabetes (T2D) on basal insulin (BI), guidelines recommend several treatment advancement options. This study compared the clinical effectiveness of once-daily iGlarLixi versus a multiple-injection BI + rapid acting insulin (RAI) regimen in adults with T2D advancing from BI therapy in real-world clinical practice.

Materials and methods: Electronic medical records from the Observational Medical Outcomes Partnership (OMOP) database were analysed retrospectively using propensity score matching to compare therapy advancement with iGlarLixi or BI + RAI in US adults ≥18 years with T2D on BI who had ≥1 valid glycated haemoglobin (HbA1c) value at baseline and at the 6-month follow-up. The primary objective was non-inferiority of iGlarLixi to BI + RAI in HbA1c change from baseline to 6 months (margin 0.3%).

Results: Propensity score matching generated cohorts with balanced baseline characteristics (N = 814 in each group). HbA1c reduction from baseline to 6 months with iGlarLixi was non-inferior to BI + RAI [mean difference (95% confidence interval): 0.1 (-0.1, 0.2)%; one-sided p = .0032]. At 6 months, weight gain was significantly lower with iGlarLixi than with BI + RAI [-0.8 (-1.3, -0.2) kg; two-sided p = .0069]. Achievement of HbA1c <7% without hypoglycaemia and weight gain were similar between groups [odds ratio (95% confidence interval): 1.15 (0.81, 1.63); p = .4280]. Hypoglycaemia was low in both groups, probably because of underreporting.

Conclusions: In real-world clinical practice, glycaemic outcomes 6 months after treatment advancement from BI are similar for people with T2D using iGlarLixi versus BI + RAI, with iGlarLixi leading to less weight gain.

概要

目的: 基礎インスリン(BI)で最適にコントロールされていない2型糖尿病(T2D)患者に対して、ガイドラインではいくつかの治療法から進めることを推奨している。本研究では、実臨床においてBI療法から移行した成人のT2D患者を対象に、1日1回投与のiGlarLixiと BI+速効型インスリン(RAI)多剤投与レジメンの臨床効果を比較した。

材料と方法: OMOPデータベースの電子カルテを傾向スコアマッチングを用いてレトロスペクティブに解析し、ベースラインおよび6ヶ月のフォローアップ時に有効なHbA1c値が1つ以上あったBI療法中の18歳以上の米国成人における、iGlarLixi または BI + RAI による治療法の進展を比較検討した。主要目的は、ベースラインから6ヵ月後までのHbA1cの変化において、iGlarLixi の BI+RAI に対する非劣性(マージン0.3%)であった。

結果:傾向スコアマッチングにより、ベースライン特性が均衡したコホートが作成された(各群N=814)。

- ・iGlarLixiのベースラインから6ヶ月までのHbA1cの減少は、BI + RAIに対して非劣性であった [平均差(95%信頼区間): 0.1 (-0.1, 0.2)%; 片側 p = 0.0032]。
- ・6ヵ月後の体重増加は、iGlarLixiがBI + RAIよりも有意に低かった [-0.8 (-1.3, -0.2) kg; 両側p = 0.0069]。
- ・低血糖を伴わないHbA1c<7%達成と体重増加は、群間で同程度であった [オッズ比(95%信頼区間):1.15(0.81、1.63); p=.4280]。
- ・低血糖は両群とも少なく, おそらく報告不足のためと思われる。

結論: 実臨床において、BIからの治療開始6ヶ月後の血糖値は、iGlarLixiとBI+RAIを使用したT2D患者において同等であり、iGlarLixiは体重増加をより少なくすることが示された。

iGlarLixi: 基礎インスリンであるインスリン グラルギンとGLP-1受容体作動薬であるリキシセナチドを1単位: 1μg配合(インスリン グラルギン100単位/mL・リキ シセナチド100μg/mL配合注射剤)として開発された1日1回皮下投与製剤。

BI + RAI: 基礎インスリン1日1回+即効性インスリン1日3回



OMOPを出力する、高性能なオープンソースの 放射線診断レポート テキスト解析システム

> IEEE Int Conf Healthc Inform. 2022 Jun;2022:288-296. doi: 10.1109/ichi54592.2022.00050. Epub 2022 Sep 8.

