

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 #48

イブニングカンファレンス(第48回)

2023.11.28



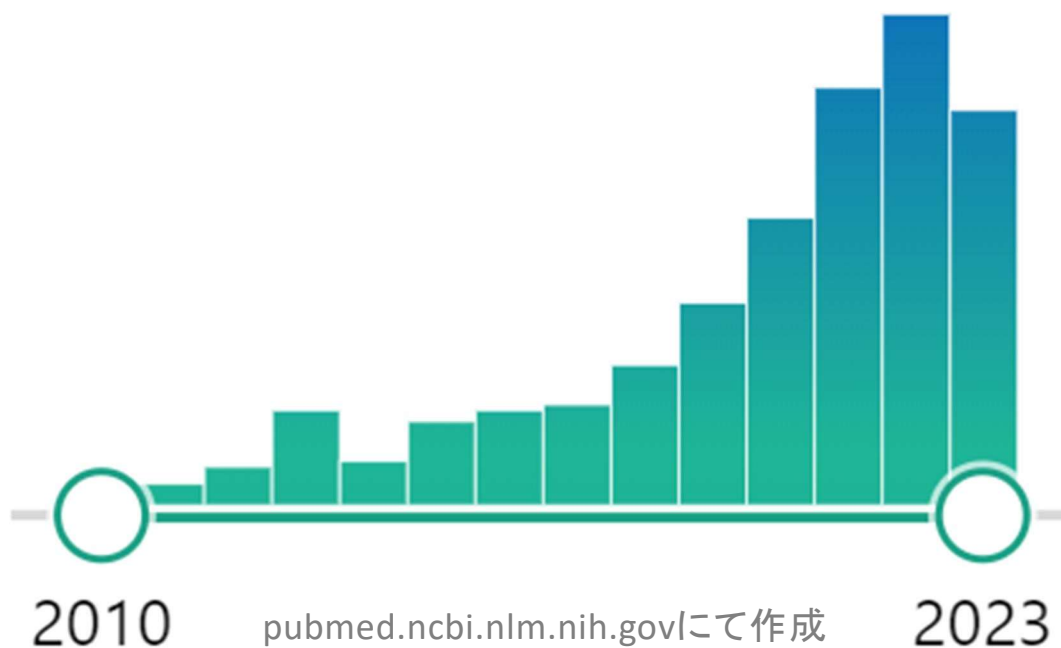
本日の内容

- OHDSI関連論文紹介
- 今月のOHDSI Global
- Global Symposium 2023 内容



OHDSI関連論文

Pubmedで“OHDSI or OMOP”を検索



全期間累計：10月370本→11月374本



ドイツ診療報酬請求データのボキャブラリ

> [JMIR Med Inform.](#) 2023 Nov 7:11:e47959. doi: 10.2196/47959.

Assessing the Use of German Claims Data Vocabularies for Research in the Observational Medical Outcomes Partnership Common Data Model: Development and Evaluation Study

Elisa Henke¹, Michéle Zoch¹, Michael Kallfelz², Thomas Ruhnke³, Liz Annika Leutner¹,
Melissa Spoden³, Christian Günster³, Martin Sedlmayr¹, Franziska Bathelt⁴

Affiliations + expand

PMID: 37942786 PMID: PMC10653283 DOI: 10.2196/47959

Affiliations

- 1 Institute for Medical Informatics and Biometry, Carl Gustav Carus Faculty of Medicine, Technische Universität Dresden, Dresden, Germany.
- 2 Odysseus Data Services GmbH, Berlin, Germany.
- 3 Wissenschaftliches Institut der AOK (AOK Research Institute), Berlin, Germany.
- 4 Thiem-Research GmbH, Cottbus, Germany.

Abstract

Background: National classifications and terminologies already routinely used for documentation within patient care settings enable the unambiguous representation of clinical information. However, the diversity of different vocabularies across health care institutions and countries is a barrier to achieving semantic interoperability and exchanging data across sites. The Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) enables the standardization of structure and medical terminology. It allows the mapping of national vocabularies into so-called standard concepts, representing normative expressions for international analyses and research. Within our project "Hybrid Quality Indicators Using Machine Learning Methods" (Hybrid-QI), we aim to harmonize source codes used in German claims data vocabularies that are currently unavailable in the OMOP CDM.

