

Oxford Summer School参加経験の共有

OHDSI Japan イブニングカンファレンス / 2024年12月27日

中外製薬株式会社

セーフティサイエンス第一部DSG 織田

免責事項

- 発表者は中外製薬株式会社に社員として所属しています。
- 本日の発表に示された意見や考えは発表者の個人的なものであり、その所属企業の公式な見解ではありません。

自己紹介

- 織田哲朗
- 中外製薬株式会社セーフティサイエンス第一部データサイエンスグループ
 - 安全性のテーマにおける，リアルワールドデータを用いた研究立案・実施・公開

アジェンダ

01

セミナー概要

02

EHDENについて

03

社内で調べたこと

Oxford Summer School 2024

*Real world evidence using the
OMOP Common Data Model*

17 - 21 June 2024

Registration is now closed



Oxford Summer Schoolは
様々なトピックで開講さ
れており、その中の一つ

Brief Description:

- RWD研究の計画・実行に必要な概念やツールをOMOP common data modelを中心に学ぶ。

Venue: Lady Margaret Hall, Norham Gardens, Oxford

AUDIENCE:

- Pharmacists, clinicians, academics, Industry (pharmacy or device) etc.

FACULTY:

- OHDSI UKをリードするBotnar Research Centre (University of Oxford)のスタッフ
- DARWIN EU関連の解析ツールの開発者
- Target trial emulationの講演者であるAssistant Prof Shirley Wang (Harvard Pilgrim)

場所



2025年はまた違う場所のようです。

スケジュール

- 午前が講義，午後が実習
- 実習はEpidemiologyとAnalyticsの2チームに分かれた。以下はAnalyticsトラックのスケジュール。

月

- 9.00h - Welcome to RWE: why do we care?
- 9.30-10.30h - Real World Data: Primary care, Hospital EHR, Claims data
 - (coffee break)
- 11.00-11.45h - The OMOP CDM
- 11.45-12.15h - Large OMOP initiatives: OHDSI, EH DEN, DARWIN EU®
- 12.15-12.30h - Theme for the week

火

- 9.15-10.00h - Study designs: cohorts and case-controls
- 10.00-10.30h - Study designs: self-controlled analyses
 - (coffee break)
- 11.00-11.30h - Phenotyping step by step
- 11.30-12.00h - Phenotyping condition-based cohorts in OMOP
- 12.00-12.30h - Phenotyping drug-based cohorts in OMOP

水

- 9.00-9.45h - Introduction to Machine Learning
- 9.45-10.30h - Introduction to Cluster Analyses
 - (coffee break)
- 11.00-11.45h - Descriptive epidemiology: incidence and prevalence
- 11.45-12.30h - Introduction to survival analysis

木

- 9.15-9.45h- Drug utilisation research: the very basics
- 9.45-10.30h- New (drug) user cohorts
- (coffee break)
- 11.00-11.15h- Confounding by indication
- 11.15-12.00h- Propensity scores, disease risk scores
- 12.00-12.30h- Negative control outcomes

金

- 9.30-10.30h- Crash intro to Health Economics
 - (coffee break)
- 11.00-11.45h- Target trial emulation [external speaker, Prof Shirley Wang]
- 11.45-12.30h- Results from DUPLICATE [external speaker, Prof Shirley Wang]

実習の内容

Day 1

Welcome

<https://dpa-pde-oxford.quarto.pub/welcome/>

OMOP and CDMConnector

<https://dpa-pde-oxford.quarto.pub/omop-and-cdmconnector/>

Bespoke programming

<https://dpa-pde-oxford.quarto.pub/bespoke-code/>

OMOPSketch and database characterisation

<https://dpa-pde-oxford.quarto.pub/omopsketch/>

Day 2

CodelistGenerator

<https://dpa-pde-oxford.quarto.pub/codelistgenerator/>

CohortConstructor

<https://dpa-pde-oxford.quarto.pub/cohortconstructors24/>

実習の内容

Day 3

CohortCharacteristics Part 1

<https://dpa-pde-oxford.quarto.pub/cohortcharacteristics-part1>

PatientProfiles

<https://dpa-pde-oxford.quarto.pub/patientprofiles-summer-school>

CohortCharacteristics Part 2

<https://dpa-pde-oxford.quarto.pub/summarisecharacteristics-summer-school/>

IncidencePrevalence

<https://dpa-pde-oxford.quarto.pub/incidenceprevalence/>

Day 4

DrugUtilisation

<https://dpa-pde-oxford.quarto.pub/drugutilisation>

CohortSurvival

<https://dpa-pde-oxford.quarto.pub/cohortsurvival/>

実習の内容

- Githubにすべて公開されている。
- Epiの方では実習は2日しかない。
- 使用するOMOP関連のパッケージはDARWIN-EUとしてOxfordの教室で作成されたもので、OHDSIコミュニティで作成されているパッケージ群とは異なる。
- {HADES}を使ったことがないので比較はできないが、文法は特殊だが慣れてしまえばかなり楽に定義整理や解析を実施できる印象。
- 基本的にはどのパッケージもちゃんとバリデートされているらしい。例えば{IncidencePrevalence}というパッケージに関しては[Paper](#)にどのような検証をしたかの記載がある。