Radiology Text Analysis System (RadText): Architecture and Evaluation

Song Wang ¹, Mingquan Lin ², Ying Ding ³, George Shih ⁴, Zhiyong Lu ⁵, Yifan Peng ²
Affiliations — collapse

Affiliations

- 1 Cockrell School of Engineering, The University of Texas at Austin, Austin, USA.
- 2 Population Health Sciences, Weill Cornell Medicine, New York, USA.
- 3 School of Information, The University of Texas at Austin, Austin, USA.
- 4 Department of Radiology, Weill Cornell Medicine, New York, USA.
- 5 National Center for Biotechnology Information (NCBI), National Library of Medicine (NLM), National Institutes of Health (NIH), Bethesda, USA.

PMID: 36128510 PMCID: PMC9484781 DOI: 10.1109/ichi54592.2022.00050

Analyzing radiology reports is a time-consuming and error-prone task, which raises the need for an efficient automated radiology report analysis system to alleviate the workloads of radiologists and encourage precise diagnosis. In this work, we present RadText, a high-performance open-source Python radiology text analysis system. RadText offers an easy-to-use text analysis pipeline, including deidentification, section segmentation, sentence split and word tokenization, named entity recognition, parsing, and negation detection. Superior to existing widely used toolkits, RadText features a hybrid text processing schema, supports raw text processing and local processing, which enables higher accuracy, better usability and improved data privacy. RadText adopts BioC as the unified interface, and also standardizes the output into a structured representation that is compatible with Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM), which allows for a more systematic approach to observational research across multiple, disparate data sources. We evaluated RadText on the MIMIC-CXR dataset, with five new disease labels that we annotated for this work. RadText demonstrates highly accurate classification performances, with a 0.91 average precision, 0.94 average recall and 0.92 average F-1 score. We also annotated a test set for the five new disease labels to facilitate future research or applications. We have made our code, documentations, examples and the test set available at https://github.com/bionlplab/radtext.

概要

放射線科医の負担を軽減し、正確な診断を行うために、効率的な放射線レポート自動解析システムの必要性が高まっている。本稿では、高性能なオープンソースのPython製放射線診断テキスト解析システムRadTextを紹介する。

RadTextは、非識別化(匿名化)、セクション分割、文分割、単語トークン化、固有表現認識、構文解析、否定文検出など、使いやすいテキスト解析パイプラインを提供する。既存のツールキットと比較して、RadTextはハイブリッドなテキスト処理スキーマを備えており、生テキスト処理とローカル処理をサポートすることで、より高い精度、より良いユーザビリティ、データのプライバシー保護を実現する。RadTextは統一インターフェースとしてBioCを採用し、さらにOMOP CDM準拠の構造化表現に出力を標準化し、複数の異種データソースにわたる観察研究へのより体系的なアプローチを可能にするものだ。我々は、MIMIC-CXRデータセットに、この研究のために新たにアノテーションした5つの疾患ラベルを用いてRadTextを評価した。RadTextは、平均適合率0.91、平均リコール0.94、平均F-1スコア0.92という高い精度の分類性能を実証している。また、今後の研究や応用を促進するために、5つの新しい疾患ラベルのテストセットもアノテーションしている。コード、ドキュメント、例題、テストセットは、https://github.com/bionlplab/radtext で公開されている。



心電図自動診断をOMOPコンセプトに変換する

> Appl Clin Inform. 2022 Aug;13(4):880-890. doi: 10.1055/s-0042-1756427. Epub 2022 Sep 21.