Objective: This study aims to increase the coverage of German vocabularies in the OMOP CDM. We aim to completely transform the source codes used in German claims data into the OMOP CDM without data loss and make German claims data usable for OMOP CDM-based research.

Methods: To prepare the missing German vocabularies for the OMOP CDM, we defined a vocabulary preparation approach consisting of the identification of all codes of the corresponding vocabularies, their assembly into machine-readable tables, and the translation of German designations into English. Furthermore, we used 2 proposed approaches for OMOP-compliant vocabulary preparation: the mapping to standard concepts using the Observational Health Data Sciences and Informatics (OHDSI) tool Usagi and the preparation of new 2-billion concepts (ie, concept_id >2 billion). Finally, we evaluated the prepared vocabularies regarding completeness and correctness using synthetic German claims data and calculated the coverage of German claims data vocabularies in the OMOP CDM.

Results: Our vocabulary preparation approach was able to map 3 missing German vocabularies to standard concepts and prepare 8 vocabularies as new 2-billion concepts. The completeness evaluation showed that the prepared vocabularies cover 44.3% (3288/7417) of the source codes contained in German claims data. The correctness evaluation revealed that the specified validity periods in the OMOP CDM are compliant for the majority (705,531/706,032, 99.9%) of source codes and associated dates in German claims data. The calculation of the vocabulary coverage showed a noticeable decrease of missing vocabularies from 55% (11/20) to 10% (2/20) due to our preparation approach.

Conclusions: By preparing 10 vocabularies, we showed that our approach is applicable to any type of vocabulary used in a source data set. The prepared vocabularies are currently limited to German vocabularies, which can only be used in national OMOP CDM research projects, because the mapping of new 2-billion concepts to standard concepts is missing. To participate in international OHDSI network studies with German claims data, future work is required to map the prepared 2-billion concepts to standard concepts.

概要

背景: 患者ケアの現場での文書化に既に日常的に使用されている国家的な分類や専門用語は、臨床情報の明確な表現を可能にする。しかし、医療機関や国によって異なる語彙の多様性は、セマンティックな相互運用性を実現し、施設間でデータを交換する際の障壁となっている。OMOP CDMは、構造と医学用語の標準化を可能にする。これにより、各国の語彙をいわゆる標準的な概念にマッピングすることができ、国際的な分析・研究のための規範的な表現となる。私たちのプロジェクト "Hybrid Quality Indicators Using Machine Learning Methods" (Hybrid-QI) では、OMOP CDMでは現在利用できないドイツのクレームデータ語彙で使用されているソースコードを調和させることを目的としている。

目的: 本研究の目的は、OMOP CDMにおけるドイツ語語彙のカバー率を高めることである。ドイツのクレームデータで使用されているソースコードを、データの損失なしにOMOP CDMに完全に変換し、ドイツのクレームデータをOMOP CDMに基づく研究に使用できるようにすることを目指す。

方法: OMOP CDMのために不足しているドイツ語の語彙を準備するため、対応する語彙の全コードの同定、機械可読な表への組み立て、ドイツ語の名称の英語への翻訳からなる語彙準備アプローチを定義した。さらに、OMOPに準拠した語彙を準備するために2つのアプローチを提案した。OHDSIツール Usagi を使用した標準概念へのマッピングと、新しい2Bコンセプト（すなわち、concept_id >20億）の準備である。最後に、準備された語彙の完全性と正しさについて、合成ドイツ請求データを用いて評価し、OMOP CDMにおけるドイツ請求データ語彙のカバー率を算出した。

結果 我々の語彙準備アプローチは、3つの欠落したドイツ語の語彙を標準的な概念にマッピングし、8つの語彙を新しい2Bコンセプトとして準備することができた。完全性の評価では、準備した語彙がドイツの請求データに含まれるソースコードの44.3% (3288/7417) をカバーしていることが示された。正しさの評価では、ドイツのクレームデータに含まれるソースコードと関連する日付の大部分 (705,531/706,032、99.9%) について、OMOP CDMで指定された有効期間が遵守されていることが明らかになった。語彙カバー率の計算では、我々の準備アプローチにより、欠落語彙が55% (11/20) から10% (2/20) に顕著に減少した。

