

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 #50

イブニングカンファレンス(第50回)

2024.1.30



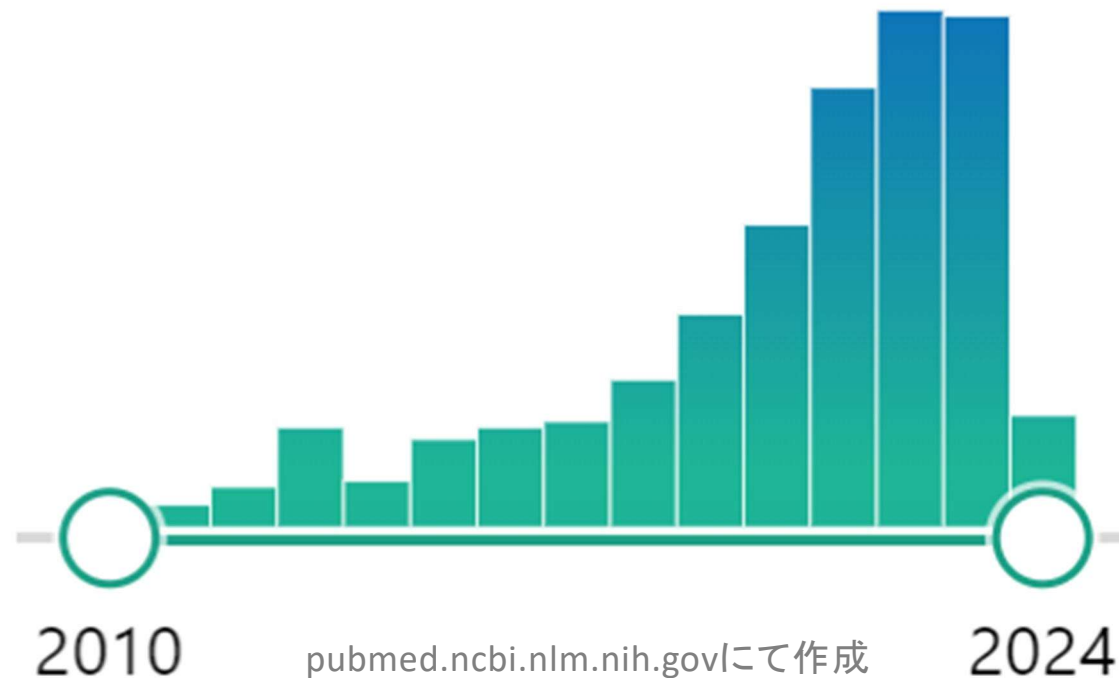
本日の内容

- OHDSI関連論文紹介
- OHDSI Global の話題から
- OHDSI関連論文紹介 part2



OHDSI関連論文

Pubmedで“OHDSI or OMOP”を検索



全期間累計：12月386本→1月399本

- 検索に漏れているものがあるため、実際は累計500本を超えている。
- 年間では100本ペース。



CPRD AURUMデータのOMOP化

> [Inform Med Unlocked](#). 2023;43:101407. doi: 10.1016/j.imu.2023.101407. Epub 2023 Nov 10.

Conversion of CPRD AURUM Data into the OMOP Common Data Model

Craig S Mayer¹

Affiliations – collapse

Affiliation

¹ Lister Hill National Center for Biomedical Communication, National Library of Medicine, NIH
Bethesda, MD, USA.

PMID: 38046363 PMCID: PMC10688258 (available on 2024-11-10)

Abstract

Introduction: Efforts to standardize clinical data using Common Data Models (CDMS) has grown in recent years. Use of CDMs allows for quicker understanding of data structure and reuse of existing tools. One CDM is the Observational Medical Outcomes Partnership (OMOP) CDM. Clinical Practice Research Datalink (CPRD) is a data collection program collecting general practitioner data in the UK.

Objective: Our objective was to convert a static copy of CPRD AURUM data into the OMOP CDM and run existing tools on the converted data.

Methods: Two methods were used to convert each CPRD file into the OMOP CDM. The first was direct mapping used when converting CPRD files that had comparable tables in the OMOP CDM. The original names were changed to the OMOP equivalent and source values converted to standardized OMOP concepts. CPRD files: Patient (to OMOP Person), Staff (to Provider), Drug Issue (to Drug Exposure) and Practice (to Care Site) were directly mapped. The second method was indirect where for the CPRD Observation file the domain of each data row was used to assign data to proper OMOP tables or columns done by converting all source values to standard concepts.

Results: The OMOP CDM conversion populated 12 tables and 20,240,453,339 rows, with the largest table being the Measurement table (5,202,579,174 data row). Mapping source values to OMOP standard concepts, we found 60.2% (46,413 of 77,149) of source concepts were also standard concepts. The Drug Exposure table had the fewest source values already in the standard form as only 4.7% (1,433 of 30,194) of the source concepts were standard concepts. On a data retention level, only 2.00% of all data rows were excluded as they did not have a clear fit in the developed CDM and were not able to stand alone without additional information which was not present.

Conclusion: CPRD AURUM was successfully converted into the OMOP CDM with minimal data loss. Existing OHDSI tools were used with the converted data to show efficacy of the converted data. The existence of a standardized version of CPRD AURUM data vastly increases its reusability in future research due to increased understanding and tools available.

概要

はじめに: 近年、共通データモデル（CDMS）を用いて臨床データを標準化する取り組みが活発化している。CDMの使用により、データ構造の迅速な理解と既存ツールの再利用が可能となる。CDMの一つにOMOP CDMがある。Clinical Practice Research Datalink（CPRD）は、英国の一般開業医データを収集するデータ収集プログラムである。

目的: 我々の目的は、CPRD AURUMデータの静的コピーをOMOP CDMに変換し、変換されたデータで既存のツールを実行することであった。

方法: 各CPRDファイルをOMOP CDMに変換するために、2つの方法が用いられた。1つ目は、OMOP CDMで同等のテーブルを持つCPRDファイルを変換する際に使用したダイレクトマッピングである。元の名称はOMOPに相当するものに変更され、ソース値は標準化されたOMOPの概念に変換された。CPRDファイル **Patient**（→OMOP Person）、**Staff**（→Provider）、**Drug Issue**（→Drug Exposure）、**Practice**（→Care Site）は直接マッピングされた。第二の方法は間接的なもので、CPRD **Observation** ファイルの各データ行のドメインを用いて、すべてのソース値を標準的な概念に変換することにより、データを適切なOMOPテーブルまたはカラムに割り当てた。

結果: OMOP CDMへの変換では、12のテーブル、20,240,453,339行が使用され、最大のテーブルはMeasurementテーブル（5,202,579,174行）であった。ソース値をOMOP標準概念にマッピングした結果、ソース概念の60.2%（77,149件のうち46,413件）が標準概念でもあることがわかりました。Drug Exposureテーブルでは、ソース概念のわずか4.7%（30,194件のうち1,433件）が標準概念であったため、標準的な形式がすでに存在するソース値が最も少なかった。データ保持のレベルでは、全データ行のわずか2.00%が、開発されたCDMに明確に適合せず、存在しない追加情報なしには自立できないとして除外された。

結論: CPRD AURUMは、最小限のデータ損失でOMOP CDMへの変換に成功した。既存のOHDSIツールは変換されたデータとともに使用され、変換されたデータの有効性が示された。CPRD AURUMデータの標準化されたバージョンが存在することで、理解と利用可能なツールが増え、将来の研究における再利用性が大幅に向上した。



エストニアの健康記録をOMOP化した

> [JAMIA Open](#). 2023 Dec 5;6(4):ooad100. doi: 10.1093/jamiaopen/ooad100. eCollection 2023 Dec.

Transforming Estonian health data to the Observational Medical Outcomes Partnership (OMOP) Common Data Model: lessons learned

Marek Oja¹, Sirli Tamm¹, Kerli Mooses¹, Maarja Pajusalu¹, Harry-Anton Talvik^{1 2}, Anne Ott¹,
Marianna Laht¹, Maria Malk¹, Marcus Lõo¹, Johannes Holm¹, Markus Haug¹,
Hendrik Šuvalov¹, Dage Särg^{1 2}, Jaak Vilo^{1 2}, Sven Laur¹, Raivo Kolde¹, Sulev Reisberg^{1 2}

Affiliations – collapse

Affiliations

¹ Institute of Computer Science, University of Tartu, 51009 Tartu, Estonia.

² STACC, 51009 Tartu, Estonia.

PMID: 38058679 PMCID: [PMC10697784](#) DOI: [10.1093/jamiaopen/ooad100](#)

Abstract

Objective: To describe the reusable transformation process of electronic health records (EHR), claims, and prescriptions data into Observational Medical Outcome Partnership (OMOP) Common Data Model (CDM), together with challenges faced and solutions implemented.

Materials and methods: We used Estonian national health databases that store almost all residents' claims, prescriptions, and EHR records. To develop and demonstrate the transformation process of Estonian health data to OMOP CDM, we used a 10% random sample of the Estonian population ($n = 150\,824$ patients) from 2012 to 2019 (MAITT dataset). For the sample, complete information from all 3 databases was converted to OMOP CDM version 5.3. The validation was performed using open-source tools.

Results: In total, we transformed over 100 million entries to standard concepts using standard OMOP vocabularies with the average mapping rate 95%. For conditions, observations, drugs, and measurements, the mapping rate was over 90%. In most cases, SNOMED Clinical Terms were used as the target vocabulary.

Discussion: During the transformation process, we encountered several challenges, which are described in detail with concrete examples and solutions.

