

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 #36

イブニング カンファレンス(第36回)

2022.11.28



## 本日の内容

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



# OHDSI関連論文

Pubmedで“OHDSI or OMOP”を検索



pubmed.ncbi.nlm.nih.govにて作成

全期間累計：10月299本→11月307本



## Study-a-thonとは

> [JAMIA Open](#). 2022 Nov 17;5(4):ooac100. doi: 10.1093/jamiaopen/ooac100. eCollection 2022 Dec.

### Evaluating a novel approach to stimulate open science collaborations: a case series of "study-a-thon" events within the OHDSI and European IMI communities

N Hughes <sup>1</sup>, P R Rijnbeek <sup>2</sup>, K van Bochove <sup>3</sup>, T Duarte-Salles <sup>4</sup>, C Steinbeisser <sup>5</sup>, D Vizcaya <sup>6</sup>, D Prieto-Alhambra <sup>7</sup>, P Ryan <sup>8</sup>

Affiliations – collapse

#### Affiliations

- 1 Epidemiology, Janssen R&D, Beerse, Belgium.
- 2 Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands.
- 3 The Hyve, Utrecht, The Netherlands.
- 4 Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain.
- 5 Steinbeisser Project Management, Munich, Germany.
- 6 Bayer Pharmaceuticals, Sant Joan Despi, Spain.
- 7 NDORMS, University of Oxford, Oxford, UK.
- 8 Epidemiology, Janssen R&D, Titusville, New Jersey, USA.

## Abstract

**Objective:** We introduce and review the concept of a study-a-thon as a catalyst for open science in medicine, utilizing harmonized real world, observation health data, tools, skills, and methods to conduct network studies, generating insights for those wishing to use study-a-thons for future research.

**Materials and methods:** A series of historical study-a-thons since 2017 to present were reviewed for thematic insights as to the opportunity to accelerate the research method to conduct studies across therapeutic areas. Review of publications and experience of the authors generated insights to illustrate the conduct of study-a-thons, key learning, and direction for those wishing to conduct future such study-a-thons.

**Results:** A review of six study-a-thons have provided insights into their scientific impact, and 13 areas of insights for those wishing to conduct future study-a-thons. Defining aspects of the study-a-thon method for rapid, collaborative research through network studies reinforce the need to clear scientific rationale, tools, skills, and methods being collaboratively to conduct a focused study. Well-characterized preparatory, execution and postevent phases, coalescing skills, experience, data, clinical input (ensuring representative clinical context to the research query), and well-defined, logical steps in conducting research via the study-a-thon method are critical.

**Conclusions:** A study-a-thon is a focused multiday research event generating reliable evidence on a specific medical topic across different countries and health systems. In a study-a-thon, a multidisciplinary team collaborate to create an accelerated contribution to scientific evidence and clinical practice. It critically accelerates the research process, without inhibiting the quality of the research output and evidence generation, through a reproducible process.

## 概要

**目的:** 医療におけるオープンサイエンスの触媒として、調和された実世界の観測健康データ、ツール、スキル、および方法を活用してネットワーク研究を行うstudy-a-thonの概念を紹介・検討し、study-a-thonを今後の研究に活用したいと考える人々への洞察を生み出す。

**材料と方法:** 2017年から現在までの一連の歴史的なstudy-a-thonをレビューし、治療領域を超えた研究を行うための研究手法を加速させる機会として、テーマ別の洞察を得た。論文や著者の経験をレビューすることで、study-a-thonの実施、重要な学習、今後このようなstudy-a-thonを実施したい人への方向性を説明するための洞察を生み出した。

**結果:** 6つのstudy-a-thonのレビューから、その科学的インパクトに関する洞察と、今後study-a-thonを実施しようとする人への13の洞察が得られた。ネットワーク研究による迅速な共同研究のためのstudy-a-thon手法の側面を定義することは、科学的根拠、ツール、スキル、そして焦点を絞った研究を共同で行う方法の必要性を強化する。準備段階、実行段階、事後段階を明確にし、スキル、経験、データ、臨床情報を統合し（研究クエリに対する代表的な臨床状況を確保する）、study-a-thon法による研究を実施するための明確に定義された、論理的なステップが重要である。

**結論:** study-a-thonは、異なる国や医療システムにおいて、特定の医療トピックに関する信頼性の高いエビデンスを生成する、集中型の多日間研究イベントである。study-a-thonでは、学際的なチームが協力して、科学的根拠と臨床実践への貢献を加速させることができる。再現性のあるプロセスを通じて、研究成果やエビデンス創出の質を阻害することなく、研究プロセスを決定的に加速させる。

- 2017: 関節リウマチの比較安全性に関する最初の 3 日間のパイロット、ニューヨーク市のコロンビア大学でのプロトタイプ ライブ イベント。 [10](#)
- 2018: OHDSI と EHDEN の間で、オックスフォード大学で開催された 5 日間のライブ イベントで、TOPKAT 臨床研究プロトコルのエミュレーションに基づいて、部分的、単コンパートメント対全膝関節置換術を評価する推定および予測研究が行われました。 [11-13](#)
- 2020: OHDSI と EHDEN の協力により、スペインのバルセロナで開催された 5 日間のライブ イベントで、新たに診断されたリウマチ患者の第一選択療法と国際的な臨床ガイドラインとの比較に関する特性評価、推定、および予測研究が行われました。 [14](#)
- 2020: COVID-19 の特徴付け、推定、および予測に対するパンデミックの制限により、オンラインで開催された 4 日間の仮想イベント。パンデミックの第一波の早い段階で政策および臨床上の意思決定を支援し、その後進行中の特徴付けと推定を継続します。プロトコル。 [7](#), [15-17](#)
- 2020年：毎年開催される米国ベースのOHDSIシンポジウム中に、心血管ケアにおける外部検証の予測に関するPROTEUSと題された仮想の2日間のオンラインイベント。 [18](#)
- 2021年：IMI EHDENおよびOHDSIと協力して、IMI PIONEERプロジェクトによって5日間の仮想イベントが促進され、治療が適応されない場合に前立腺がんの進行の「経過観察」を開始するという決定の実際の結果を評価しました。表現型、特徴付け、および予測研究を利用しています。このスタディ・アソシエーションは、患者の経験の観点から研究作業を文脈化するための患者参加のレベルでも注目に値しました。 [19](#), [20](#) --





## FHIR→OMOP変換の大量処理機構を開発した

> *Int J Med Inform.* 2022 Nov 10;169:104925. doi: 10.1016/j.ijmedinf.2022.104925.  
Online ahead of print.