結論: 10種類の語彙を準備することで、我々のアプローチがソースデータセットで使用されるあらゆるタイプの語彙に適用可能であることを示した。なぜなら、2Bコンセプトの標準概念へのマッピングが欠けているからである。ドイツのクレームデータを用いた国際的なOHDSIネットワーク研究に参加するためには、今後、準備された2Bコンセプトを標準概念にマッピングする作業が必要である。



EHDENでのOMOP CDM構築

› [J Am Med Inform Assoc. 2023 Nov 10:ocad214. doi: 10.1093/jamia/ocad214. Online ahead of print.](#)

European Health Data & Evidence Network—learnings from building out a standardized international health data network

Erica A Voss^{1 2 3}, Clair Blacketer^{1 2 3}, Sebastiaan van Sandijk^{1 4}, Maxim Moinat^{1 2}, Michael Kallfelz^{1 4}, Michel van Speybroeck³, Daniel Prieto-Alhambra^{1 2 5}, Martijn J Schuemie^{1 3 6}, Peter R Rijnbeek^{1 2}

PMID: 37952118

Affiliations

- 1 OHDSI Collaborators, Observational Health Data Sciences and Informatics (OHDSI), New York, NY, United States.
- 2 Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, the Netherlands.
- 3 Janssen Pharmaceutical Research and Development LLC, Raritan, NJ 08869, United States.
- 4 Odysseus Data Services, Prague, Czech Republic.
- 5 Centre for Statistics in Medicine, NDORMS, University of Oxford, Oxford, United Kingdom.
- 6 Department of Biostatistics, University of California, Los Angeles, CA 90095, United States.

Abstract

Objective: Health data standardized to a common data model (CDM) simplifies and facilitates research. This study examines the factors that make standardizing observational health data to the Observational Medical Outcomes Partnership (OMOP) CDM successful.

Materials and methods: Twenty-five data partners (DPs) from 11 countries received funding from the European Health Data Evidence Network (EHDEN) to standardize their data. Three surveys, DataQualityDashboard results, and statistics from the conversion process were analyzed qualitatively and quantitatively. Our measures of success were the total number of days to transform source data into the OMOP CDM and participation in network research.

Results: The health data converted to CDM represented more than 133 million patients. 100%, 88%, and 84% of DPs took Surveys 1, 2, and 3. The median duration of the 6 key extract, transform, and load (ETL) processes ranged from 4 to 115 days. Of the 25 DPs, 21 DPs were considered applicable for analysis of which 52% standardized their data on time, and 48% participated in an international collaborative study.

Discussion: This study shows that the consistent workflow used by EHDEN proves appropriate to support the successful standardization of observational data across Europe. Over the 25 successful transformations, we confirmed that getting the right people for the ETL is critical and vocabulary mapping requires specific expertise and support of tools. Additionally, we learned that teams that proactively prepared for data governance issues were able to avoid considerable delays improving their ability to finish on time.

Conclusion: This study provides guidance for future DPs to standardize to the OMOP CDM and participate in distributed networks. We demonstrate that the Observational Health Data Sciences and Informatics community must continue to evaluate and provide guidance and support for what ultimately develops the backbone of how community members generate evidence.

概要

目的: 共通データモデル（CDM）に標準化された健康データは、研究を簡素化し促進する。本研究では、OMOPのCDMに標準化された観察健康データを成功させる要因について検討する。

材料と方法: 11カ国の25のデータパートナー（DP）がEuropean Health Data Evidence Network（EHDEN）から資金提供を受け、データの標準化を行った。3回の調査、DataQualityDashboardの結果、変換プロセスの統計を定性的・定量的に分析した。成功の尺度は、ソースデータをOMOP CDMに変換するのに要した総日数と、ネットワーク研究への参加であった。

結果: CDMに変換された健康データは、1億3,300万人以上の患者に相当した。DPの100%、88%、84%が調査1、2、3を受けた。6つの主要な抽出、変換、ロード（ETL）プロセスの期間中央値は4～115日であった。25のDPのうち、21のDPが解析に適していると考えられ、そのうちの52%が期限内にデータを標準化し、48%が国際共同研究に参加した。