{PatientProfiles}の例

addDemographics()

```
cdm$sinusitis |>  
  addDemographics()
```

```
# Source:   table<og_018_1718807112> [?? x 8]  
# Database: DuckDB v0.10.0 [martics@Windows 10 x64:R 4.2.3/C:\Users\martics\AppData\Local\Temp\RtmpaC46NE\file9ac843d33dbf.duckdb]  
  cohort_definition_id subject_id cohort_start_date cohort_end_date  age sex  prior_observation future_observation  
                <int>      <int> <date>          <date>          <int> <chr>          <int>          <int>  
1                1        280 1995-11-27      1995-11-27      57 Female        20849         8590  
2                1        393 1961-11-20      1961-11-20       4 Female        1776         20869  
3                1       2156 1994-04-03      1994-04-03      29 Female       10846         9118  
4                1       2392 1932-08-18      1932-08-18      16 Female        6040        31618  
5                1       3761 1980-08-27      1980-08-27      24 Female        8894        13821  
6                1       4764 2005-06-11      2005-06-11      36 Male         13443         4827  
7                2         97 1952-12-28      1952-12-28       9 Female        3450        23983  
8                2        163 2005-12-17      2005-12-17      39 Female       14417         4588  
9                2        183 1998-11-01      1998-11-01      24 Male         8774         7440  
10               2        338 1974-09-29      1974-09-29       7 Female        2720        16285  
# i more rows
```

{PatientProfiles}の例

```
cdm <- generateConceptCohortSet(  
  cdm = cdm,  
  name = "bronchitis",  
  conceptSet = list(bronchitis = c(258780, 260139)),  
  end = 0,  
  limit = "all"  
)
```

```
cdm$bronchitis <- cdm$bronchitis |>  
  addSex() |>  
  filter(sex == "Male") |>  
  select(-"sex") |>  
  compute(name = "bronchitis", temporary = FALSE) |>  
  recordCohortAttrition("Just Male") |>  
  addAge() |>  
  filter(age >= 25) |>  
  select(-"age") |>  
  compute(name = "bronchitis", temporary = FALSE) |>
```

A tibble: 4 × 7

	cohort_definition_id	number_records	number_subjects	reason_id	reason	excluded_records	excluded_subjects
	<int>	<int>	<int>	<int>	<chr>	<int>	<int>
1	1	8232	2546	1	Initial qualifying events	0	0
2	1	4115	1245	2	Just Male	4117	1301
3	1	2235	1039	3	25 or above	1880	206
4	1	2235	1039	4	At least 365 days of prior obs...	0	0