Conclusion: For a representative 10% random sample, we successfully transferred complete records from 3 national health databases to OMOP CDM and created a reusable transformation process. Our work helps future researchers to transform linked databases into OMOP CDM more efficiently, ultimately leading to better real-world evidence.

概要

目的: 電子カルテ（EHR）、請求データ、処方データをOMOP（Observational Medical Outcome Partnership）共通データモデル（CDM）に再利用可能に変換するプロセスを、直面した課題と実施した解決策とともに記述すること。

材料と方法: ほぼすべての住民の請求、処方、EHR記録が保存されているエストニアの全国健康データベースを使用した。エストニアの健康データのOMOP CDMへの変換プロセスを開発し実証するために、2012年から2019年までのエストニア人口の10%無作為サンプル（n = 150,824患者）を使用した（MAITTデータセット）。サンプルについては、3つのデータベースすべてからの完全な情報をOMOP CDMバージョン5.3に変換した。検証はオープンソースツールを用いて行った。

結果: 合計で1億件以上のエントリーを、標準的なOMOP語彙を用いて標準的な概念に変換し、平均マッピング率は95%であった。病態、観察、薬物、測定については、90%以上のマッピング率であった。ほとんどの場合、SNOMED Clinical Terms がターゲット語彙として使用された。

考察: 変換の過程でいくつかの課題に遭遇したが、具体的な例と解決策とともに詳しく説明する。

結論: 代表的な10%の無作為サンプルについて、3つの国民健康データベースからOMOP CDMへの完全なレコードの転送に成功し、再利用可能な変換プロセスを作成した。我々の研究は、将来の研究者がリンクされたデータベースをより効率的にOMOP CDMに変換するのに役立ち、最終的にはより良い実世界のエビデンスにつながる。



医用画像CDM (OMOPの拡張として)

› JCO Clin Cancer Inform. 2023 Sep;7:e2300101. doi: 10.1200/CCI.23.00101.

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

Varvara Kalokyri¹, Haridimos Kondylakis¹, Stelios Sfakianakis¹, Katerina Nikiforaki¹, Ioannis Karatzanis¹, Simone Mazzetti^{1 2 3}, Nikolaos Tachos^{1 4}, Daniele Regge^{1 3}, Dimitrios I Fotiadis^{1 4}, Konstantinos Marias¹, Manolis Tsiknakis¹

PMID: 38061012

Affiliations

- 1 Institute of Computer Science, Foundation of Research and Technology Hellas, Heraklion, Greece.
- 2 Department of Surgical Sciences, University of Turin, Turin, Italy.
- 3 Radiology Unit, Candiolo Cancer Institute, FPO-IRCCS, Candiolo, Italy.
- 4 Biomedical Research Institute, Foundation of Research and Technology Hellas, University Campus of Ioannina, Ioannina, Greece.

Abstract

Purpose: The explosion of big data and artificial intelligence has rapidly increased the need for integrated, homogenized, and harmonized health data. Many common data models (CDMs) and standard vocabularies have appeared in an attempt to offer harmonized access to the available information, with Observational Medical Outcomes Partnership (OMOP)-CDM being one of the most prominent ones, allowing the standardization and harmonization of health care information. However, despite its flexibility, still capturing imaging metadata along with the corresponding clinical data continues to pose a challenge. This challenge arises from the absence of a comprehensive standard representation for image-related information and subsequent image curation processes and their interlinkage with the respective clinical information. Successful resolution of this challenge holds the potential to enable imaging and clinical data to become harmonized, quality-checked, annotated, and ready to be used in conjunction, in the development of artificial intelligence models and other data-dependent use cases.

Methods: To address this challenge, we introduce medical imaging (MI)-CDM—an extension of the OMOP-CDM specifically designed for registering medical imaging data and curation-related processes. Our modeling choices were the result of iterative numerous discussions among clinical and AI experts to enable the integration of imaging and clinical data in the context of the ProCAncer-I project, for answering a set of clinical questions across the prostate cancer's continuum.

Results: Our MI-CDM extension has been successfully implemented for the use case of prostate cancer for integrating imaging and curation metadata along with clinical information by using the OMOP-CDM and its oncology extension.

概要

目的: ビッグデータと人工知能の爆発的な普及により、統合され、均質化され、調和された健康データの必要性が急速に高まっている。利用可能な情報への調和されたアクセスを提供する試みとして、多くの共通データモデル（CDM）と標準語彙が登場しており、中でもOMOP-CDMは最も著名なものの1つで、医療情報の標準化と調和を可能にしている。しかし、その柔軟性にもかかわらず、対応する臨床データとともに画像メタデータを取り込むことは依然として課題となっている。この課題は、画像関連情報の包括的な標準表現がないこと、その後の画像キュレーションプロセスとそれぞれの臨床情報との相互リンクに起因する。この課題の解決に成功すれば、画像データと臨床データを調和させ、品質をチェックし、注釈を付け、人工知能モデルやその他のデータに依存するユースケースの開発で併用できるようになる可能性がある。

方法: この課題に対処するため、我々は医用画像（MI）-CDMを導入する。これは、医用画像データとキュレーション関連プロセスを登録するために特別に設計されたOMOP-CDMの拡張である。我々のモデリングの選択は、ProCAncer-Iプロジェクトにおいて、前立腺癌の連続性を横断する一連の臨床的疑問に答えるために、画像データと臨床データの統合を可能にするために、臨床とAIの専門家が何度も議論を繰り返した結果である。

結果: 我々のMI-CDM拡張機能は、OMOP-CDMとその腫瘍学拡張機能を用いて、臨床情報とともに画像とキュレーションのメタデータを統合する前立腺がんのユースケースに実装することに成功した。



滲出型の加齢黄斑変性の医療経済負荷

> [JMIR Public Health Surveill. 2023 Dec 8;9:e49852. doi: 10.2196/49852.](#)

Patient-Centered Economic Burden of Exudative Age-Related Macular Degeneration: Retrospective Cohort Study

Kyungseon Choi ^{# 1 2}, Sang Jun Park ^{# 3}, Sola Han ⁴, Yongseok Mun ⁵, Da Yun Lee ³,
Dong-Jin Chang ⁶, Seok Kim ⁷, Sooyoung Yoo ⁷, Se Joon Woo ³, Kyu Hyung Park ⁸,
Hae Sun Suh ^{1 2 9}

PMID: 38064251

- ¹Department of Regulatory Science, Graduate School, Kyung Hee University, Seoul, Republic of Korea.
- ²Institute of Regulatory Innovation Through Science, Kyung Hee University, Seoul, Republic of Korea.
- ³Department of Ophthalmology, Seoul National University College of Medicine, Seoul National University Bundang Hospital, Seongnam, Republic of Korea.
- ⁴Health Outcomes Division, College of Pharmacy, University of Texas at Austin, Austin, TX, United States.
- ⁵Department of Ophthalmology, Kangnam Sacred Heart Hospital, Hallym University College of Medicine, Seoul, Republic of Korea.
- ⁶Department of Ophthalmology and Visual Science, Yeouido St Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea.
- ⁷Healthcare ICT Research Center, Office of eHealth Research and Businesses, Seoul National University Bundang Hospital, Seongnam, Republic of Korea.
- ⁸Department of Ophthalmology, Seoul National University College of Medicine, Seoul National University Hospital, Seoul, Republic of Korea.
- ⁹College of Pharmacy, Kyung Hee University, Seoul, Republic of Korea.

Abstract

Background: Exudative age-related macular degeneration (AMD), one of the leading causes of blindness, requires expensive drugs such as anti-vascular endothelial growth factor (VEGF) agents. The long-term regular use of effective but expensive drugs causes an economic burden for patients with exudative AMD. However, there are no studies on the long-term patient-centered economic burden of exudative AMD after reimbursement of anti-VEGFs.

Objective: This study aimed to evaluate the patient-centered economic burden of exudative AMD for 2 years, including nonreimbursement and out-of-pocket costs, compared with nonexudative AMD using the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM).

Methods: This retrospective cohort study was conducted using the OMOP CDM, which included 2,006,478 patients who visited Seoul National University Bundang Hospital from June 2003 to July 2019. We defined the exudative AMD group as patients aged >50 years with a diagnosis of exudative AMD and a prescription for anti-VEGFs or verteporfin. The control group was defined as patients aged >50 years without a diagnosis of exudative AMD or a prescription for anti-VEGFs or verteporfin. To adjust for selection bias, controls were matched by propensity scores using regularized logistic regression with a Laplace prior. We measured any medical cost occurring in the hospital as the economic burden of exudative AMD during a 2-year follow-up period using 4 categories: total medical cost, reimbursement cost, nonreimbursement cost, and out-of-pocket cost. To estimate the average cost by adjusting the confounding variable and overcoming the positive skewness of costs, we used an exponential conditional model with a generalized linear model.

Results: We identified 931 patients with exudative AMD and matched 783 (84.1%) with 2918 patients with nonexudative AMD. In the exponential conditional model, the total medical, reimbursement, nonreimbursement, and out-of-pocket incremental costs were estimated at US \$3426, US \$3130, US \$366, and US \$561, respectively, in the first year and US \$1829, US \$1461, US \$373, and US \$507, respectively, in the second year. All incremental costs in the exudative AMD group were 1.89 to 4.25 and 3.50 to 5.09 times higher in the first and second year, respectively, than those in the control group ($P < .001$ in all cases).

Conclusions: Exudative AMD had a significantly greater economic impact ($P < .001$) for 2 years on reimbursement, nonreimbursement, and out-of-pocket costs than nonexudative AMD after adjusting for baseline demographic and clinical characteristics using the OMOP CDM. Although economic policies could relieve the economic burden of patients with exudative AMD over time, the out-of-pocket cost of exudative AMD was still higher than that of nonexudative AMD for 2 years. Our findings support the need for expanding reimbursement strategies for patients with exudative AMD given the significant economic burden faced by patients with incurable and fatal diseases both in South Korea and worldwide.