# An ETL-process design for data harmonization to participate in international research with German real-world data based on FHIR and OMOP CDM

Yuan Peng <sup>1</sup>, Elisa Henke <sup>2</sup>, Ines Reinecke <sup>3</sup>, Michéle Zoch <sup>3</sup>, Martin Sedlmayr <sup>3</sup>, Franziska Bathelt <sup>3</sup>

Affiliations – collapse

### Affiliations

- <sup>1</sup> Institute for Medical Informatics and Biometry at Carl Gustav Carus Faculty of Medicine at Technische Universität Dresden, Germany. Electronic address: yuan.peng@tu-dresden.de.
- <sup>2</sup> Institute for Medical Informatics and Biometry at Carl Gustav Carus Faculty of Medicine at Technische Universität Dresden, Germany. Electronic address: elisa.henke@tu-dresden.de.
- <sup>3</sup> Institute for Medical Informatics and Biometry at Carl Gustav Carus Faculty of Medicine at Technische Universität Dresden, Germany.

## Abstract

**Background:** International studies are increasingly needed in order to gain more unbiased evidence from real-world data. To achieve this goal across the European Union, the EMA set up the DARWIN EU project based on OMOP CDM established by the OHDSI community. The harmonization of heterogeneous local health data in OMOP CDM is an essential step to participate in such networks. Using the widespread communication standard HL7 FHIR can reduce the complexity of the transformation process to OMOP CDM. Enabling German university hospitals to participate in such networks requires an Extract, Transform and Load (ETL)-process that satisfies the following criteria: 1) transforming German patient data from FHIR to OMOP CDM, 2) processing huge amount of data at once and 3) flexibility to cope with changes in FHIR profiles.

**Method:** A mapping of German patient data from FHIR to OMOP CDM was accomplished, validated by an interdisciplinary team and checked through the OHDSI Data Quality Dashboard (DQD). To satisfy criteria 2-3, we decided to use SpringBatch-Framework according to its chunk-oriented design and reusable functions for processing large amounts of data.

**Results:** We have successfully developed an ETL-process that fulfills the defined criteria of transforming German patient data from FHIR into OMOP CDM. To measure the validity of the mapping conformance and performance of the ETL-process, it was tested with 392,022 FHIR resources. The ETL execution lasted approximately-one minute and the DQD result shows 99% conformance in OMOP CDM.

**Conclusions:** Our ETL-process has been successfully tested and integrated at 10 German university hospitals. The data harmonization utilizing international recognized standards like FHIR and OMOP fosters their ability to participate in international observational studies. Additionally, the ETL process can help to prepare more German hospitals with their data harmonization journey based on existing standards.

## 概要

背景 実データからより公平なエビデンスを得るために、国際的な研究がますます必要とされている。この目標を欧州連合全体で達成するために、EMA は OHDSI コミュニティによって確立された OMOP CDM に基づいて DARWIN EU プロジェクトを立ち上げた。このようなネットワークに参加するためには、OMOP CDMで異種の地域健康データを調和させることが不可欠なステップです。広く普及している通信規格である HL7 FHIR を使用することで、OMOP CDM への変換プロセスの複雑さを軽減することができます。ドイツの大学病院がこのようなネットワークに参加できるようにするには、以下の基準を満たす抽出、変換、読み込み (ETL) プロセスが必要です。1) ドイツの患者データをFHIRからOMOP CDMに変換する、2) 膨大なデータを一度に処理する、3) FHIRプロファイルの変更に対応できる柔軟性がある。

**方法:** ドイツの患者データをFHIRからOMOP CDMにマッピングし、学際的なチームによって検証され、OHDSI Data Quality Dashboard (DQD) を通じて確認された。基準2-3を満たすために、大量のデータを処理するためのチャンク指向の設計と再利用可能な関数に従って、SpringBatch-Frameworkを使用することにしました。

**結果:** ドイツの患者データをFHIRからOMOP CDMに変換するために定義された基準を満たすETLプロセスの開発に成功した。マッピングの適合性とETLプロセスのパフォーマンスを測定するため、392,022のFHIRリソースを用いてテストしました。ETLの実行時間は約1分で、DQDの結果はOMOP CDMに99%適合している。

**結論:** 我々の ETL プロセスは、ドイツの 10 の大学病院においてテストされ、統合に成功した。FHIRやOMOPのような国際的に認知された標準を利用したデータ調和は、国際的な観察研究への参加能力を促進させる。さらに、ETLプロセスは、既存の標準に基づくデータ調和の旅で、より多くのドイツの病院の準備に役立つことができます。





## PPIはアルツハイマー病のリスクか？

➤ Ther Adv Neurol Disord. 2022 Nov 8;15:17562864221135700. doi: 10.1177/17562864221135700.  
eCollection 2022.