考察: 本研究は、EHDENで使用されている一貫したワークフローが、ヨーロッパ全域での観察データの標準化を成功させるのに適切であることを示している。25件の変換を成功させる中で、ETLには適切な人材を確保することが重要であり、語彙マッピングには特定の専門知識とツールのサポートが必要であることが確認された。さらに、データガバナンスの問題に対して積極的に準備したチームは、大幅な遅延を回避し、予定通りに完了する能力を向上させることができた。

結論: 本研究は、将来DPがOMOP CDMに標準化し、分散ネットワークに参加するための指針を提供するものである。我々は、OHDSIコミュニティが、最終的にコミュニティメンバーがエビデンスを生成する方法のバックボーンとなるものを評価し、ガイダンスとサポートを提供し続けなければならないことを実証した。



Observational Study > J Med Internet Res. 2023 Nov 13:25:e42259. doi: 10.2196/42259.

Development and Validation of a Prognostic Classification Model Predicting Postoperative Adverse Outcomes in Older Surgical Patients Using a Machine Learning Algorithm: Retrospective Observational Network Study

Jung-Yeon Choi ^{# 1}, Sooyoung Yoo ^{# 2}, Wongeun Song ^{# 2 3}, Seok Kim ², Hyunyoung Baek ², Jun Suh Lee ⁴, Yoo-Seok Yoon ^{5 6}, Seonghae Yoon ⁷, Hae-Young Lee ^{8 9}, Kwang-Il Kim ^{1 9}

Affiliations

- 1 Department of Internal Medicine, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea.
- 2 Office of eHealth Research and Business, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea.
- 3 Department of Health Science and Technology, Graduate School of Convergence Science and Technology, Seoul National University, Seongnam-si, Republic of Korea.
- 4 Department of Surgery, G Sam Hospital, Gunpo, Republic of Korea.
- 5 Department of Surgery, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea.
- 6 Department of Surgery, Seoul National University College of Medicine, Seoul, Republic of Korea.
- 7 Department of Clinical Pharmacology and Therapeutic, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea.
- 8 Department of Internal Medicine, Seoul National University Hospital, Seoul, Republic of Korea.
- 9 Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea.

PMID: 37955965

Abstract

Background: Older adults are at an increased risk of postoperative morbidity. Numerous risk stratification tools exist, but effort and manpower are required.

Objective: This study aimed to develop a predictive model of postoperative adverse outcomes in older patients following general surgery with an open-source, patient-level prediction from the Observational Health Data Sciences and Informatics for internal and external validation.

Methods: We used the Observational Medical Outcomes Partnership common data model and machine learning algorithms. The primary outcome was a composite of 90-day postoperative all-cause mortality and emergency department visits. Secondary outcomes were postoperative delirium, prolonged postoperative stay (≥ 75 th percentile), and prolonged hospital stay (≥ 21 days). An 80% versus 20% split of the data from the Seoul National University Bundang Hospital (SNUBH) and Seoul National University Hospital (SNUH) common data model was used for model training and testing versus external validation. Model performance was evaluated using the area under the receiver operating characteristic curve (AUC) with a 95% CI.

Results: Data from 27,197 (SNUBH) and 32,857 (SNUH) patients were analyzed. Compared to the random forest, Adaboost, and decision tree models, the least absolute shrinkage and selection operator logistic regression model showed good internal discriminative accuracy (internal AUC 0.723, 95% CI 0.701-0.744) and transportability (external AUC 0.703, 95% CI 0.692-0.714) for the primary outcome. The model also possessed good internal and external AUCs for postoperative delirium (internal AUC 0.754, 95% CI 0.713-0.794; external AUC 0.750, 95% CI 0.727-0.772), prolonged postoperative stay (internal AUC 0.813, 95% CI 0.800-0.825; external AUC 0.747, 95% CI 0.741-0.753), and prolonged hospital stay (internal AUC 0.770, 95% CI 0.749-0.792; external AUC 0.707, 95% CI 0.696-0.718). Compared with age or the Charlson comorbidity index, the model showed better prediction performance.

Conclusions: The derived model shall assist clinicians and patients in understanding the individualized risks and benefits of surgery.