データ分析は、ATLAS (ver 2.10.1)、R ソフトウェア (ver 4.0.3)、我々が開発した Health Resources Econometric Analysis Tool (HERMES) を使用して実施した。

概要

背景: 失明の主な原因の一つである滲出型加齢黄斑変性（AMD）には、抗血管内皮増殖因子（VEGF）薬などの高価な薬剤が必要である。有効だが高価な薬剤を長期間常用することは、滲出型AMD患者にとって経済的負担となる。しかし、抗VEGF薬の償還後の滲出型AMDの患者中心の長期的な経済的負担に関する研究はない。

目的: 本研究では、OMOP CDMを用いて、非滲出型AMDと比較して、償還外費用や自己負担費用を含む滲出型AMDの2年間の患者中心の経済的負担を評価することを目的とした。

方法: この後方視的コホート研究は、2003年6月から2019年7月までにソウル大学盆唐病院を受診した患者2,006,478人を含むOMOP CDMを用いて実施した。50歳以上で滲出型AMDと診断され、抗VEGF薬またはバーテポルフィンが処方された患者を滲出型AMD群と定義した。対照群は、滲出型AMDと診断されず、抗VEGFsまたはバーテポルフィンの処方を受けていない50歳を超える患者と定義した。選択バイアスを調整するため、対照群はラプラス事前分布を用いた正規化ロジスティック回帰を用いて傾向スコアでマッチさせた。2年間の追跡期間中に滲出型AMDの経済的負担として病院で発生したあらゆる医療費を、総医療費、償還費、償還外費用、自己負担費用の4つのカテゴリーを用いて測定した。交絡変数を調整し、費用の正の歪度を克服して平均費用を推定するために、一般化線形モデルを用いた指数条件付きモデルを用いた。

結果: 滲出型AMD患者931例を同定し、783例（84.1%）を非滲出型AMD患者2918例とマッチさせた。指数条件付きモデルでは、医療費、償還費、非償還費、自己負担額の総増加費用は、1年目にそれぞれ3426USドル、3130USドル、366USドル、561USドル、2年目にそれぞれ1829USドル、1461USドル、373USドル、507USドルと推定された。滲出型AMD群のすべての増分費用は、対照群と比較して、1年目は1.89~4.25倍、2年目は3.50~5.09倍であった（すべての症例で $P<.001$ ）。

結論: 滲出型AMDは、OMOP CDMを用いてベースラインの人口統計学的特徴および臨床的特徴を調整した後、非滲出型AMDと比較して、償還、非償還、および自己負担費用において、2年間の経済的影響が有意に大きかった（ $P<.001$ ）。経済政策により、滲出型AMD患者の経済的負担は時間とともに軽減される可能性があるが、滲出型AMDの自己負担額は2年間、非滲出型AMDよりも依然として高かった。我々の知見は、滲出型AMD患者に対する償還戦略を拡大する必要性を支持するものである。これは、韓国と世界の両方において、難病で致命的な疾患を有する患者が直面する大きな経済的負担を考慮したものである。



ATLASは治験適格基準判定につかえるか？

› [Sci Rep. 2023 Dec 18;13\(1\):22457. doi: 10.1038/s41598-023-49560-w.](#)

Implementation of inclusion and exclusion criteria in clinical studies in OHDSI ATLAS software

Romina Blasini ¹, Kornelia Marta Buchowicz ^{2 3}, Henning Schneider ^{2 3}, Birgit Samans ³,
Keywan Sohrabi ^{2 3}

Affiliations – collapse

Affiliations

- 1 Institute of Medical Informatics, Justus Liebig University, Giessen, Germany.
romina.blasini@informatik.med.uni-giessen.de.
- 2 Institute of Medical Informatics, Justus Liebig University, Giessen, Germany.
- 3 Faculty of Health Sciences, University of Applied Sciences, Giessen, Germany.

PMID: 38105303 PMCID: [PMC10725886](#) DOI: [10.1038/s41598-023-49560-w](#)

Abstract

Clinical trials are essential parts of a medical study process, but studies are often cancelled due to a lack of participants. Clinical Trial Recruitment Support Systems are systems that help to increase the number of participants by seeking more suitable subjects. The software ATLAS (developed by Observational Health Data Sciences and Informatics) can support the launch of a clinical trial by building cohorts of patients who fulfill certain criteria. The correct use of medical classification systems aiming at clearly defined inclusion and exclusion criteria in the studies is an important pillar of this software. The aim of this investigation was to determine whether ATLAS can be used in a Clinical Trial Recruitment Support System to portray the eligibility criteria of clinical studies. Our analysis considered the number of criteria feasible for integration with ATLAS and identified its strengths and weaknesses. Additionally, we investigated whether nonrepresentable criteria were associated with the utilized terminology systems. We analyzed ATLAS using 223 objective eligibility criteria from 30 randomly selected trials conducted in the last 10 years. In the next step, we selected appropriate ICD, OPS, LOINC, or ATC codes to feed the software. We classified each criterion and study based on its implementation capability in the software, ensuring a clear and logical progression of information. Based on our observations, 51% of the analyzed inclusion criteria were fully implemented in ATLAS. Within our selected example set, 10% of the studies were classified as fully portrayable, and 73% were portrayed to some extent. Additionally, we conducted an evaluation of the software regarding its technical limitations and interaction with medical classification systems. To improve and expand the scope of criteria within a cohort definition in a practical setting, it is recommended to work closely with personnel involved in the study to define the criteria precisely and to carefully select terminology systems. The chosen criteria should be combined according to the specific setting. Additional work is needed to specify the significance and amount of the extracted criteria.

概要

臨床試験は、医学研究のプロセスにおいて必要不可欠なものであるが、参加者不足のために試験が中止されることも少なくない。臨床試験参加者募集支援システムは、より適切な被験者を探すことによって参加者数を増やすことを支援するシステムである。ソフトウェア ATLASは、一定の基準を満たす患者のコホートを構築することで、臨床試験の開始をサポートすることができる。このソフトウェアの重要な柱は、明確に定義された組み入れ基準および除外基準を目指した医学的分類システムの正しい使用である。

この調査の目的は、**臨床試験の適格基準**を描写するためにATLASを臨床試験募集支援システムで使用できるかどうかを決定することであった。

我々の分析では、ATLASとの統合が可能な基準の数を検討し、その長所と短所を明らかにした。さらに、表現不可能な基準が利用されている用語体系と関連しているかどうかを調査した。過去10年間に実施された試験の中から無作為に抽出した30試験から、223の客観的適格基準を用いてATLASを分析した。次のステップでは、適切なICD、OPS、LOINC、ATCコードを選択し、ソフトウェアに入力した。我々は、各基準と試験をソフトウェアへの実装能力に基づいて分類し、情報の明確かつ論理的な進行を確保した。

我々の観察によると、分析した包含基準の51%はATLASに完全に実装されていた。我々が選択した事例セットの中で、10%の研究が完全に描写可能であると分類され、73%はある程度描写可能であった。さらに、ソフトウェアの技術的境界と医学分類システムとの相互作用に関する評価を行った。

実用的な設定でコホート定義内の基準の範囲を改善し拡大するためには、研究関係者と緊密に協力して基準を正確に定義し、用語体系を慎重に選択することが推奨される。選択した基準は、特定の設定に応じて組み合わせるべきである。抽出された基準の意義と量を特定するためには、さらなる作業が必要である。



フランスの総合がんセンターのデータ戦略

> PLOS Digit Health. 2023 Dec 19;2(12):e0000415. doi: 10.1371/journal.pdig.0000415.
eCollection 2023 Dec.

From data strategy to implementation to advance cancer research and cancer care: A French comprehensive cancer center experience

Pierre Heudel ¹, Hugo Crochet ², Thierry Durand ³, Philippe Zrounba ⁴, Jean-Yves Blay ^{1 5}

Affiliations – collapse

Affiliations

- 1 Department of Medical Oncology, Centre Léon Bérard, Lyon, France.
- 2 Data and Artificial Intelligence Team, Centre Léon Bérard, Lyon, France.
- 3 Data protection officer, Centre Léon Bérard, Lyon, France.
- 4 Department of Surgical Oncology, Centre Léon Bérard, Lyon, France.
- 5 General Director, Centre Léon Bérard, Lyon, France.

Abstract

In a comprehensive cancer center, effective data strategies are essential to evaluate practices, and outcome, understanding the disease and prognostic factors, identifying disparities in cancer care, and overall developing better treatments. To achieve these goals, the Center Léon Bérard (CLB) considers various data collection strategies, including electronic medical records (EMRs), clinical trial data, and research projects. Advanced data analysis techniques like natural language processing (NLP) can be used to extract and categorize information from these sources to provide a more complete description of patient data. Data sharing is also crucial for collaboration across comprehensive cancer centers, but it must be done securely and in compliance with regulations like GDPR. To ensure data is shared appropriately, CLB should develop clear data sharing policies and share data in a controlled, standardized format like OSIRIS RWD, OMOP and FHIR. The UNICANCER initiative has launched the CONSORE project to support the development of a structured and standardized repository of patient data to improve cancer research and patient outcomes. Real-world data (RWD) studies are vital in cancer research as they provide a comprehensive and accurate picture of patient outcomes and treatment patterns. By incorporating RWD into data collection, analysis, and sharing strategies, comprehensive cancer centers can take a more comprehensive and patient-centered approach to cancer research. In conclusion, comprehensive cancer centers must take an integrated approach to data collection, analysis, and sharing to enhance their understanding of cancer and improve patient outcomes. Leveraging advanced data analytics techniques and developing effective data sharing policies can help cancer centers effectively harness the power of data to drive progress in cancer research.