# Long-term use of proton-pump inhibitor on Alzheimer's disease: a real-world distributed network analysis of six observational Korean databases using a Common Data Model

Yerim Kim <sup>1</sup>, Seung In Seo <sup>2 3</sup>, Kyung Joo Lee <sup>4</sup>, Jinseob Kim <sup>5</sup>, Jong Jin Yoo <sup>6</sup>, Won-Woo Seo <sup>7</sup>,  
Hyung Seok Lee <sup>8</sup>, Woon Geon Shin <sup>9 3</sup>

Affiliations + expand

PMID: 36389281 PMCID: PMC9647297 DOI: 10.1177/17562864221135700

## Abstract

**Background:** Dementia has a crucial impact on the quality of life of elderly patients and their caregivers. Proton-pump inhibitors (PPIs) are the most frequently prescribed treatment, but they have been shown to be associated with dementia. The data are inconsistent, however.

**Objective:** To investigate the association between PPIs use and Alzheimer's disease (AD) or all-cause dementia in six observational Korean databases using a Common Data Model (CDM) and to perform a distributed network analysis.

**Methods:** Subjects aged over 18 years between 1 January 2004 and 31 December 2020. Among 7,293,565 subjects from 6 cohorts, 41,670 patients met the eligibility criteria. A total of 2206 patients who were included in both cohorts or with a history of dementia were excluded. After propensity matching, 5699 propensity-matched pairs between the PPIs and histamine-2 receptor antagonist (H<sub>2</sub>RA) users were included in this study. The primary outcome was the incidence of AD at least 365 days after drug exposure. The secondary outcome was the incidence of all-cause dementia at least 365 days after drug exposure.

**Results:** In the 1:1 propensity score matching, the risk of AD or all-cause dementia was not significantly different between the PPIs and H<sub>2</sub>RA groups in all six databases. In the distributed network analysis, the long-term PPI users ( $\geq 365$  days) were unassociated with AD [hazard ratio (HR) = 0.92, 95% confidence interval (CI) = 0.68-1.23;  $I^2 = 0\%$ ] and all-cause dementia (HR = 1.04, 95% CI = 0.82-1.31;  $I^2 = 0\%$ ) compared with H<sub>2</sub>RA users.

**Conclusion:** In the distributed network analysis of six Korean hospital databases using Observational Medical Outcomes Partnership (OMOP)-CDM data, the long-term use of PPI was not associated with a statistically significantly increased risk of AD or all-cause dementia. Therefore, we suggest that physicians should not avoid these medications because of concern about dementia risk.

## 概要

**背景:** 認知症は、高齢の患者さんとその介護者のQOLに決定的な影響を与える。プロトンポンプ阻害剤（PPI）は最も頻繁に処方される治療薬であるが、認知症との関連が示されている。しかし、そのデータは一貫していない。

**目的:** 韓国の6つの観察データベースにおいて、PPIの使用とアルツハイマー病（AD）または全原因性認知症との関連を、共通データモデル（CDM）を用いて調査し、分散ネットワーク分析を行うことである。

**方法:** 2004年1月1日から2020年12月31日の間に18歳以上となった対象者。6つのコホートの7,293,565人の被験者のうち、41,670人が適格基準を満たした。両コホートに含まれる、または認知症の既往がある計2206名を除外した。傾向マッチングの結果、PPIとヒスタミン2受容体拮抗薬（H<sub>2</sub>RA）使用者の間で傾向マッチングされた5699組がこの研究に含まれた。主要アウトカムは、薬物曝露後少なくとも365日におけるADの発生率とした。副次的アウトカムは、薬物曝露から少なくとも365日後の全原因性認知症の発生率とした。

**結果:** 1:1の傾向スコアマッチングでは、6つのデータベースすべてにおいて、PPI群とH<sub>2</sub>RA群との間でADまたは全原因性認知症のリスクに有意差はなかった。分散型ネットワーク解析では、長期PPI使用者（365日）はH<sub>2</sub>RA使用者と比較して、AD [ハザード比（HR）= 0.92, 95%信頼区間（CI）= 0.68-1.23,  $I^2 = 0\%$ ] および全死亡認知症 [HR = 1.04, 95% CI = 0.82-1.31,  $I^2 = 0\%$ ] と関連しないことが明らかにされた。

**結論:** OMOP CDMデータを用いた韓国の6つの病院データベースの分散型ネットワーク解析では、PPIの長期使用は、ADおよび全原因性認知症のリスクを統計的に有意に増加させるとは関連しなかった。したがって、医師は認知症リスクを懸念してこれらの薬剤を避けるべきでないことが示唆された。



## 晩期うつ病患者の社会的・臨床的特徴

Observational Study > BMC Psychiatry. 2022 Nov 2;22(1):677. doi: 10.1186/s12888-022-04339-7.

### The sociodemographic characteristics and clinical features of the late-life depression patients: results from the Beijing Anding Hospital mental health big data platform

Xiao Wang <sup>1</sup>, Wenwang Rao <sup>2</sup>, Xueyan Chen <sup>1</sup>, Xinqiao Zhang <sup>1</sup>, Zeng Wang <sup>1</sup>, Xianglin Ma <sup>1</sup>, Qinge Zhang <sup>3</sup>

Affiliations – collapse

#### Affiliations

- 1 The National Clinical Research Center for Mental Disorders & Beijing Key Laboratory of Mental Disorders & Department of Psychiatry, Capital Medical University& Beijing Anding Hospital, Capital Medical University, 5 Ankang Lane, Dewai Avenue, Xicheng District, Beijing, China.
- 2 Unit of Psychiatry, Institute of Translational Medicine, Faculty of Health Sciences, University of Macau, Macao SAR, China.
- 3 The National Clinical Research Center for Mental Disorders & Beijing Key Laboratory of Mental Disorders & Department of Psychiatry, Capital Medical University& Beijing Anding Hospital, Capital Medical University, 5 Ankang Lane, Dewai Avenue, Xicheng District, Beijing, China. zqe81@126.com.