Conversion of Automated 12-Lead Electrocardiogram Interpretations to OMOP CDM Vocabulary

Sunho Choi 1, Hyung Joon Joo 2 3, Yoojoong Kim 4, Jong-Ho Kim 2 3, Junhee Seok 1

Affiliations – collapse

Affiliations

- 1 School of Electrical Engineering, Korea University, Seoul, South Korea.
- 2 Korea University Research Institute for Medical Bigdata Science, Korea University, Seoul, South Korea.
- 3 Department of Cardiology, Cardiovascular Center, Korea University College of Medicine, Seoul, South Korea.
- 4 School of Computer Science and Information Engineering, The Catholic University of Korea, Seoul, South Korea.

PMID: 36130711 PMCID: PMC9492322 DOI: 10.1055/s-0042-1756427

Background: A computerized 12-lead electrocardiogram (ECG) can automatically generate diagnostic statements, which are helpful for clinical purposes. Standardization is required for big data analysis when using ECG data generated by different interpretation algorithms. The common data model (CDM) is a standard schema designed to overcome heterogeneity between medical data. Diagnostic statements usually contain multiple CDM concepts and also include non-essential noise information, which should be removed during CDM conversion. Existing CDM conversion tools have several limitations, such as the requirement for manual validation, inability to extract multiple CDM concepts, and inadequate noise removal.

Objectives: We aim to develop a fully automated text data conversion algorithm that overcomes limitations of existing tools and manual conversion.

Methods: We used interpretations printed by 12-lead resting ECG tests from three different vendors: GE Medical Systems, Philips Medical Systems, and Nihon Kohden. For automatic mapping, we first constructed an ontology-lexicon of ECG interpretations. After clinical coding, an optimized tool for converting ECG interpretation to CDM terminology is developed using term-based text processing. **Results:** Using the ontology-lexicon, the cosine similarity-based algorithm and rule-based hierarchical algorithm showed comparable conversion accuracy (97.8 and 99.6%, respectively), while an integrated algorithm based on a heuristic approach, ECG2CDM, demonstrated superior performance (99.9%) for datasets from three major vendors.

Conclusion: We developed a user-friendly software that runs the ECG2CDM algorithm that is easy to use even if the user is not familiar with CDM or medical terminology. We propose that automated algorithms can be helpful for further big data analysis with an integrated and standardized ECG dataset.

概要

背景: コンピュータによる12誘導心電図(ECG)は診断文を自動生成することができ、臨床に有用である。 異なる解釈アルゴリズムによって生成されたECGデータを使用する場合、ビッグデータ解析には標準化が 必要である。共通データモデル(CDM)は、医療データ間の異質性を克服するために設計された標準ス キーマである。診断文には通常、複数のCDMコンセプトが含まれ、また本質的でないノイズ情報も含まれ るため、CDM変換の際に除去する必要がある。既存のCDM変換ツールは、手作業による検証、複数の CDMコンセプトの抽出、ノイズ除去の不十分さなど、いくつかの制限がある。

目的: 既存の変換ツールや手動変換の限界を克服した、完全自動のテキストデータ変換アルゴリズムを開発することを目的とする。

方法: 3つのベンダーの12誘導安静時心電図検査で出力された解釈を使用した。GEメディカルシステムズ、フィリップスメディカルシステムズ、日本光電工業の3社から提供された12誘導安静時心電図を使用した。自動マッピングのために、まず心電図解釈のオントロジーレキシコンを構築した。その後、用語ベースのテキスト処理により、心電図解釈をCDM用語に変換するための最適化ツールを開発した。

結果: オントロジーレキシコンを用いて、コサイン類似度アルゴリズムとルールベース階層型アルゴリズムは同程度の変換精度(それぞれ97.8%と99.6%)を示したが、ヒューリスティックアプローチに基づく統合アルゴリズムECG2CDMは主要3社のデータセットで優れた性能(99.9%)を実証した。

結論: CDMや医学用語に詳しくないユーザでも使いやすいECG2CDMアルゴリズムを実行するソフトウェアを開発した。統合され標準化されたECGデータセットを用いたさらなるビッグデータ解析に、自動化されたアルゴリズムが役立つことを提案する。