概要

総合的ながんセンターでは、効果的なデータ戦略が診療や転帰を評価し、疾患や予後因子を理解し、がん医療における格差を明らかにし、全体としてより良い治療法を開発するために不可欠である。これらの目標を達成するため、Center Léon Bérard (CLB) では、電子カルテ (EMR)、臨床試験データ、研究プロジェクトなど、さまざまなデータ収集戦略を検討している。自然言語処理 (NLP) のような高度なデータ解析技術は、これらのソースから情報を抽出・分類し、患者データをより完全に記述するために使用することができる。

データ共有もまた、包括的ながんセンター間の協力にとって極めて重要であるが、GDPRのような規制を遵守し、安全に行われなければならない。データが適切に共有されるよう、CLBは明確なデータ共有ポリシーを策定し、OSIRIS RWD、OMOP、FHIRのような管理された標準化されたフォーマットでデータを共有すべきである。

UNICANCERイニシアチブは、がん研究と患者の転帰を改善するために、患者データの構造化・標準化されたリポジトリの開発を支援するCONSOIREプロジェクトを立ち上げた。

リアルワールドデータ (RWD) 研究は、患者の転帰や治療パターンを包括的かつ正確に把握できるため、がん研究において不可欠である。データ収集、分析、共有戦略にRWDを取り入れることで、総合がんセンターはがん研究により包括的で患者中心のアプローチをとることができる。

結論として、総合がんセンターは、がんに対する理解を深め、患者の転帰を改善するために、データの収集、分析、共有に統合的なアプローチをとらなければならない。高度なデータ解析技術を活用し、効果的なデータ共有方針を策定することで、がんセンターはデータの力を効果的に活用し、がん研究の進展を促進することができる。



OHDSI標準化ボキャブラリ

> J Am Med Inform Assoc. 2024 Jan 4:ocad247. doi: 10.1093/jamia/ocad247. Online ahead of print.

OHDSI Standardized Vocabularies—a large-scale centralized reference ontology for international data harmonization

Christian Reich^{1 2 3}, Anna Ostropolets^{1 4 5}, Patrick Ryan^{1 4 6}, Peter Rijnbeek^{1 3},
Martijn Schuemie^{1 6}, Alexander Davydov^{1 5}, Dmitry Dymshyts^{1 6}, George Hripcsak^{1 4}

PMID: 38175665

Affiliations

- 1 Coordinating Center, Observational Health Data Sciences and Informatics, New York City NY 10032, United States.
- 2 OHDSI Center at the Roux Institute, Northeastern University, Portland ME 04101, United States.
- 3 Department of Medical Informatics, Erasmus University Medical Center, 3015 GD Rotterdam, The Netherlands.
- 4 Department of Biomedical Informatics, Columbia University Medical Center, New York City NY 10032, United States.
- 5 Odysseus Data Services, Cambridge MA 02142, United States.
- 6 Observational Health Data Analytics, Janssen Research & Development, Titusville NJ 08560, United States.

Abstract

Importance: The Observational Health Data Sciences and Informatics (OHDSI) is the largest distributed data network in the world encompassing more than 331 data sources with 2.1 billion patient records across 34 countries. It enables large-scale observational research through standardizing the data into a common data model (CDM) (Observational Medical Outcomes Partnership [OMOP] CDM) and requires a comprehensive, efficient, and reliable ontology system to support data harmonization.

Materials and methods: We created the OHDSI Standardized Vocabularies—a common reference ontology mandatory to all data sites in the network. It comprises imported and de novo-generated ontologies containing concepts and relationships between them, and the praxis of converting the source data to the OMOP CDM based on these. It enables harmonization through assigned domains according to clinical categories, comprehensive coverage of entities within each domain, support for commonly used international coding schemes, and standardization of semantically equivalent concepts.

Results: The OHDSI Standardized Vocabularies comprise over 10 million concepts from 136 vocabularies. They are used by hundreds of groups and several large data networks. More than 8600 users have performed 50 000 downloads of the system. This open-source resource has proven to address an impediment of large-scale observational research—the dependence on the context of source data representation. With that, it has enabled efficient phenotyping, covariate construction, patient-level prediction, population-level estimation, and standard reporting.

Discussion and conclusion: OHDSI has made available a comprehensive, open vocabulary system that is unmatched in its ability to support global observational research. We encourage researchers to exploit it and contribute their use cases to this dynamic resource.

概要

重要性: OHDSIは、34カ国にまたがる21億の患者記録、331以上のデータソースを包含する世界最大の分散データネットワークである。OHDSIは、データを共通データモデル OMOP CDM に標準化することにより、大規模な観察研究を可能にしており、データのハーモナイゼーションをサポートするために、包括的、効率的、かつ信頼性の高いオントロジーシステムを必要としている。

材料と方法 : OHDSI Standardized Vocabulariesは、ネットワーク内の全データサイトに必須の共通参照オントロジーである。このオントロジーは、インポートされたオントロジーと新規に生成されたオントロジーで構成され、概念とそれらの間の関係を含み、これらに基づいてソースデータをOMOP CDMに変換する実践的なものである。これにより、臨床カテゴリーに従ったドメインの割り当て、各ドメイン内のエンティティの包括的なカバー、一般的に使用されている国際的なコーディングスキームのサポート、および意味的に等価な概念の標準化による調和が可能になる。

結果: OHDSI標準化語彙は、136の語彙からなる1,000万以上の概念で構成されている。数百のグループといくつかの大規模なデータネットワークで使用されている。8600人以上のユーザーが50,000回のダウンロードを行っている。このオープンソースリソースは、大規模な観察研究の障害である、ソースデータ表現の文脈への依存を解決することが証明されている。これにより、効率的な表現型分類、共変量構築、患者レベルの予測、集団レベルの推定、標準的な報告が可能になった。

考察と結論: OHDSIは、グローバルな観察研究をサポートする能力において、他に類を見ない包括的でオープンな語彙システムを利用可能にした。このダイナミックなリソースを研究者が活用し、ユースケースを投稿することを奨励する。



ワクチンはlong COVIDの予防効果もある

➤ [Lancet Respir Med.](#) 2024 Jan 11:S2213-2600(23)00414-9. doi: 10.1016/S2213-2600(23)00414-9.
Online ahead of print.

The effectiveness of COVID-19 vaccines to prevent long COVID symptoms: staggered cohort study of data from the UK, Spain, and Estonia

Martí Català¹, Núria Mercadé-Besora², Raivo Kolde³, Nhung T H Trinh⁴, Elena Roel²,
Edward Burn¹, Trishna Rathod-Mistry¹, Kristin Kostka¹, Wai Yi Man¹, Antonella Delmestri¹,
Hedvig M E Nordeng⁵, Anneli Uusküla⁶, Talita Duarte-Salles⁷, Daniel Prieto-Alhambra⁸,
Annika M Jödicke¹

PMID: 38219763

Affiliations

- ¹Pharmaco- and Device Epidemiology Group, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK.
- ²Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina, Barcelona, Spain.
- ³Institute of Computer Science, University of Tartu, Tartu, Estonia.
- ⁴Pharmacoepidemiology and Drug Safety Research Group, Department of Pharmacy, Faculty of Mathematics and Natural Sciences, University of Oslo, Oslo, Norway.
- ⁵Pharmacoepidemiology and Drug Safety Research Group, Department of Pharmacy, Faculty of Mathematics and Natural Sciences, University of Oslo, Oslo, Norway; Department of Child Health and Development, Norwegian Institute of Public Health, Oslo, Norway.
- ⁶Department of Family Medicine and Public Health, University of Tartu, Tartu, Estonia.
- ⁷Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina, Barcelona, Spain; Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, Netherlands.
- ⁸Pharmaco- and Device Epidemiology Group, Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK; Oxford National Institute for Health and Care Research Biomedical Research Centre, University of Oxford, Oxford, UK; Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, Netherlands. Electronic address: daniel.prietoalhambra@ndorms.ox.ac.uk.

Abstract

Background: Although vaccines have proved effective to prevent severe COVID-19, their effect on preventing long-term symptoms is not yet fully understood. We aimed to evaluate the overall effect of vaccination to prevent long COVID symptoms and assess comparative effectiveness of the most used vaccines (ChAdOx1 and BNT162b2).

Methods: We conducted a staggered cohort study using primary care records from the UK (Clinical Practice Research Datalink [CPRD] GOLD and AURUM), Catalonia, Spain (Information System for Research in Primary Care [SIDIAP]), and national health insurance claims from Estonia (CORIVA database). All adults who were registered for at least 180 days as of Jan 4, 2021 (the UK), Feb 20, 2021 (Spain), and Jan 28, 2021 (Estonia) comprised the source population.

Vaccination status was used as a time-varying exposure, staggered by vaccine rollout period. Vaccinated people were further classified by vaccine brand according to their first dose received. The primary outcome definition of long COVID was defined as having at least one of 25 WHO-listed symptoms between 90 and 365 days after the date of a PCR-positive test or clinical diagnosis of COVID-19, with no history of that symptom 180 days before SARS-Cov-2 infection. Propensity score overlap weighting was applied separately for each cohort to minimise confounding. Sub-distribution hazard ratios (sHRs) were calculated to estimate vaccine effectiveness against long COVID, and empirically calibrated using negative control outcomes. Random effects meta-analyses across staggered cohorts were conducted to pool overall effect estimates.