## Abstract

**Background:** The sociodemographic characteristics and clinical features of the Late-life depression (LLD) patients in psychiatric hospitals have not been thoroughly studied in China. This study aimed to explore the psychiatric outpatient attendance of LLD patients at a psychiatric hospital in China, with a subgroup analysis, such as with or without anxiety, gender differences.

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

**Results:** In a sample of 47,334 unipolar depression patients, 31,854 (67.30%) were women, and 15,480 (32.70%) were men. The main comorbidities of LDD are generalized anxiety disorder (GAD) (83.62%) and insomnia (74.52%). Among patients with unipolar depression, of which benzodiazepines accounted for the largest proportion (77.77%), Selective serotonin reuptake inhibitors (SSRIs) accounted for 59.00%, a noradrenergic and specific serotonergic antidepressant (NaSSAs) accounted for 36.20%. The average cost of each visit was approximately 646.27 yuan, and the cost of each visit was primarily attributed to Western medicine (22.97%) and Chinese herbal medicine (19.38%). For the cost of outpatient visits, depression comorbid anxiety group had a higher average cost than the non-anxiety group ( $p < 0.05$ ). There are gender differences in outpatient costs, men spend more than women, for western medicine, men spend more than women, for Chinese herbal medicine, women spend more than men (all  $p < 0.05$ ). The utilization rate of SSRIs and benzodiazepines in female patients is significantly higher than that in male patients ( $p < 0.05$ ).

**Conclusion:** LLD patients are more commonly women than men and more commonly used SSRIs and NaSSAs. Elderly patients with depression often have comorbid generalized anxiety. LLD patients spend most of their visits on medicines, and while the examination costs are lower.

## 概要

**背景:** 精神科病院におけるLate-life depression (LLD) 患者の社会人口統計学的特徴や臨床的特徴については、中国では十分に研究されていない。本研究は、中国の精神科病院におけるLLD患者の精神科外来通院状況を、不安の有無、性差などのサブグループ分析を行いながら調査することを目的とした。

**方法:** 本研究は、北京安定医院のOMOP-CDMのデータを用いて、2013年1月から2019年8月のLLDの外来患者を調べた後方視的研究である。年齢、性別、受診回数、薬剤の使用状況、併存する疾患は診療記録から抽出した。

**結果:** 単極性うつ病患者47,334人のうち、31,854人(67.30%)が女性で、15,480人(32.70%)が男性であった。LDDの主な併存疾患は全般性不安障害(GAD)(83.62%)、不眠症(74.52%)。単極性うつ病患者のうち、ベンゾジアゼピン系が最も多く(77.77%)、選択的セロトニン再取込阻害薬(SSRI)が59.00%、ノルアドレナリン作動性・特異的セロトニン作動性抗うつ薬(NaSSA)が36.20%となり、1回の診察の平均費用は約646.27円で、その主な内訳は西洋医学(22.97%)、漢方薬(19.38%)であった。外来受診費用については、うつ病併発不安症群は非不安症群より平均費用が高かった( $p < 0.05$ )。外来診療費には性差があり、男性は女性より多く、西洋医学では男性が女性より多く、漢方薬では女性が男性より多くかかった(いずれも $p < 0.05$ )。SSRIとベンゾジアゼピン系薬剤の使用率は、女性患者の方が男性患者より有意に高い( $p < 0.05$ )。

**結論:** LLD患者は男性よりも女性に多く、SSRIやNaSSAをよく使用していた。高齢のうつ病患者には全般性不安が併存していることが多い。LLD患者は受診料のほとんどを薬代に費やしており、一方で診察料は安い。





## 全英 COVID-19 データネットワーク

> J Med Internet Res. 2022 Nov 1. doi: 10.2196/40035. Online ahead of print.

### CO-CONNECT: A hybrid architecture to facilitate rapid discovery and access to UK wide data in the response to the COVID-19 pandemic

Emily Jefferson <sup>1</sup>, Christian Cole <sup>1</sup>, Shahzad Mumtaz <sup>1</sup>, Sam Cox <sup>2</sup>, Tom Giles <sup>2</sup>, Samuel Adejumo <sup>2</sup>, Esmond Urwin <sup>2</sup>, Daniel Lea <sup>2</sup>, Calum McDonald <sup>3</sup>, Joseph Best <sup>2 4</sup>, Erum Masood <sup>1</sup>, Gordon Milligan <sup>1</sup>, Jenny Johnston <sup>1</sup>, Scott Horban <sup>1</sup>, Ipek Birced <sup>1</sup>, Christopher Hall <sup>1</sup>, Aaron Jackson <sup>1</sup>, Clare Collins <sup>2</sup>, Sam Rising <sup>2</sup>, Charlotte Dodsley <sup>2</sup>, Jill Hampton <sup>1</sup>, Andrew Hadfield <sup>2</sup>, Roberto Santos <sup>2</sup>, Simon Tarr <sup>2</sup>, Vasiliki Panagi <sup>2</sup>, Joseph Lavagna <sup>2</sup>, Tracy Jackson <sup>3</sup>, Antony Chuter <sup>5</sup>, Jillian Beggs <sup>1</sup>, Magdalena Martinez-Queipo <sup>6</sup>, Helen Ward <sup>7</sup>, Julie von Ziegenweidt <sup>8 9</sup>, Frances Burns <sup>10</sup>, Jo Martin <sup>11</sup>, Neil Sebire <sup>12</sup>, Carole Morris <sup>13</sup>, Declan Bradley <sup>14 15</sup>, Rob Baxter <sup>16</sup>, Anni Ahonen-Bishop <sup>17</sup>, Amelia Shoemark <sup>18</sup>, Ana Valdes <sup>19</sup>, Benjamin J Ollivere <sup>19</sup>, Charlotte Manisty <sup>20</sup>, David William Eyre <sup>21</sup>, Stephanie Gallant <sup>18</sup>, George Joy <sup>22</sup>, Andrew McAuley <sup>23</sup>, David W Connell <sup>24</sup>, Kate Northstone <sup>25</sup>, Katie Jm Jeffery <sup>26 27</sup>, Emanuele Di Angelantonio <sup>28 29 30 31 32</sup>, Amy McMahon <sup>28 30</sup>, Matthew Walker <sup>28 30</sup>, Malcolm Gracie Semple <sup>33 34</sup>, Jessica Mai Sims <sup>35</sup>, Emma Lawrence <sup>36</sup>, Bethan Davies <sup>7</sup>, J Kenneth Baillie <sup>37</sup>, Ming Tang <sup>38</sup>, Gary Leeming <sup>39</sup>, Linda Power <sup>40</sup>, Thomas Breeze <sup>41</sup>, Natalie Gilson <sup>42</sup>, Duncan J Murray <sup>43 44</sup>, Chris Orton <sup>45</sup>, Iain Pierce <sup>46 47</sup>, Ian Hall <sup>48</sup>, Shamez Ladhani <sup>49</sup>, Matthew Whitaker <sup>7</sup>, Laura Shallcross <sup>50</sup>, David Seymour <sup>4</sup>, Susheel Varma <sup>4</sup>, Gerry Reilly <sup>4</sup>, Andrew Morris <sup>4</sup>, Susan Hopkins <sup>40</sup>, Aziz Sheikh <sup>51</sup>, Philip Quinlan <sup>2 19</sup>