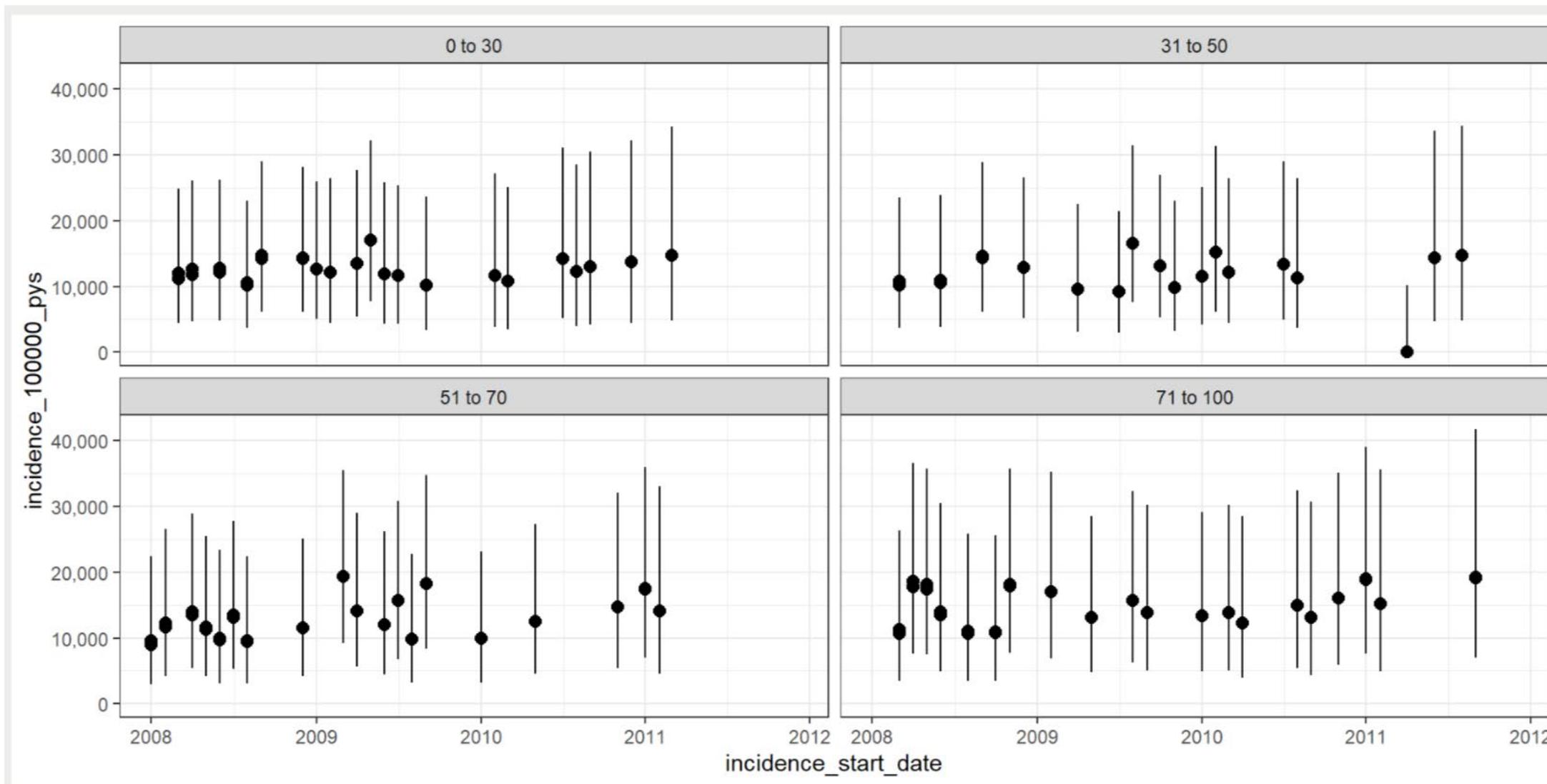
{IncidencePrevalence}の例

estimateIncidence()

```
inc <- estimateIncidence(  
  cdm = cdm,  
  denominatorTable = "denominator",  
  outcomeTable = "outcome",  
  interval = c("Months"),  
  outcomeWashout = c(0, 365),  
  repeatedEvents = FALSE  
)
```

```
plot <- plotIncidence(inc,  
  facet = "denominator_age_group")
```

{IncidencePrevalence}の例



アジェンダ

01

セミナー概要

02

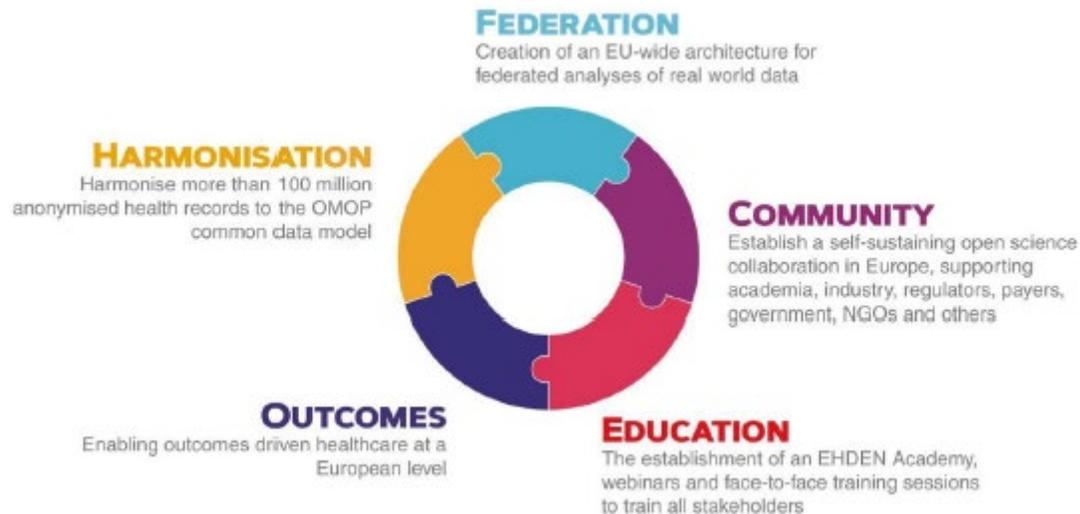
EHDENについて

03

社内で調べたこと

The European Health Data & Evidence Network (EHDEN)