Findings: A total of 1 618 395 (CPRD GOLD), 5 729 800 (CPRD AURUM), 2 744 821 (SIDIAP), and 77 603 (CORIVA) vaccinated people and 1 640 371 (CPRD GOLD), 5 860 564 (CPRD AURUM), 2 588 518 (SIDIAP), and 302 267 (CORIVA) unvaccinated people were included. Compared with unvaccinated people, overall HRs for long COVID symptoms in people vaccinated with a first dose of any COVID-19 vaccine were 0·54 (95% CI 0·44-0·67) in CPRD GOLD, 0·48 (0·34-0·68) in CPRD AURUM, 0·71 (0·55-0·91) in SIDIAP, and 0·59 (0·40-0·87) in CORIVA. A slightly stronger preventative effect was seen for the first dose of BNT162b2 than for ChAdOx1 (sHR 0·85 [0·60-1·20] in CPRD GOLD and 0·84 [0·74-0·94] in CPRD AURUM).

Interpretation: Vaccination against COVID-19 consistently reduced the risk of long COVID symptoms, which highlights the importance of vaccination to prevent persistent COVID-19 symptoms, particularly in adults.

概要

背景: ワクチンは重篤なCOVID-19の予防に有効であることが証明されているが、長期症状の予防効果についてはまだ十分に解明されていない。我々は、COVIDの長期症状を予防するためのワクチン接種の総合的な効果を評価し、最も使用されているワクチン（ChAdOx1とBNT162b2）の比較効果を評価することを目的とした。

方法: 英国のプライマリケア記録（Clinical Practice Research Datalink [CPRD] GOLDおよびAURUM）、スペインのカタルーニャ（Information System for Research in Primary Care [SIDIAP]）、およびエストニアの国民健康保険請求（CORIVAデータベース）を用いて時間差のあるコホート研究を実施した。2021年1月4日（英国）、2021年2月20日（スペイン）、2021年1月28日（エストニア）の時点で180日以上登録されていたすべての成人を対象集団とした。ワクチン接種状況は、ワクチン展開期間によって時系列的に変化する曝露として使用された。ワクチン接種者はさらに、最初に接種したワクチンブランドで分類した。長期COVIDの主要アウトカム定義は、PCR陽性検査またはCOVID-19の臨床診断日から90～365日の間に、WHOにリストアップされた25の症状のうち少なくとも1つを有し、SARS-Cov-2感染180日前にその症状の既往がないこととした。交絡を最小化するために、各コホートに対して個別に傾向スコア重複加重を適用した。長期のCOVIDに対するワクチン効果を推定するために部分分布ハザード比（sHR）を算出し、陰性対照転帰を用いて経験的に較正した。全体的な効果推定値をプールするために、時差のあるコホート間のランダム効果メタアナリシスを実施した。

結果: ワクチン接種者1 618 395人（CPRD GOLD）、5 729 800人（CPRD AURUM）、2 744 821人（SIDIAP）、77 603人（CORIVA）、ワクチン未接種者1 640 371人（CPRD GOLD）、5 860 564人（CPRD AURUM）、2 588 518人（SIDIAP）、302 267人（CORIVA）を対象とした。ワクチン未接種者と比較して、いずれかのCOVID-19ワクチンを初回接種した人のCOVID症状長期化の全体HRは、CPRD GOLDでは0.54（95%CI 0.44-0.67）、CPRD AURUMでは0.48（0.34-0.68）、SIDIAPでは0.71（0.55-0.91）、CORIVAでは0.59（0.40-0.87）であった。BNT162b2の初回投与では、ChAdOx1よりもやや強い予防効果が認められた（CPRD GOLDではsHR 0.85 [0.60-1.20]、CPRD AURUMでは0.84 [0.74-0.94]）。

解釈: COVID-19に対するワクチン接種は、長期にわたるCOVID症状のリスクを一貫して減少させたことから、特に成人におけるCOVID-19症状の持続を予防するためのワクチン接種の重要性が強調された。



文書をメタデータでLOINC分類する

› [AMIA Annu Symp Proc. 2024 Jan 11:2023:1017-1026. eCollection 2023.](#)

Mapping Clinical Documents to the *Logical Observation Identifiers, Names and Codes (LOINC) Document Ontology* using Electronic Health Record Systems Structured Metadata

Huzaiifa Khan ^{1 2}, Abu Saleh Mohammad Mosa ², Vyshnavi Paka ², Md Kamruz Zaman Rana ²,
Vasanthi Mandhadi ², Soliman Islam ², Hua Xu ^{3 4}, James C McClay ², Sraboni Sarker ⁵,
Praveen Rao ⁵, Lemuel R Waitman ²

PMID: 38222329

Affiliations

- 1 MU Institute of Data Science and Informatics, University of Missouri-Columbia.
- 2 Department of Health Management and Informatics, School of Medicine, University of Missouri-Columbia.
- 3 Yale University, New Haven, CT, USA.
- 4 OHDSI Consortium, Natural Language Processing Working Group.
- 5 Department of Electrical and Computer Science, School of Engineering, University of Missouri-Columbia.

Abstract

As Electronic Health Record (EHR) systems increase in usage, organizations struggle to maintain and categorize clinical documentation so it can be used for clinical care and research. While prior research has often employed natural language processing techniques to categorize free text documents, there are shortcomings relative to computational scalability and the lack of key metadata within notes' text. This study presents a framework that can allow institutions to map their notes to the LOINC document ontology using a Bag of Words approach. After preliminary manual value- set mapping, an automated pipeline that leverages key dimensions of metadata from structured EHR fields aligns the notes with the dimensions of the document ontology. This framework resulted in 73.4% coverage of EHR documents, while also mapping 132 million notes in less than 2 hours; an order of magnitude more efficient than NLP based methods.

概要

電子カルテ(EHR)システムの利用が増加するにつれ、組織は臨床文書を維持・分類し、臨床ケアや研究に利用できるようにすることに奮闘している。先行研究では、フリーテキスト文書を分類するために自然言語処理技術がしばしば採用されているが、計算の拡張性やメモのテキスト内の主要なメタデータの欠如に関連する欠点がある。本研究では、Bag of Wordsのアプローチを用いて、医療機関がノートをLOINC文書オントロジーにマッピングできるフレームワークを提示する。手作業による予備的な値セットのマッピングの後、構造化されたEHRフィールドからのメタデータの主要な次元を活用する自動化されたパイプラインが、ノートを文書オントロジーの次元に整合させる。このフレームワークにより、EHR文書の73.4%がカバーされるとともに、1億3,200万件のノートが2時間未満でマッピングされた。

補足

臨床文書をLOINC文書コードで分類するタスク。

この研究における私たちの目標は、利用可能なメタデータを通じてのみ臨床文書を LOINC コードに迅速かつ正確にマッピングできる一般化されたフレームワークを作成することでした。このようなフレームワークの能力は、メタデータの可用性と品質に基づいて制限されますが、これにより、品質や速度を犠牲にすることなく、方法の複雑さが大幅に軽減されます。

すべての臨床ノートの 73.4% をカバーしていることを考えると、私たちのフレームワークは非常にうまく機能していると言えます。残りの 26.6% のメモを詳しく調べたところ、かなりの数のメモが電話メッセージ、ポータルメモ、または初期のマッピング作業の実際の焦点ではないその他の種類の外れ値メモであることがわかりました。



妊娠中の時期を特定するPhenotyping

> [AMIA Annu Symp Proc. 2024 Jan 11:2023:1096-1104. eCollection 2023.](#)

Computational Phenotyping of OMOP CDM Normalized EHR for Prenatal and Postpartum Episodes: An Informatics Framework and Clinical Implementation on All of Us

Tianchu Lyu ¹, Chen Liang ^{1 2}

Affiliations – collapse

Affiliations

1 University of South Carolina, Columbia, South Carolina, USA.

2 National Institutes of Health, Bethesda, Maryland, USA.

PMID: 38222375 PMCID: [PMC10785883](#)

Abstract

The use of Electronic Health Records (EHR) in pregnancy care and obstetrics-gynecology (OB/GYN) research has increased in recent years. In pregnancy, timing is important because clinical characteristics, risks, and patient management are different in each stage of pregnancy. However, the difficulty of accurately differentiating pregnancy episodes and temporal information of clinical events presents unique challenges for EHR phenotyping. In this work, we introduced the concept of time relativity and proposed a comprehensive framework of computational phenotyping for prenatal and postpartum episodes based on the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). We implemented it on the All of Us national EHR database and identified 6,280 pregnancies with accurate start and end dates among 5,399 female patients. With the ability to identify different episodes in pregnancy care, this framework provides new opportunities for phenotyping complex clinical events and gestational morbidities for pregnant women, thus improving maternal and infant health.

概要

近年、妊娠ケアや産婦人科（OB/GYN）研究における電子カルテ（EHR）の利用が増加している。妊娠では、臨床的特徴、リスク、患者管理が妊娠の各段階で異なるため、タイミングが重要である。しかし、妊娠エピソードと臨床イベントの時間的情報を正確に区別することは困難であり、EHRフェノタイピングには独特の課題がある。本研究では、時間相対性の概念を導入し、OMOP CDMに基づく、出生前および分娩後エピソードの計算によるフェノタイピングの包括的な枠組みを提案した。我々はこれをAll of Us 全国EHRデータベースに実装し、5,399人の女性患者のうち、正確な開始日と終了日を持つ6,280件の妊娠を同定した。妊娠ケアにおけるさまざまなエピソードを同定する能力により、このフレームワークは、妊婦の複雑な臨床イベントや妊娠中の病的状態を表現型化する新たな機会を提供し、母体と乳児の健康を改善する。



文書をメタデータでLOINC分類する(その2)

> AMIA Annu Symp Proc. 2024 Jan 11:2023:834-843. eCollection 2023.