Affiliations + expand

PMID: 36322788 DOI: 10.2196/40035

## Abstract

**Background:** COVID-19 data have been generated across the UK as a by-product of clinical care and public health provision, and numerous bespoke and repurposed research endeavours. Analysis of these data has underpinned the UK's response to the pandemic and informed public health policies and clinical guidelines. However, these data are held by different organisations and this fragmented landscape has presented challenges for public health agencies and researchers as they struggle to find, navigate permissions to access and interrogate the data they need to inform the pandemic response at pace.

**Objective:** To transform UK COVID-19 diagnostic datasets to be Findable, Accessible, Interoperable and Reusable (FAIR).

**Methods:** A federated infrastructure model was rapidly built to enable the automated and reproducible mapping of health Data Partners' pseudonymised data to the OMOP common data model without the need for any data to leave the data controllers' secure environments and to support federated cohort discovery queries and meta-analysis.

**Results:** 56 datasets from 19 organisations are being connected to the federated network. The data includes research cohorts and COVID-19 data collected through routine health care provision linked to longitudinal healthcare records and demographics. The infrastructure is live, supporting aggregate level querying of data across the UK.

**Conclusions:** CO-CONNECT was developed by a multidisciplinary team enabling rapid COVID-19 data discovery, instantaneous meta-analysis across data sources, and is researching streamlined data extraction for egress into a Trusted Research Environment (TRE) for research and public health analysis. CO-CONNECT has the potential to make UK health data more interconnected and better able to answer national-level research questions whilst maintaining patient confidentiality and local governance procedures.

## 概要

**背景:** COVID-19データは、臨床ケアや公衆衛生提供の副産物として、また多くの特注や再利用の研究努力によって英国全土で生成されてきた。これらのデータの解析は、パンデミックに対する英国の対応を支え、公衆衛生政策や臨床ガイドラインに情報を提供してきた。しかし、これらのデータは様々な組織によって保有されており、この断片的な状況は、公衆衛生機関や研究者がパンデミックへの対応を迅速に伝えるために必要なデータを見つけ、アクセス許可を得て、調査するのに苦労しているという課題を示している。

**目的:** 英国のCOVID-19診断データセットをFAIR (Findable, Accessible, Interoperable and Reusable) に変換することである。

**方法:** データ管理者の安全な環境からデータを持ち出すことなく、医療データパートナーの仮名化データを OMOP 共通データモデルに自動的かつ再現可能にマッピングし、コホート発見クエリーとメタアナリシスの統合をサポートする統合インフラモデルを迅速に構築した。

**結果:** 19の組織から56のデータセットが連携ネットワークに接続されている。データには研究コホートと、縦断的な医療記録や人口統計にリンクした日常的な医療提供によって収集されたCOVID-19データが含まれる。インフラは稼働しており、英国内のデータの総計レベルのクエリーをサポートしている。

**結論:** CO-CONNECTは学際的なチームによって開発され、迅速なCOVID-19データの発見、データソース間の即時メタ分析を可能にし、研究および公衆衛生分析のために信頼できる研究環境 (TRE) への進出のための合理的なデータ抽出を研究途中である。CO-CONNECTは、英国の健康データをより相互接続し、患者の機密性と地域の統治手続きを維持しながら、国家レベルの研究課題に答えることができるようにする可能性を持っています。





## 妊娠週数を抽出するアルゴリズム開発

> PLoS One. 2022 Oct 31;17(10):e0276923. doi: 10.1371/journal.pone.0276923. eCollection 2022.

# Temporal Events Detector for Pregnancy Care (TED-PC): A rule-based algorithm to infer gestational age and delivery date from electronic health records of pregnant women with and without COVID-19

Tianchu Lyu <sup>1</sup>, Chen Liang <sup>1</sup>, Jihong Liu <sup>2</sup>, Berry Campbell <sup>3</sup>, Peiyin Hung <sup>1</sup>, Yi-Wen Shih <sup>1</sup>, Nadia Ghumman <sup>1</sup>, Xiaoming Li <sup>4</sup>, National COVID Cohort Collaborative Consortium

Affiliations – collapse

### Affiliations

- 1 Department of Health Services Policy and Management, Arnold School of Public Health, University of South Carolina, Columbia, South Carolina, United States of America.
- 2 Department of Epidemiology & Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, South Carolina, United States of America.
- 3 Department of Obstetrics and Gynecology, School of Medicine, University of South Carolina, Columbia, South Carolina, United States of America.
- 4 Department of Health Promotion Education and Behaviors, Arnold School of Public Health, University of South Carolina, Columbia, South Carolina, United States of America.