OHDSI シンポジウム

■ OHDSI Global Symposium

・10月14-16日

Oct. 14 – Main Conference

Oct. 15 – Tutorial

Oct. 15-16 – Workgroup Activities

■ OHDSI APAC Symposium

・11月12-13日に台北で。

12日 プレ日 チュートリアルday

13日 本開催 9:00-17:00

・会場

臺北醫學大學 綜合醫學大樓 4F

General Medicine Building 4F, Taipei Medical University https://goo.gl/maps/hBzhfiQmp5TYPPwe9

OHDSI APAC Symposium

ここから参加登録できます。Virtual & In-person.

https://ohdsi.org/2022-apac-ohdsi-symposium/

Publications

Support & Sponsorship

✓ 2022 OHDSI Symposium

✓ 2022 APAC Symposium

Newsletters ∨ Follow OHDSI on Social ∨

2022 APAC OHDSI Symposium

Nov. 12 - 13 • Taipei Medical University



We are excited to announce that registration and collaborator showcase submissions for the 2022 OHDSI APAC Symposium, which will be held Nov. 12-13, is now open! This two-day event will take place in Taipei, Taiwan at the Taipei Medical University and also be broadcasted virtually for those who are not able to participate in person.

OHDSI APAC Symposium 現在の予定

11月12日 土曜日

08:30-09:00 Registration

09:00-12:00 OHDSI Intro - CDM & Vocab

12:00-13:00 Lunch & Poster session

13:00-17:00 ETL & DQ / Phenotype Development

11月13日 日曜日

08:00-09:00 Registration & Light breakfast

09:00-09:20 Welcome session

09:20-09:40 Group Photo

Session 1: Envisioning of OHDSI Global & EU

09:40-10:00 Key note - OHDSI Global Presentation

10:00-10:20 DARWIN EU

10:20-10:30 Break

Session 2: The Challenges of Research in OHDSI APAC

10:30-10:50 OHDSI APAC Introduction

10:50-11:30 Researches in OHDSI APAC

11:30-11:45 Researches using Taiwan national data

11:45-12:00 Researches using TMUCRD data

12:00-13:00 Lunch & Poster Presentation

Session 3: The Implication Experiences in OHDSI Region

13:00-14:00 Panel - Standardization & common data models

14:00-15:00 Panel - APAC Regional adaption to standardization

15:00-15:15 Break

15:15-16:15 Poster session / Network Session

16:15-17:00 Closing remarks



今月のCommunity Call

- APAC Community Call テーマ
- Aug. 25 Workgroup Updates
 (Natural Language Processing,
 Phenotype Development & Evaluation, FHIR & OMOP)
- Sept. 8 EU Chapter Sharing Session (Greece, Israel)
- Global Community Call テーマ
- Aug. 30 EHDEN Portal and EHDEN Academy
- Sept. 6 OHDSI Study Publications
- Sept. 13 Clinical Registry Efforts in OHDSI
- Sept. 20 2022 Symposium Preview
- Sept. 27 2022 HTA Challenge



PheValuator 2.0

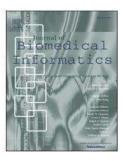
PheValuatorがバージョンアップした



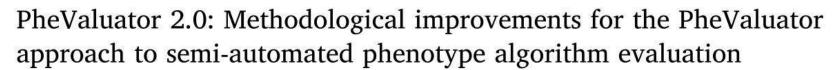
Contents lists available at ScienceDirect

Journal of Biomedical Informatics





Original Research





^a Janssen Research and Development, Titusville, NJ, USA



^b Columbia University, New York, NY, USA

^c Observational Health Data Sciences and Informatics (OHDSI), New York, NY

Phenotypingとは

《背景》

- ある疾病の患者群を「病名」で抽出しても臨床的には正しくない。(特にレセプト:請求のための患者状態であって臨床的状態とはやや異なる)
- 病名として記載されない患者状態がほしいときもある。