Standardizing Multi-site Clinical Note Titles to LOINC Document Ontology: A Transformer-based Approach

Xu Zuo¹, Yujia Zhou¹, Jon Duke^{2 3}, George Hripcsak^{4 3}, Nigam Shah^{5 3}, Juan M Banda^{6 3}, Ruth Reeves^{7 3}, Timothy Miller^{8 3}, Lemuel R Waitman⁹, Karthik Natarajan^{4 3}, Hua Xu^{10 3}

Affiliations

- 1 University of Texas Health Science Center at Houston, Houston, TX, USA.
- 2 Georgia Institute of Technology, Atlanta, GA, USA.
- 3 OHDSI Consortium, Natural Language Processing Working Group.
- 4 Columbia University, New York City, NY, USA.
- 5 Stanford University, Stanford, CA, USA.
- 6 Georgia State University, Atlanta, GA, USA.
- 7 Vanderbilt University Medical Center, Nashville, TN, USA.
- 8 Boston Children's Hospital, Boston, MA, USA.
- 9 University of Missouri, Columbia, MO, USA.
- 10 Yale University, New Haven, CT, USA.

Abstract

The types of clinical notes in electronic health records (EHRs) are diverse and it would be great to standardize them to ensure unified data retrieval, exchange, and integration. The LOINC Document Ontology (DO) is a subset of LOINC that is created specifically for naming and describing clinical documents. Despite the efforts of promoting and improving this ontology, how to efficiently deploy it in real-world clinical settings has yet to be explored. In this study we evaluated the utility of LOINC DO by mapping clinical note titles collected from five institutions to the LOINC DO and classifying the mapping into three classes based on semantic similarity between note *titles* and LOINC DO codes. Additionally, we developed a standardization pipeline that automatically maps clinical note titles from multiple sites to suitable LOINC DO codes, without accessing the content of clinical notes. The pipeline can be initialized with different large language models, and we compared the performances between them. The results showed that our automated pipeline achieved an accuracy of 0.90. By comparing the manual and automated mapping results, we analyzed the coverage of LOINC DO in describing multi-site clinical note titles and summarized the potential scope for extension.

概要

電子カルテ（EHR）における臨床記録の種類は多様であり、統一されたデータ検索、交換、統合を保証するためにそれらを標準化することは素晴らしいことである。LOINC文書オントロジー（DO）はLOINCのサブセットであり、特に臨床文書の命名と記述のために作成された。このオントロジーの普及と改良の努力にもかかわらず、実世界の臨床環境で効率的に展開する方法はまだ検討されていない。本研究では、5つの施設から収集した臨床ノートのタイトルをLOINC DOにマッピングし、ノートのタイトルとLOINC DOコードの意味的類似性に基づいてマッピングを3つのクラスに分類することで、LOINC DOの有用性を評価した。さらに、臨床ノートの内容にアクセスすることなく、複数の施設からの臨床ノートのタイトルを適切なLOINC DOコードに自動的にマッピングする標準化パイプラインを開発した。このパイプラインは、異なる大規模言語モデルで初期化することができ、それらの性能を比較した。その結果、我々の自動化パイプラインは0.90の精度を達成した。手動マッピングと自動マッピングの結果を比較することで、LOINC DOが多施設のクリニカルノートタイトルを記述する際のカバレッジを分析し、拡張の可能性をまとめた。



Phenotypingにおいて2番目のイベントを
追加することによる影響

> [J Pharm Pharm Sci](#). 2024 Jan 3;26:12095. doi: 10.3389/jpps.2023.12095. eCollection 2023.

Comparing broad and narrow phenotype algorithms: differences in performance characteristics and immortal time incurred

Joel N Swerdel ^{1 2}, Mitchell M Conover ^{1 2}

Affiliations – collapse

Affiliations

- 1 Observational Health Data Analytics, Global Epidemiology, Janssen Research and Development, Titusville, NJ, United States.
- 2 Observational Health Data Sciences and Informatics, New York, NY, United States.

PMID: 38235322 PMCID: [PMC10791821](#) DOI: [10.3389/jpps.2023.12095](#)

Abstract

Introduction: When developing phenotype algorithms for observational research, there is usually a trade-off between definitions that are sensitive or specific. The objective of this study was to estimate the performance characteristics of phenotype algorithms designed for increasing specificity and to estimate the immortal time associated with each algorithm. **Materials and methods:** We examined algorithms for 11 chronic health conditions. The analyses were from data from five databases. For each health condition, we created five algorithms to examine performance (sensitivity and positive predictive value (PPV)) differences: one broad algorithm using a single code for the health condition and four narrow algorithms where a second diagnosis code was required 1-30 days, 1-90 days, 1-365 days, or 1- all days in a subject's continuous observation period after the first code. We also examined the proportion of immortal time relative to time-at-risk (TAR) for four outcomes. The TAR's were: 0-30 days after the first condition occurrence (the index date), 0-90 days post-index, 0-365 days post-index, and 0-1,095 days post-index. Performance of algorithms for chronic health conditions was estimated using PheValuator (V2.1.4) from the OHDSI toolstack. Immortal time was calculated as the time from the index date until the first of the following: 1) the outcome; 2) the end of the outcome TAR; 3) the occurrence of the second code for the chronic health condition. **Results:** In the first analysis, the narrow phenotype algorithms, i.e., those requiring a second condition code, produced higher estimates for PPV and lower estimates for sensitivity compared to the single code algorithm. In all conditions, increasing the time to the required second code increased the sensitivity of the algorithm. In the second analysis, the amount of immortal time increased as the window used to identify the second diagnosis code increased. The proportion of TAR that was immortal was highest in the 30 days TAR analyses compared to the 1,095 days TAR analyses. **Conclusion:** Attempting to increase the specificity of a health condition algorithm by adding a second code is a potentially valid approach to increase specificity, albeit at the cost of incurring immortal time.

概要

はじめに: 観察研究用の表現型アルゴリズムを開発する場合、通常、感度の高い定義か特異性の高い定義かのトレードオフが存在する。本研究の目的は、特異性を高めるように設計された表現型アルゴリズムの性能特性を推定し、各アルゴリズムに関連する不死時間を推定することである。

材料と方法: 11の慢性的な健康状態に関するアルゴリズムを検討した。解析は5つのデータベースのデータから行った。各健康状態について、性能（感度および陽性的中率（PPV））の違いを調べるために5つのアルゴリズムを作成した：健康状態の単一のコードを使用する1つの広範なアルゴリズムと、最初のコード後の被験者の連続観察期間において1～30日、1～90日、1～365日、または1～全日に2つ目の診断コードが必要とされる4つの狭義のアルゴリズム。また、4つのアウトカムについて、タイムアットリスク（TAR）に対する不死時間の割合を調べた。TARは以下の通りである：TARは、最初の状態発生（指標日）から0～30日後、指標後0～90日後、指標後0～365日後、指標後0～1,095日後である。慢性健康状態に関するアルゴリズムの性能は、OHDSIツールスタックのPheValuator（V2.1.4）を用いて推定した。不死時間は、指標日から次のうち最初のものまでの時間として計算した：1)転帰、2)転帰TARの終了、3)慢性健康状態の2つ目のコードの発生。

結果: 最初の解析では、狭い表現型アルゴリズム、すなわち2つ目の病態コードを必要とするアルゴリズムは、単一コードアルゴリズムと比較してPPVの推定値が高く、感度の推定値が低かった。すべての条件において、必要とされる2つ目のコードまでの時間を長くすると、アルゴリズムの感度が高くなった。2つ目の解析では、2つ目の診断コードを特定するためのウィンドウが大きくなるにつれて、不死時間の量が増加した。不死であったTARの割合は、1,095日のTAR解析と比較して30日のTAR解析で最も高かった。

結論: 2つ目のコードを追加することで健康状態アルゴリズムの特異性を高めようとする試みは、不死時間を要するものの、特異性を高める有効なアプローチである可能性がある。

結論

ここで紹介する研究は、疫学研究において狭い表現型アルゴリズムと広い表現型アルゴリズムを使用するコストと利点の証拠を提供します。

狭い表現型の利点は、対象となる健康状態に対するより高い陽性的中率で示され、研究者が正しい状態を研究していることを保証します。そのコストは、研究設計に不死の時間を費やすことから発生します。これは、死亡などの重篤な転帰を調べる場合には重要な問題ですが、死亡率が低い転帰を調べる場合には影響は小さいことが示されています。

狭い表現型アルゴリズムが必要な研究デザインを行う研究者は、診断コード後の非常に短い期間内に特定の手順や他の治療を必要とするなど、2番目のコードを使用する代替アプローチを検討する可能性があります。



再入院予測に在宅医療診療記録を活用

> [Stud Health Technol Inform.](#) 2024 Jan 25:310:1438-1439. doi: 10.3233/SHTI231233.

Prediction Models for Readmission Using Home Healthcare Notes and OMOP-CDM

Sujin Gan ¹, Chungsoo Kim ¹, Dong Yun Lee ², Rae Woong Park ^{1 2}

Affiliations – collapse

Affiliations

- 1 Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, Korea.
- 2 Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Korea.

PMID: 38269685 DOI: [10.3233/SHTI231233](#)

Abstract

This study developed readmission prediction models using Home Healthcare (HHC) documents via natural language processing (NLP). An electronic health record of Ajou University Hospital was used to develop prediction models (A reference model using only structured data, and an NLP-enriched model with structured and unstructured data). Among 573 patients, 63 were readmitted to the hospital. Five topics were extracted from HHC documents and improved the model performance (AUROC 0.740).