## Abstract

**Objective:** Identifying the time of SARS-CoV-2 viral infection relative to specific gestational weeks is critical for delineating the role of viral infection timing in adverse pregnancy outcomes. However, this task is difficult when it comes to Electronic Health Records (EHR). In combating the COVID-19 pandemic for maternal health, we sought to develop and validate a clinical information extraction algorithm to detect the time of clinical events relative to gestational weeks.

**Materials and methods:** We used EHR from the National COVID Cohort Collaborative (N3C), in which the EHR are normalized by the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM). We performed EHR phenotyping, resulting in 270,897 pregnant women (June 1st, 2018 to May 31st, 2021). We developed a rule-based algorithm and performed a multi-level evaluation to test content validity and clinical validity, and extreme length of gestation (<150 or >300).

**Results:** The algorithm identified 296,194 pregnancies (16,659 COVID-19, 174,744 without COVID-19) in 270,897 pregnant women. For inferring gestational age, 95% cases ( $n = 40$ ) have moderate-high accuracy (Cohen's Kappa = 0.62); 100% cases ( $n = 40$ ) have moderate-high granularity of temporal information (Cohen's Kappa = 1). For inferring delivery dates, the accuracy is 100% (Cohen's Kappa = 1). The accuracy of gestational age detection for the extreme length of gestation is 93.3% (Cohen's Kappa = 1). Mothers with COVID-19 showed higher prevalence in obesity or overweight (35.1% vs. 29.5%), diabetes (17.8% vs. 17.0%), chronic obstructive pulmonary disease (0.2% vs. 0.1%), respiratory distress syndrome or acute respiratory failure (1.8% vs. 0.2%).

**Discussion:** We explored the characteristics of pregnant women by different gestational weeks of SARS-CoV-2 infection with our algorithm. TED-PC is the first to infer the exact gestational week linked with every clinical event from EHR and detect the timing of SARS-CoV-2 infection in pregnant women.

**Conclusion:** The algorithm shows excellent clinical validity in inferring gestational age and delivery dates, which supports multiple EHR cohorts on N3C studying the impact of COVID-19 on pregnancy.

## 概要

**目的:** SARS-CoV-2ウイルスの感染時期と特定の妊娠週数との関係を明らかにすることは、ウイルス感染時期が妊娠の有害事象に及ぼす役割を明らかにするために重要である。しかし、この作業は電子健康記録（EHR）に関しては困難である。母体の健康のためのCOVID-19のパンデミックに対抗するために、我々は妊娠週数に対する臨床イベントの時期を検出する臨床情報抽出アルゴリズムを開発し、検証することを目指した。

**材料と方法:** 全米COVIDコホート共同体（N3C）のEHRを使用し、EHRはOMOP-CDMにより正規化されている。EHRフェノタイピングを実施し、270,897人の妊婦を得た（2018年6月1日～2021年5月31日）。ルールベースのアルゴリズムを開発し、内容の妥当性と臨床的妥当性、極端な妊娠期間（150未満または300以上）を検証するために多段階評価を行った。

**結果:** アルゴリズムは270,897人の妊婦の296,194件の妊娠（COVID-19 16,659件、COVID-19なし174,744件）を同定した。妊娠週数の推定では、95%（ $n = 40$ ）が中程度の高い精度（CohenのKappa = 0.62）を有し、100%（ $n = 40$ ）が時間情報の粒度が中程度の高い（CohenのKappa = 1）ものであった。分娩日の推定については、精度は100%（CohenのKappa = 1）であった。極端な妊娠期間に対する妊娠週数検出の精度は93.3%である（CohenのKappa = 1）。COVID-19を持つ母親は、肥満または過体重（35.1% vs 29.5%）、糖尿病（17.8% vs 17.0%）、慢性閉塞性肺疾患（0.2% vs 0.1%）、呼吸困難症候群または急性呼吸不全（1.8% vs 0.2%）において高い有病率を示した。

**考察:** 我々は、SARS-CoV-2感染の異なる妊娠週数による妊婦の特徴をアルゴリズムで探索した。TED-PCは、EHRからすべての臨床イベントとリンクした正確な妊娠週数を推論し、妊婦のSARS-CoV-2感染時期を検出する初めての方法である。

**結論:** このアルゴリズムは、妊娠週数と出産日の推定において優れた臨床的妥当性を示しており、COVID-19の妊娠への影響を研究するN3Cの複数のEHRコホートを裏付けている。



## CDMメタデータの横断検討

➤ [Sci Data](#). 2022 Oct 28;9(1):659. doi: 10.1038/s41597-022-01792-7.

# Harvesting metadata in clinical care: a crosswalk between FHIR, OMOP, CDISC and openEHR metadata

Caroline Bönisch <sup>1</sup>, Dorothea Kesztyüs <sup>2</sup>, Tibor Kesztyüs <sup>2</sup>

Affiliations – collapse

### Affiliations

- 1 Medical Data Integration Center, Department of Medical Informatics, University Medical Center Göttingen, Robert-Koch-Str. 40, 37075, Göttingen, Germany. [caroline.boenisch@med.uni-goettingen.de](mailto:caroline.boenisch@med.uni-goettingen.de).
- 2 Medical Data Integration Center, Department of Medical Informatics, University Medical Center Göttingen, Robert-Koch-Str. 40, 37075, Göttingen, Germany.