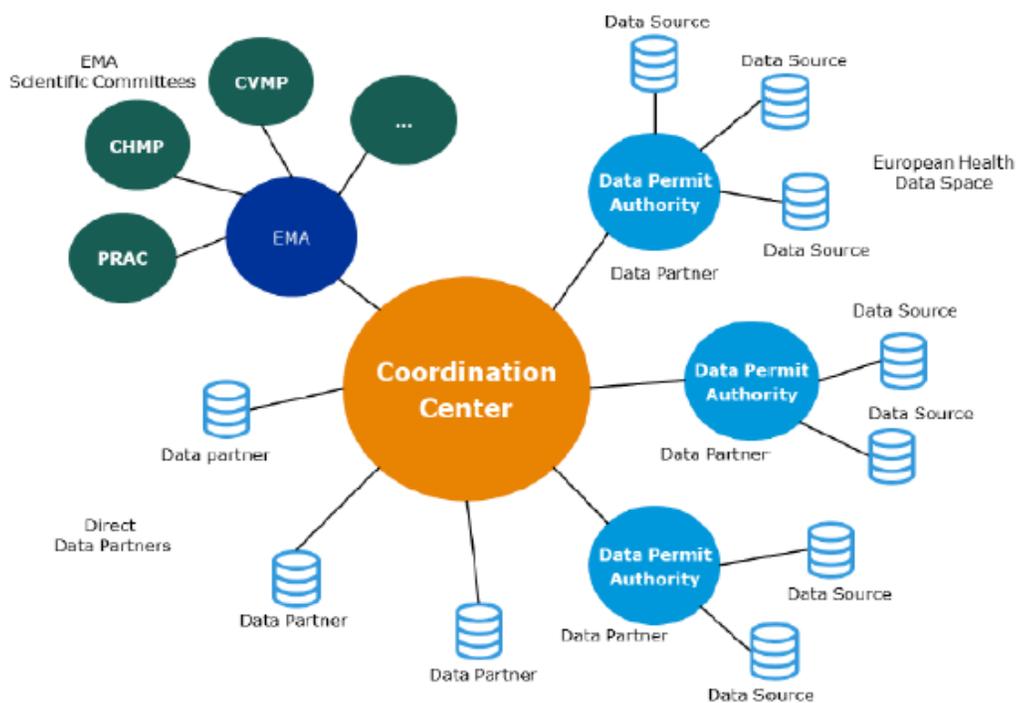
Providing a new paradigm for the discovery and analysis of health data in Europe by building a large-scale, federated network of data sources standardised to a common data model (OMOP), significantly speeding up the generation of reliable evidence.



Geographic spread of data partners. The shade of blue indicates the # of data partners in that country (darker = more)

DARWIN EU

DARWIN EU® is a federated **network** of **data**, **expertise** and **services** that supports better decision-making throughout the product lifecycle by generating reliable **evidence from real world healthcare data**