概要

本研究では、自然言語処理（NLP）を用いて在宅医療（HHC）文書を用いた再入院予測モデルを開発した。予測モデルの開発には、亜城大学病院の電子カルテを用いた（構造化データのみを用いた参照モデルと、構造化データと非構造化データを用いたNLP強化モデル）。573人の患者のうち、63人が再入院した。HHC文書から5つのトピックが抽出され、モデル性能を向上させた（AUROC 0.740）。

補足

対象は1日以上入院患者。アウトカムは90日以内の再入院。

構造化データは電子カルテのデータ（おそらく退院後の外来/在宅データも含む？）。

NLP強化は、退院後の在宅医療診療録を自然言語処理したものを追加。

- ・ 構造化データのみ AUROC 0.654 (0.499-0.810).
- ・ NLP強化モデル AUROC 0.740 (0.630-0.852)



R-CDMを使った眼科画像多施設研究

> Stud Health Technol Inform. 2024 Jan 25;310:48-52. doi: 10.3233/SHTI230925.

Multi-Institutional Collaborative Research Using Ophthalmic Medical Image Data Standardized by Radiology Common Data Model (R-CDM)

ChulHyoung Park¹, Sang Jun Park², Da Yun Lee², Seng Chan You³, Kihwang Lee⁴,
Rae Woong Park^{1 5}

PMID: 38269763

Affiliations

- 1 Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Republic of Korea.
- 2 Department of Ophthalmology, Seoul National University Bundang Hospital, Seongnam, Republic of Korea.
- 3 Department of Biomedicine Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea.
- 4 Department of Ophthalmology, Ajou University School of Medicine, Suwon, Republic of Korea.
- 5 Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, Republic of Korea.

Abstract

Observational Medical Outcome Partners - Common Data Model (OMOP-CDM) is an international standard model for standardizing electronic medical record data. However, unstructured data such as medical image data which is beyond the scope of standardization by the current OMOP-CDM is difficult to be used in multi-institutional collaborative research. Therefore, we developed the Radiology-CDM (R-CDM) which standardizes medical imaging data. As a proof of concept, 737,500 Optical Coherence Tomography (OCT) data from two tertiary hospitals in South Korea is standardized in the form of R-CDM. The relationship between chronic disease and retinal thickness was analyzed by using the R-CDM. Central macular thickness and retinal nerve fiber layer (RNFL) thickness were significantly thinner in the patients with hypertension compared to the control cohort. It is meaningful in that multi-institutional collaborative research using medical image data and clinical data simultaneously can be conducted very efficiently.

概要

OMOP-CDMは、電子カルテデータを標準化するための国際標準モデルである。しかし、医用画像データのような非構造化データは、現在のOMOP-CDMでは標準化の対象外であり、多施設共同研究での利用は困難である。そこで我々は、医用画像データを標準化したRadiology-CDM (R-CDM)を開発した。概念実証として、韓国の2つの3次病院から737,500件の光干渉断層計 (OCT) データをR-CDMの形で標準化した。慢性疾患と網膜厚の関係をR-CDMを用いて解析した。中心黄斑厚と網膜神経線維層 (RNFL) 厚は、対照コホートと比較して高血圧患者で有意に薄かった。医用画像データと臨床データを同時に用いた多施設共同研究が非常に効率的に実施できる点で有意義である。



用語マッピングの標準的で反復可能なプロセス

> [Stud Health Technol Inform.](#) 2024 Jan 25;310:68-73. doi: 10.3233/SHTI230929.

Promoting Learning Health System Cycles by Optimizing EHR Data Clinical Concept Encoding Processes

Piper Ranallo ¹, Bronwyn Southwell ², Christopher Tignanelli ², Steven G Johnson ², Richard Krueger ¹, Tess Severeid-Groth ¹, Adam Carvel ¹, Genevieve B Melton ^{1 2}

Affiliations – collapse

Affiliations

¹ Fairview Health Services, Minneapolis, MN USA.

² University of Minnesota, Minneapolis, MN USA.

PMID: 38269767 DOI: [10.3233/SHTI230929](#)

Abstract

Electronic health records (EHRs) and other real-world data (RWD) are critical to accelerating and scaling care improvement and transformation. To efficiently leverage it for secondary uses, EHR/RWD should be optimally managed and mapped to industry standard concepts (ISCs). Inherent challenges in concept encoding usually result in inefficient and costly workflows and resultant metadata representation structures outside the EHR. Using three related projects to map data to ISCs, we describe the development of standard, repeatable processes for precisely and unambiguously representing EHR data using appropriate ISCs within the EHR platform lifecycle and mappings specific to SNOMED-CT for Demographics, Specialty and Services. Mappings in these 3 areas resulted in ISC mappings of 779 data elements requiring 90 new concept requests to SNOMED-CT and 738 new ISCs mapped into the workflow within an accessible, enterprise-wide EHR resource with supporting processes.

概要

電子カルテ(EHR)とその他の実世界データ(RWD)は、ケアの改善と変革を加速し、拡大するために不可欠である。二次利用のために効率的に活用するためには、EHR/RWDは最適に管理され、業界標準概念(ISC)にマッピングされるべきである。概念エンコーディングにおける本質的な課題は、通常、非効率的でコストのかかるワークフローと、その結果としてのEHR外のメタデータ表現構造に帰結する。データをISCにマッピングするための3つの関連プロジェクトを用いて、EHRプラットフォームのライフサイクルの中で適切なISCを使用してEHRデータを正確かつ明確に表現するための標準的で反復可能なプロセスの開発と、人口統計、専門分野、サービスのSNOMED-CTに特化したマッピングについて述べる。これら3つの領域におけるマッピングの結果、SNOMED-CTへの90の新しいコンセプト要求を必要とする779のデータ要素のISCマッピングと、それをサポートするプロセスとともに、アクセス可能な全社的なEHRリソース内のワークフローにマッピングされた738の新しいISCが得られた。

補足

用語マッピングの標準的で反復可能なプロセスの開発を目指すため、米国の特定地域10の医療施設での3つのマッピングを伴う既存プロジェクトを実例に、SNOMED-CTに特化したワークフローを検討した。

- ・気をつけたこと：共有リポジトリ（できるだけEHR内）への保存、可能であればデータエンコーディングをEHR構築へ不可欠化すること、そのデータビルドの専門家の事前特定とプロジェクトへの参加。
- ・院内コードから標準コードへのマッピングプロセスにおいて参考になる。



値集合の適切さ

› [Stud Health Technol Inform.](#) 2024 Jan 25:310:164-168. doi: 10.3233/SHTI230948.

Toward Real-World Reproducibility: Verifying Value Sets for Clinical Research

Scott L DuVall¹, Craig G Parker², Amanda R Shields³, Patrick R Alba⁴, Julie A Lynch¹, Michael E Matheny⁵, Aaron W C Kamaau²

Affiliations – collapse

Affiliations

- 1 Department of Veterans Affairs, United States.
- 2 Navidence, United States.
- 3 Quanten LLC, United States.
- 4 University of Utah, United States.
- 5 Vanderbilt University Medical Center, United States.

PMID: 38269786 DOI: [10.3233/SHTI230948](#)

Abstract

Standardized operational definitions are an important tool to improve reproducibility of research using secondary real-world healthcare data. This approach was leveraged for studies evaluating the effectiveness of AZD7442 as COVID-19 pre-exposure prophylaxis across multiple healthcare systems. Value sets were defined, grouped, and mapped. Results of this exercise were reviewed and recorded. Value sets were updated to reflect findings.

概要

標準化された操作的定義は、RWDを用いた研究の再現性を向上させるための重要なツールである。このアプローチは、複数の医療システムにおけるCOVID-19の曝露前予防としてのAZD7442の有効性を評価する研究に活用された。バリューセットが定義され、グループ化され、マッピングされた。その結果をレビューし、記録した。バリューセットは結果を反映して更新された。

補足

OMOPでいえば、コホート定義で使うコンセプトセットの話。
AZD7442の評価のために定めた様々な参照ソースからの用語/コードの“値集合”をOMOPコンセプトで表現し、VAのOMOP DBにに対して適用してレビューした。

この報告では、表現・適用に関しての工夫、OMOPコンセプトにマッピングすることでオリジナルの値から指すものが変わってしまう問題点（粒度の違いなど）、実データには存在しない値(用語)が多数あること、等が指摘されている。

用語の値集合の作り方（実装との乖離、実データとの乖離）や、実装(OMOPコンセプト)と欲しい値集合との差異は0にはできないであろうこと、など考えさせられる点が多い。



より高度な品質測定システム

> Stud Health Technol Inform. 2024 Jan 25;310:349-353. doi: 10.3233/SHTI230985.

Development of Integrated Data Quality Management System for Observational Medical Outcomes Partnership Common Data Model

Seol Whan Oh ^{1 2}, Soo Jeong Ko ^{1 2}, Yun Seon Im ^{1 2}, Surin Jung ¹, Bo Yeon Choi ^{1 2},
Jae Yoon Kim ^{1 2}, Sunghyeon Park ^{1 2}, Wona Choi ¹, In Young Choi ¹

Affiliations – collapse

Affiliations

- 1 Department of Medical Informatics, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea.
- 2 Department of Biomedicine & Health Sciences, The Catholic University of Korea, Seoul, Republic of Korea.