PMID: 36307424 PMCID: [PMC9616884](#) DOI: [10.1038/s41597-022-01792-7](#)

## Abstract

Metadata describe information about data source, type of creation, structure, status and semantics and are prerequisite for preservation and reuse of medical data. To overcome the hurdle of disparate data sources and repositories with heterogeneous data formats a metadata crosswalk was initiated, based on existing standards. FAIR Principles were included, as well as data format specifications. The metadata crosswalk is the foundation of data provision between a Medical Data Integration Center (MeDIC) and researchers, providing a selection of metadata information for research design and requests. Based on the crosswalk, metadata items were prioritized and categorized to demonstrate that not one single predefined standard meets all requirements of a MeDIC and only a maximum data set of metadata is suitable for use. The development of a convergence format including the maximum data set is the anticipated solution for an automated transformation of metadata in a MeDIC.

## 概要

メタデータは、データソース、作成タイプ、構造、ステータス、セマンティクスに関する情報を記述し、医療データの保存と再利用のための前提条件となるものです。異種データ形式を持つ異種データソースやリポジトリのハードルを克服するために、既存の標準に基づくメタデータのクロスウォークが開始されました。データ形式の仕様と同様に、FAIR原則も含まれています。

メタデータ・クロスウォークは、医療データ統合センター（MeDIC）と研究者の間のデータ提供の基盤となり、研究設計やリクエストのためのメタデータ情報のセレクションを提供するものです。クロスウォークに基づき、メタデータ項目の優先順位付けと分類を行い、一つの定義済み規格がMeDICの全ての要求を満たすわけではなく、最大限のデータセットであるメタデータのみが使用に適していることを実証しました。最大限のデータセットを含むコンバージェンスフォーマットの開発は、MeDICにおけるメタデータの自動変換のためのソリューションとして期待されるものです。

対象：OMOP、openEHR、FHIR、CDISC



# 今月のCommunity Call

## ● APAC Community Call テーマ

Nov. 3 2022 APAC Study Quarterly Updates

- Characterization of health by OHDSI AP chapter to identify temporal effect of the pandemic (CHAPTER) by Seng Chan You
- Comparison of mortality, morbidities & healthcare resources utilisation between patients with and without a diagnosis of COVID-19 by Ivan Chun Hang Lam
- Data quality of OHDSI APAC: CDM inspection study by Chungsoo Kim  
APAC Study Quarterly Updates

## ● Global Community Call テーマ

Nov. 1 Meet the Titans

Nov. 8 Collaborator Showcase Presentations

Nov. 15 Network Studies + ARES Demo

Nov. 22 10-Minute Tutorials

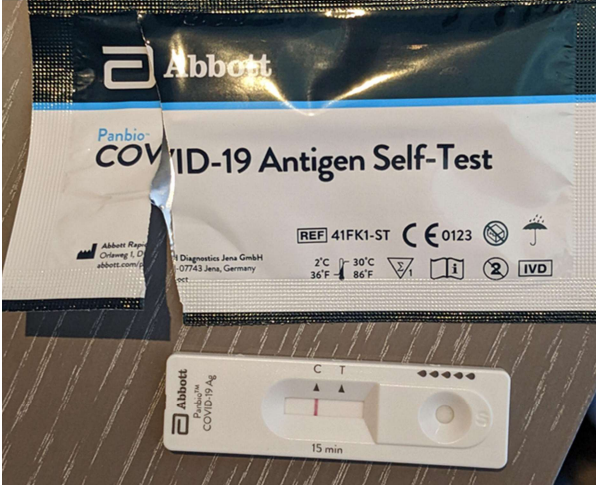
PHOEBE 2.0, Automated Comparator Selection, Strategus, Einstein-ATLAS, Broadsea





# 2022 APAC OHDSI Symposium

Nov. 12 - 13 • Taipei Medical University 臺北醫學大學







# Preliminary Substudy Analysis of the Characterization of Health by OHDSI Asia-Pacific chapter to identify Temporal Effect of the Pandemic (CHAPTER) study focusing on hematologic diseases

Eri Matsuki<sup>1,2</sup>, Jing Li<sup>3</sup>, Can Yin<sup>3</sup>, Mui Van Zandt<sup>3</sup>, Seng Chan You<sup>4,5</sup>

1 Division of Hematology, Department of Medicine, Keio University School of Medicine, Tokyo, Japan; 2 Clinical and Translational Research Center, Keio University Hospital, Tokyo, Japan; 3 IQVIA, Durham, NC, USA; 4 Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, South Korea; 5 Institute for Innovation in Digital Healthcare, Yonsei University, Seoul, South Korea

## Background

- A global pandemic of coronavirus disease of 2019 (COVID-19) has affected more than millions of lives worldwide since the first report of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in December, 2019.
- The rapidly changing nature of the disease including the different types of virus strain causing the pandemic, the symptoms and infectious rate, as well as change in governmental policies, it has been difficult to recapitulate the full impact of SARS-CoV-2 infection to date.
- Especially, its effect on healthcare resource utilization not limited to the treatment of COVID-19 but also its impact on non-communicable diseases has not been well elucidated<sup>1</sup>.
- Recently, studies performed on a federated network have been shown to provide a broader picture on how the healthcare systems across different regions and countries have been utilized during this pandemic.
- Previous studies have shown that data assets mapped to the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) provide a unique opportunity to make a difference in the current crisis, allowing for robust analyses to be performed in a timely across a network of sites<sup>2</sup>.
- The OHDSI Asian Pacific regional chapter has launched the Characterization of Health by OHDSI Asia-Pacific chapter to identify Temporal Effect of the Pandemic (CHAPTER) study to describe the temporal change in the incidence of diseases and healthcare patterns before and after the emergence of COVID-19.
- Multiple myeloma (MM) is one of the most common hematologic malignancies, with a reported annual incidence of 114,252 new cases and 80,119 deaths worldwide in 2012, comprising 0.8% and 1% of all cancers, respectively.
- Its incidence is known to vary by ethnicity, with Asians showing a relatively lower incidence than Caucasians, however, recently, several reports have shown that the incidence of MM is increasing in Asian countries, including Korea, Taiwan, and Thailand.
- It has a unique clinical course of starting from a preceding pre-malignant state called monoclonal gammopathy of undetermined significance (MGUS) to an indolent disease state with no symptoms to a more aggressive disease requiring long-term treatment with chemotherapeutic agents.
- In this study, we assessed the impact of COVID-19 on the diagnostic pattern of this disease in the Asia-Pacific Region.