概要

背景: 高齢者は術後の罹患リスクが高い。リスク層別化ツールは数多く存在するが、労力とマンパワーが必要である。

目的: 本研究では、一般外科手術後の高齢患者における**術後有害転帰の予測モデル**を開発することを目的とし、OHDSIによるオープンソースの患者レベルの予測で、内部および外部検証を行った。

方法: OMOPの共通データモデルと機械学習アルゴリズムを用いた。主要アウトカムは**術後90日の全死因死亡率と救急外来受診率**の複合とした。副次アウトカムは、術後せん妄、術後在院日数の延長（75パーセントイル以上）、在院日数の延長（21日以上）であった。ソウル大学盆唐病院（SNUBH）とソウル大学病院（SNUH）の共通データモデルからのデータを80%対20%に分割し、モデルのトレーニングおよびテスト対外部検証に使用した。モデルの性能は、95%信頼区間を有する受信者動作特性曲線下面積（AUC）を用いて評価した。

結果: 27,197例（SNUBH）と32,857例（SNUH）のデータが解析された。ランダムフォレスト、Adaboost、および決定木モデルと比較して、最小絶対縮小および選択演算子ロジスティック(LASSO)回帰モデルは、主要アウトカムについて良好な内部識別精度（内部AUC 0.723、95%CI 0.701-0.744）および移植性（外部AUC 0.703、95%CI 0.692-0.714）を示した。このモデルはまた、術後せん妄（内部AUC 0.754、95%CI 0.713-0.794；外部AUC 0.750、95%CI 0.727-0.772）、術後滞在延長（内部AUC 0.813、95%CI 0.800-0.825；外部AUC 0.747、95%CI 0.741-0.753）、入院期間の延長（内部AUC 0.770、95%CI 0.749-0.792；外部AUC 0.707、95%CI 0.696-0.718）であった。年齢やシャルソン併存疾患指数と比較して、このモデルはより優れた予測性能を示した。

結論: 導き出されたモデルは、臨床医と患者が手術の個々のリスクとベネフィットを理解するのに役立つであろう。



スタチン使用と頭頸部がんの転帰

Meta-Analysis > Sci Rep. 2023 Nov 13;13(1):19770. doi: 10.1038/s41598-023-45654-7.

Effect of statin use on head and neck cancer prognosis in a multicenter study using a Common Data Model

Soobeen Seol ^{# 1}, Jung Ran Choi ^{# 2}, Byungjin Choi ¹, Sungryeal Kim ³, Ja Young Jeon ⁴,
Ki Nam Park ⁵, Jae Hong Park ⁶, Min Woo Park ⁷, Young-Gyu Eun ⁸, Jung Je Park ^{9 10},
Byung-Joo Lee ¹¹, Yoo Seob Shin ², Chul-Ho Kim ², Rae Woong Park ^{12 13}, Jeon Yeob Jang ^{14 15}

Affiliations + expand

PMID: 37957229 PMID: PMC10643676 DOI: 10.1038/s41598-023-45654-7

Abstract

Few studies have found an association between statin use and head and neck cancer (HNC) outcomes. We examined the effect of statin use on HNC recurrence using the converted Observational Medical Outcome Partnership (OMOP) Common Data Model (CDM) in seven hospitals between 1986 and 2022. Among the 9,473,551 eligible patients, we identified 4669 patients with HNC, of whom 398 were included in the target cohort, and 4271 were included in the control cohort after propensity score matching. A Cox proportional regression model was used. Of the 4669 patients included, 398 (8.52%) previously received statin prescriptions. Statin use was associated with a reduced rate of 3- and 5-year HNC recurrence compared to propensity score-matched controls (risk ratio [RR], 0.79; 95% confidence interval [CI], 0.61-1.03; and RR 0.89; 95% CI 0.70-1.12, respectively). Nevertheless, the association between statin use and HNC recurrence was not statistically significant. A meta-analysis of recurrence based on subgroups, including age subgroups, showed similar trends. The results of this propensity-matched cohort study may not provide a statistically significant association between statin use and a lower risk of HNC recurrence. Further retrospective studies using nationwide claims data and prospective studies are warranted.