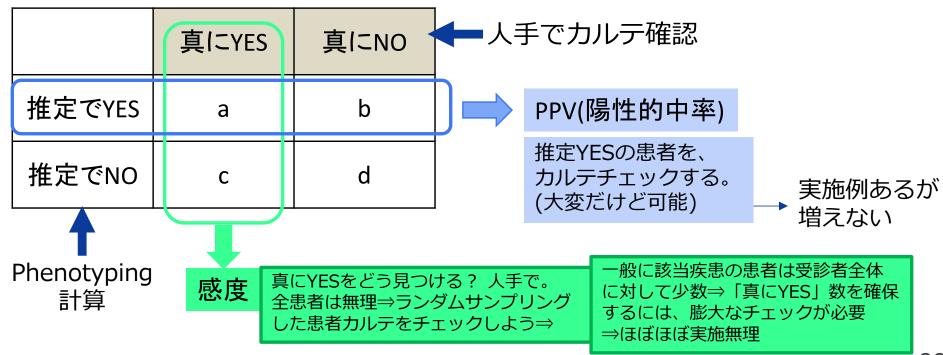
DB中の複数の項目から、欲しい患者状態を推定する

- 臨床的な表現型(Phenotype)を推定する。
- 人間が判定するのではなく、計算機で大量に、高速に。
- EHR-Phenotyping、e-Phenotypingとも。

大きく3つの手法: 条件式、ルールベース、機械学習

Phenotypingの妥当性評価

- 「推定」なので、どれくらい真実に近いか知りたいよね。
- ◆ 分析の信頼性確保のためには必須だよね。
 - 一般的な評価の方法:専門家による人手カルテチェック



Phenotypingと妥当性評価例(validation)