PMID: 38269823 DOI: 10.3233/SHTI230985

Abstract

The amount of research on the gathering and handling of healthcare data keeps growing. To support multi-center research, numerous institutions have sought to create a common data model (CDM). However, data quality issues continue to be a major obstacle in the development of CDM. To address these limitations, a data quality assessment system was created based on the representative data model OMOP CDM v5.3.1. Additionally, 2,433 advanced evaluation rules were created and incorporated into the system by mapping the rules of existing OMOP CDM quality assessment systems. The data quality of six hospitals was verified using the developed system and an overall error rate of 0.197% was confirmed. Finally, we proposed a plan for high-quality data generation and the evaluation of multi-center CDM quality.

概要

医療データの収集と取り扱いに関する研究は増え続けている。多施設共同研究をサポートするために、多くの機関が共通データモデル（CDM）の作成を模索してきた。しかし、データの質の問題はCDMの開発における大きな障害であり続けている。これらの限界に対処するため、代表的なデータモデルであるOMOP CDM v5.3.1に基づいてデータ品質評価システムが作成された。さらに、既存のOMOP CDM品質評価システムのルールをマッピングすることにより、2,433の高度評価ルールを作成し、システムに組み込んだ。開発したシステムを用いて6病院のデータ品質を検証し、全体のエラー率は0.197%であることを確認した。最後に、高品質なデータ作成と多施設CDM品質評価のための計画を提案した。



HADES 紹介

> [Stud Health Technol Inform.](#) 2024 Jan 25;310:966-970. doi: 10.3233/SHTI231108.

Health-Analytics Data to Evidence Suite (HADES): Open-Source Software for Observational Research

Martijn Schuemie^{1 2 3}, Jenna Reps^{1 2 4}, Adam Black^{1 5}, Frank Defalco^{1 2}, Lee Evans^{1 6},
Egill Fridgeirsson^{1 4}, James P Gilbert^{1 2}, Chris Knoll^{1 2}, Martin Lavalley^{1 7},
Gowtham A Rao^{1 2}, Peter Rijnbeek^{1 4}, Katy Sadowski^{1 8}, Anthony Sena^{1 2 4}, Joel Swerdel^{1 2},
Ross D Williams^{1 4}, Marc Suchard^{1 3 9}

PMID: 38269952

Affiliations

- 1 Observational Health Data Science and Informatics, New York, NY, USA.
- 2 Observational Health Data Analytics, Johnson & Johnson, Titusville, NJ, USA.
- 3 Department of Biostatistics, UCLA, Los Angeles, CA, USA.
- 4 Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands.
- 5 Odysseus Data Services Inc., Cambridge, MA, USA.
- 6 LTS Computing LLC, West Chester, PA, USA.
- 7 Virginia Commonwealth University, Richmond, VA, USA.
- 8 TrialSpark Inc., New York, NY, USA.
- 9 VA Informatics and Computing Infrastructure, Department of Veterans Affairs, Salt Lake City, UT, USA.

Abstract

The Health-Analytics Data to Evidence Suite (HADES) is an open-source software collection developed by Observational Health Data Sciences and Informatics (OHDSI). It executes directly against healthcare data such as electronic health records and administrative claims, that have been converted to the Observational Medical Outcomes Partnership (OMOP) Common Data Model. Using advanced analytics, HADES performs characterization, population-level causal effect estimation, and patient-level prediction, potentially across a federated data network, allowing patient-level data to remain locally while only aggregated statistics are shared. Designed to run across a wide array of technical environments, including different operating systems and database platforms, HADES uses continuous integration with a large set of unit tests to maintain reliability. HADES implements OHDSI best practices, and is used in almost all published OHDSI studies, including some that have directly informed regulatory decisions.

概要

Health-Analytics Data to Evidence Suite (HADES)は、OHDSIによって開発されたオープンソースのソフトウェアコレクションである。HADESは、OMOP共通データモデルに変換された電子カルテや行政請求などのヘルスケアデータに対して直接実行されます。HADESは、高度な分析を用いて、特性解析、集団レベルの因果関係推定、患者レベルの予測を行い、連携データネットワーク全体で、患者レベルのデータはローカルに残し、集計された統計情報のみを共有することが可能です。HADESは、さまざまなオペレーティングシステムやデータベースプラットフォームなど、幅広い技術環境で実行できるように設計されており、信頼性を維持するために、大規模なユニットテストと継続的インテグレーションを使用しています。HADESはOHDSIのベストプラクティスを実装しており、規制当局の決定に直接影響を与えたものも含め、発表されたほぼすべてのOHDSI研究で使用されています。

HADESは、高度な分析のためにC++、Java、Pythonを採用し、Rパッケージを通して実現されている。例えば、コアとなる回帰エンジンCyclopsは、C++で回帰モデルを最適化し、大規模データセットを処理する。SQLはデータ操作を管理し、さまざまなプラットフォームに変換される。HADESパッケージの一部はCRANに、その他はGitHubにある。

連携ネットワークにおける患者のプライバシーを保護するため、主なパッケージは、セル内の数が少ない場合の目隠しなどのプライバシー対策を提供している。データは人間のレビュー可能なCSVファイルで共有される。

HADESのドキュメントは、Rの標準であるroxygen2とpkgdownを採用し、リファレンスマニュアルとビネットを含む。Windows、MacOS、Linuxにまたがる継続的な統合テストにより、システム間の互換性が保証されている。

コホートはHADES解析の中核をなす要素であり、ある期間にわたって特定の基準を満たす個人を捕捉する。コホートは、曝露（例：ワルファリン曝露）、結果（例：出血症例）、または特別なグループ（例：妊婦）を意味する。

HADESは入力としてコホートを必要とし、洗練されたロジックはパッケージによって管理される： Caprは定義の作成に、PhenotypeLibraryは承認されたコホート定義の保存に、CirceRはSQL/ヒト可読変換に、CohortGeneratorはCDMと互換性のあるインスタンス化に、CohortDiagnosticsは評価のためにPheValuatorとともに使用される。

主な HADES 分析パッケージは次のとおりです。

● **DataQualityDashboard** は、広範なテストを通じて適合性、完全性、妥当性をチェックします。

● **PatientLevelPrediction** は、OHDSI の予測モデル フレームワークに準拠し、CDM データからの幅広い予測子を利用します。回帰や勾配ブースティングなどのさまざまなアルゴリズムをサポートし、OHDSI ネットワークでの迅速な外部検証を可能にします。

● **CohortMethod** は、交絡調整に大規模傾向スコア (LSPS) を利用して、因果効果推定に比較コホート設計を適用します。

● **EvidenceSynthesis** は、メタ分析を通じて複数のデータベースの結果を結合します。これには、カウントが低いかゼロの場合に Cox モデルを組み合わせるための最近の統計的アプローチが含まれています。

● **EmpiricalCalibration** は、科学的精度のための不確実性を組み込んで、因果関係の推定値を強化するためにネガティブ コントロール効果の推定値を採用します。



> [Stud Health Technol Inform.](#) 2024 Jan 25:310:1349-1351. doi: 10.3233/SHTI231189.

Feasibility of Applying the OMOP Common Data Model to Traditional Eastern Asian Medicine Dataset

Jiyun Cha ¹, Eun Kyoung Ahn ¹, Young-Heum Yoon ², Man Young Park ¹

Affiliations – collapse

Affiliations

- 1 Korea Institute of Oriental Medicine, Daejeon, 34054, Republic of Korea.
- 2 National Institute for Korean Medicine Development, Seoul, 04516, Republic of Korea.

PMID: 38270038 DOI: [10.3233/SHTI231189](#)

Abstract

To evaluate the feasibility of applying the Observational Medical Outcome Partnership (OMOP) Common Data Model (CDM) to databases of traditional East Asian medicine (TEAM), we composed a TEAM dataset and transformed it to the OMOP CDM. We found that some important TEAM information entities could not be transformed to the OMOP CDM (version 6.0) data fields. We suggest to develop data fields and guideline for transforming TEAM data to the OMOP CDM.

概要

東アジア伝統医学（TEAM）のデータベースにOMOP CDMを適用する可能性を評価するため、TEAMのデータセットを構成し、OMOP CDMに変換した。その結果、いくつかの重要なTEAM情報エンティティがOMOP CDM（バージョン6.0）のデータフィールドに変換できないことがわかった。我々は、TEAMデータをOMOP CDMに変換するためのデータフィールドとガイドラインを開発することを提案する。

補足

3つの疾患CPGから重複を除いて合計52のエンティティが抽出された。うち42のエンティティは、医療分野（西洋医学、TEAM）に関係なく共通の臨床情報であった。しかし、10項目は従来の西洋医学では考慮されていないTEAM特有の情報であった。

テーブルレベルでは、すべてのエンティティがOMOP CDMテーブルに適切にマッピングされた。しかし、フィールドレベルでは、3つのエンティティがOMOP CDMの適切なフィールドに変換できなかった。これらはTEAM固有の情報であった：生薬処方のためのTEAM診断情報、TEAM症候、TEAM式の処方形式である。

例えば、処方された医薬品（生薬配合剤）のサブ情報として、生薬の加工や配合比に関する情報は、非常に重要である。例えば、処方薬（生薬配合剤）のサブ情報として、生薬の加工や配合比の情報は薬の効能に極めて重要であるが、既存のOMOP CDMではデータフィールドに含まれていない。OMOP CDMを適用するためには、補足的なデータフィールドが必要となる。



今月のOHDSI Global

- APAC Call テーマ

Jan. 18 APAC 2024 Kickoff; Training Session#7

- Global Community Call テーマ

Jan. 9 Welcome Back! What Can OHDSI Accomplish in 2024?

Jan. 16 Connections for Collaborations

Jan. 23 Lessons Learned from the UK Studyathon

参考 study-a-thon解説

<https://www.thehyve.nl/articles/studyathon-multidisciplinary-collaboration-in-biomedical-open-science>