要旨

スタチン使用と頭頸部がん（HNC）の転帰との関連を明らかにした研究はほとんどない。われわれは、1986年から2022年の間に7つの病院において、変換されたOMOP CDMを用いて、HNC再発に対するスタチン使用の効果を検討した。対象患者9,473,551人のうち、4669人のHNC患者を同定し、そのうち398人を対象コホートに含め、傾向スコアマッチング後に4271人を対照コホートに含めた。Cox比例回帰モデルを用いた。対象となった4669人の患者のうち、398人（8.52%）が以前にスタチンの処方を受けていた。スタチンの使用は、傾向スコアマッチングを行った対照群と比較して、3年および5年HNC再発率の低下と関連していた（リスク比[RR]、それぞれ0.79；95%信頼区間[CI]、0.61-1.03；およびRR 0.89；95%CI、0.70-1.12）。それにもかかわらず、スタチン使用とHNC再発との関連は統計学的に有意ではなかった。年齢サブグループを含むサブグループに基づく再発のメタアナリシスでも同様の傾向が示された。この傾向一致コホート研究の結果は、スタチン使用とHNC再発リスク低下との統計学的に有意な関連を示していない可能性がある。全国的な請求データを用いたさらなるレトロスペクティブ研究とプロスペクティブ研究が正当化される。



今月のOHDSI Global

- APAC Call テーマ

Nov. 16 Global Symposium Recap and Training Session#6

- Global Community Call テーマ

Oct. 31 Welcome to OHDSI, part 2

Nov. 7 Meet The Titans

Nov. 14 Collaborator Showcase Honorees

Nov. 21 Collaborator Showcase Software Demos

Plus 紹介 : 日本語資料 OHDSI Tool Documents



2023 OHDSI Global Symposium

- 日程 10月20~22日
- 場所 Hilton East Brunswick Hotel & Executive Meeting Center
(East Brunswick, New Jersey, USA)

<https://ohdsi.org/ohdsi2023/>



State of the Community

Various leaders within OHDSI shared a presentation on the state of the community, with specific focuses on data standards, vocabulary enhancements and open-source development. **Speakers included:**

George Hripcsak, Columbia University

Clair Blacketer, Johnson & Johnson

Alexander Davydov, Odysseus Data Services

Katy Sadowski, Boehringer Ingelheim

Peter Rijnbeek, Erasmus MC

Mengling 'Mornin' Feng, National University of Singapore



OHDSIの多くのリーダーたちが、データ標準、語彙の強化、オープンソース開発などに焦点を当て、コミュニティの現状についてプレゼンテーションを行った。

Plenary: Improving the reliability and scale of case validation

症例検証は、規制グレードのエビデンスに必要な要素であると考えられているが、ソースレコードの人手による判定によって症例検証を行うことは、時間とリソースを必要とし、性能は未知数であり、症例セット全体のレビューや適切な定量的バイアス分析ができないような方法で実施されることが多い。

本プレナリーでは、OHDSIの共同研究者が、症例検証プロセスの信頼性と拡張性を向上させるための革新的な方法論研究とオープンソースの開発について発表し、OMOP CDM (KEEPER) の構造化データからインフォーマティクスによって強化された患者プロフィールによって、ソースレコードを置き換えることが可能であること、また、測定誤差を推定し、差分誤分類を特定するための大規模言語モデルの使用によって、人間のレビューを補完することが可能であることを示した。

KEEPER + LLMは、2つの異なるデータソースを用いた3つの実験を通して、10疾患において実証的に評価され、人間のレビュアー間の一致にはかなりの不均一性があり得るが、LLMは人間同士の一貫と同じくらい人間とも一致することが明らかになった。

講演者は以下の通り： Patrick Ryan (ジョンソン・エンド・ジョンソン、コロンビア大学)、Anna Ostropolets (Odysseus Data Service)、Martijn Schuemie (ジョンソン・エンド・ジョンソン、カリフォルニア大学ロサンゼルス校)



Panel: Lessons learned from OHDSI network studies

Network studies are challenging, but they provide a critical opportunity to generate robust, reliable real-world evidence. Leaders from recent OHDSI network studies share insights, challenges and possible solutions from recent network studies, including from the SOS Challenge. **Speakers included:**

Marc Suchard, University of California-Los Angeles

Cindy Cai, Johns Hopkins University

Seng Chan You, Yonsei University

Anthony Sena, Johnson & Johnson

Moderator: **Sarah Seager**, IQVIA



ネットワーク研究は困難ではあるが、強固で信頼できる実世界のエビデンスを生み出す重要な機会である。最近のOHDSIネットワーク研究のリーダーたちが、SOSチャレンジを含む最近のネットワーク研究から得た洞察、課題、可能な解決策を共有する。