●甲状腺機能亢進症 のレセプトからのPPV

病名のみ:27.1%

治療薬のみ:65.5%

病名&治療薬:69.4%

病名&診療行為&治療薬:72.4%

●急性腎不全

病名「急性腎不全」があり、

かつ、血清Cre値の<u>急激な上昇</u>※がある。

※前1ヶ月〜後1ヶ月の血清Cre値が、過去3ヶ月以内の血清Creの値(**前値**)より 0.3mg/dL以上高値あるいは150%以上の値への変化。

SS-MIX2データからのPPVは

前値≦1.2mg/dL に限定したとき 89.5%

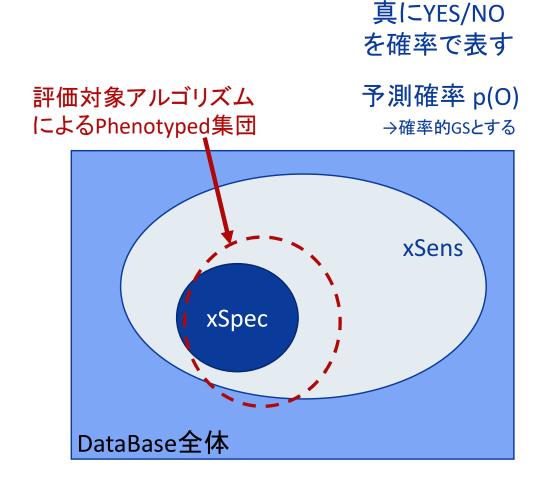
前値≦2.0mg/dL に限定したとき 80.5%

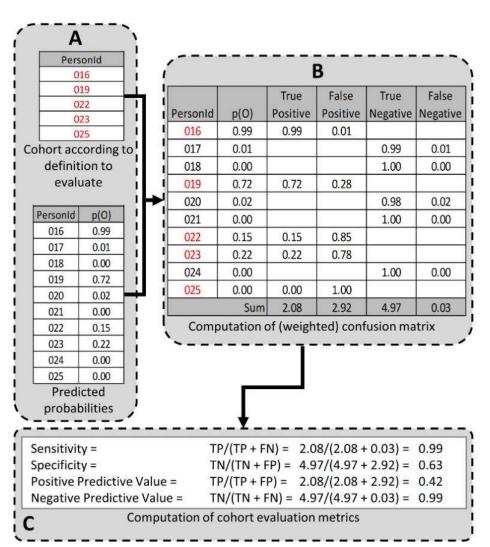
前値の制限がないとき 53.7%

※血清Cre正常値はおおよそ、男性1.2mg/dL以下、女性1.0mg/dL以下。

PheValuatorとは

カルテチェックをしないPhenotyping妥当性評価方法





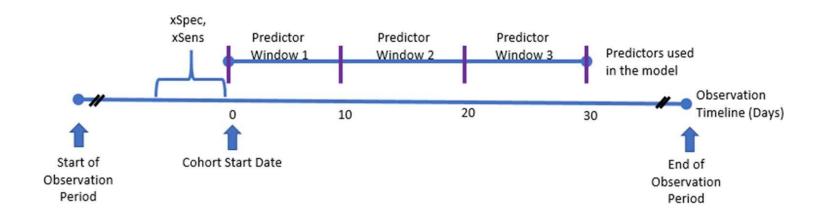
2.0で何が変わったか

- ■1.0の限界
- カルテチェックの結果と比較して、PPVの系統的な過小評価がありそう。
- 確率的GSを計算するモデルの説明変数に制限がある。 xSens/xSpecに使った事象を説明変数に使用することができないが、 その事象は推定にとても重要な変数のはずである。

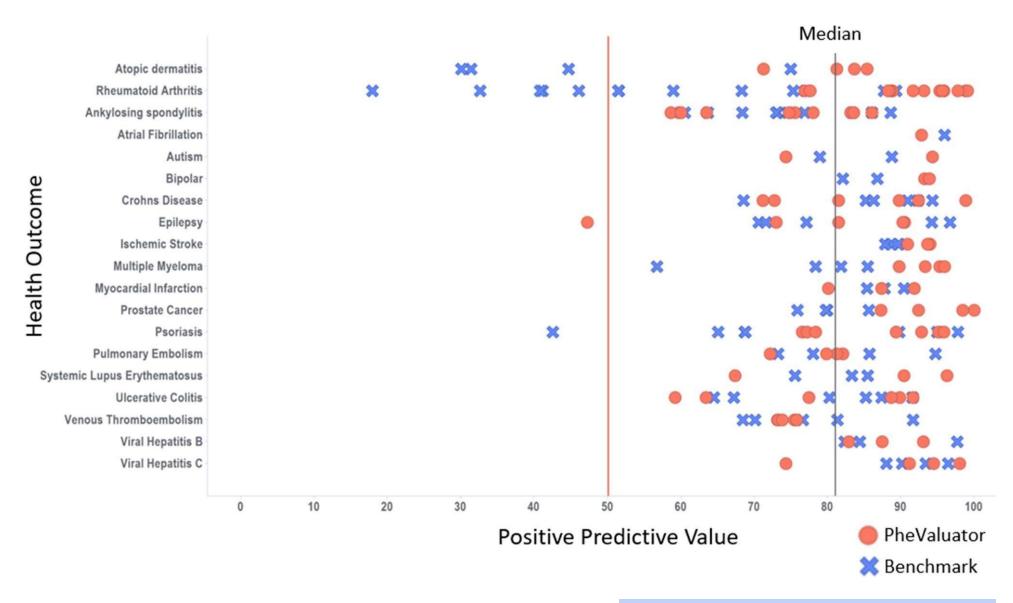
■2.0では

時間窓(Time Window)を導入

- モデル作成時とは異なる時間枠で評価することで、変数制限がなくなり、 推定に重要な変数を使用できる。
- 観察期間全体を使用しないので、急性や反復性のイベントを取り扱える。
- 全体のイベント回数も増える。



結果として精度が良くなった



幅広いPhenotypingで検証され、 PheValuatorの一般化可能性に信憑性を与える。 適切なPhenotyping アルゴリズムの選択のために。