## Methods

- Databases from the Australia LPD and Japan Claims converted to the OMOP-CDM were adapted as our data source in this retrospective study.
- For the definition of MM, we leveraged the digital phenotype definition from the previous Phenotype Phebruary project initiated by the Observational Health Data Science and Informatics (OHDSI) community<sup>3</sup>.
- We employed interrupted time series analysis to describe the incidence trend of MM before and after the COVID-19 pandemic.

## References

1. Patti KG, Kohli P. COVID's impact on non-communicable diseases: What we do not know may hurt us. *Curr Cardiol Rep.* 2022 May 7, Online ahead of print.
2. Burn E, You SC, Sena AG, Kostka K, Abedtash H, Abrahão MTF, et al. Deep phenotyping of 34,128 adult patients hospitalised with COVID-19 in an international network study. *Nat Commun.* 2020 Oct 6;11(1):5009.
3. Patrick Ryan, et al. "Phenotype Phebruary Day 4 – Multiple Myeloma." OHDSI Forums; 5 Feb. 2022. <https://forums.ohdsi.org/t/phenotype-phebruary-day-4-multiple-myeloma/15801>.
4. Greene G, Griffiths R, Han J, Akbari A, Jones M, Lyons J, et al. Impact of the SARS-CoV-2 pandemic on female breast, colorectal and non-small cell lung cancer incidence, stage and healthcare pathway to diagnosis during 2020 in Wales, UK, using a national cancer clinical record system. *Br J Cancer.* 2022 May 2, Online ahead of print.

## Results

- In this preliminary analysis, both the Australia LPD (Fig 1A) and Japan Claims (Fig 1B) showed a decrease in the incidence of MM in 2020, compared to both the incidence in the previous year as well as what could be expected from the trend in previous years.
- MM is a malignant disease that can be slowly progressive, starting from a smoldering disease that does not require treatment to symptomatic disease with possible bone fracture, renal dysfunction, hypercalcemia, or increased susceptibility to infectious diseases. *Our current result may reflect the reduced number of diagnoses due to asymptomatic patients avoiding hospital visits during the COVID-19 pandemic.*
- In terms of treatment, once treatment is required, many patients would require continuous treatment that could also increase the patients' susceptibility to infection. While our current phenotypic description of MM only uses diagnostic code for MM definition, other types of cohort definition established during the Phenotype Phebruary project include MM-specific treatments. Therefore, it will be of interest to compare the change in the incidence rate based on different cohort definitions to figure out whether the decrease of multiple myeloma incidence is purely due to decreased number of newly diagnosed cases versus whether there had been any delayed treatment under the concern of COVID-19 infection as well.

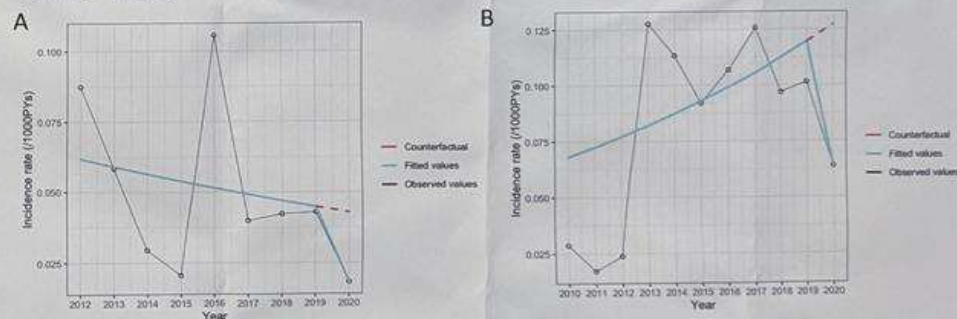


Figure 1. Change in the incidence rate before and after COVID-19 pandemic. The red dotted line of counterfactual shows the expected incidence rate derived from the data in pre-pandemic era. The fitted values shown in green solid line are estimated based on the Poisson regression model with adjusting time vector (years).

## Conclusions

Our current result suggests that the incidence of MM in the Asia-Pacific region changed with the COVID-19 pandemic. Further investigation of the CHAPTER study group will provide more concrete evidence of the true incidence of MM and other hematologic diseases and provide more scientific relevant and detailed information across the OHDSI network. Efforts to refine the cohort definition of different hematologic diseases and recruit data partners to join this study are ongoing.





# APAC Symposium プログラム

Day 2

## Welcome Session

### **Session 1: Envisioning of OHDSI Global & EU**

Keynote - OHDSI Global Presentation

DARWIN EU

### **Session 2: The Challenges of Research in OHDSI APAC**

OHDSI APAC Introduction

Research in OHDSI APAC (10min each)

Research using Taiwan National Data

Research using TMUCRD Data

### **Session 3: The Implication Experiences in OHDSI Region**

Panel – Standardization & Common Data Models (CDM)

OMOP, FHIR, HADES

Panel - APAC Regional Adaption to Standardization

Taiwan, Australia, China, Japan, Korea, Singapore

## **Poster Session/Network Session**

## **Closing Remarks**



# 各国の状況

## ● 台湾



## ● ほかの国

- 韓国
- シンガポール
- オーストラリア



# 来年以降の APAC Symposium

## ■ 2023 Sydney, Australia



Medinfo 2023: 7月8日～12日

APAC Symposium: **直後** 7月13日～14日（仮り決め日程）

## ■ 2024 Singapore