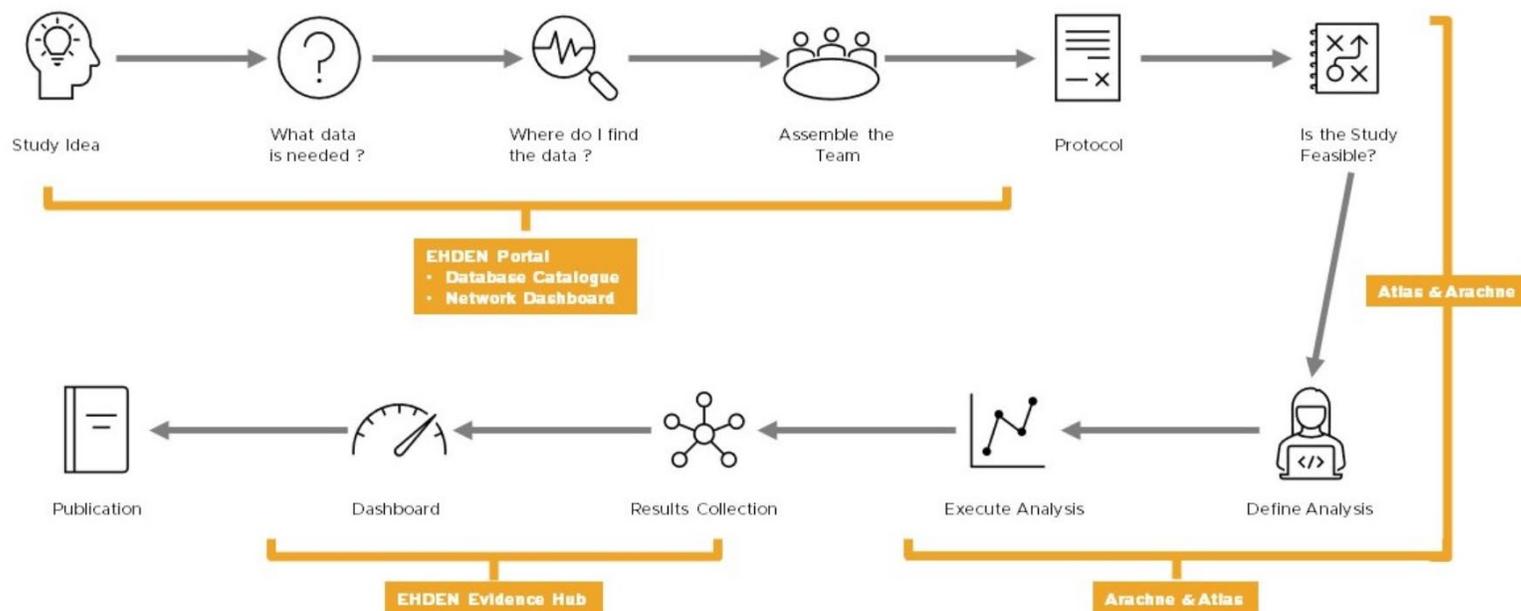


Standard Analysis	Regulatory example
Population-level disease epidemiology	<ul style="list-style-type: none"> Prevalence of rare disease/s Background rates of AESI or DMEs
Patient-level disease epidemiology	<ul style="list-style-type: none"> Natural history/prognosis Current practice/treatment patterns
Population-level DUS	<ul style="list-style-type: none"> Incidence and prevalence of use of medicine/s over time
Patient-level DUS	<ul style="list-style-type: none"> Describing indication/s for drug/s Treatment duration, cumulative use

※This presentation represents the views of the DARWIN EU® Coordination Centre only and cannot be interpreted as reflecting those of the European Medicines Agency or the European Medicines Regulatory Network.

OMOP CDM-mappedデータを使った研究の始め方

- EHDEN Portalからフローを提供している。
- 複数のデータベースではなく1つのデータベースに興味があった場合は直接そのデータパートナーに連絡して研究実施依頼をすることは可能。
- EHDENのスポンサーになっていなくても研究を依頼することができる（費用はもちろんかかる）。



EHDEN Portal

Overview Demographics Data Domains Data Provenance Visit Concept Browser About



Countries

30

Data Sources

193

Patients

363M

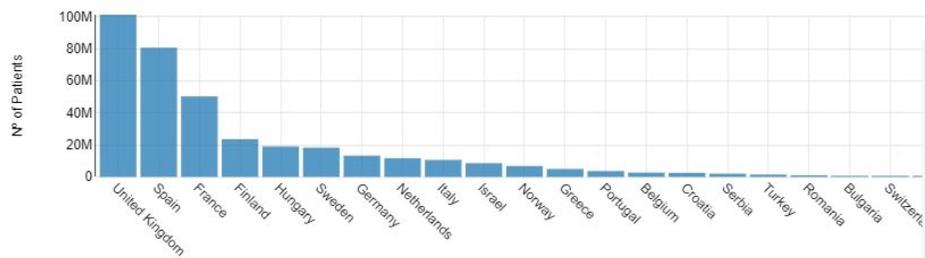
Datasource Types

Hospital Registry Hospital + Lab Resu 1/29



Total number of patients per country (only the records that were mapped into OMOP CDM)

Patients by Country



Meta Data

Show All entries

Search 193 records...

acronym	name	country	database_type	number_of_patients	source_release_date	cdm_release_date	cdm_version	vocabulary_version	execution_date	package
NNRD	National Neonatal Research Database	United Kingdom	Hospital	1180103	N/A	2021-12-30	v5.3.1	v5.0 16-DEC-21	2022-03-25	10.2
DCCG	Danish Colorectal Cancer Group	Denmark	Registry	76849	2021-05-21	2021-05-21	v5.3	v5	2021-06-25	10.2
DDW	The Dutch Data Warehouse of Amsterdam UMC	Netherlands	Hospital	1834	2020-11-23	2020-12-03	v5.3.1	v5.0 06-NOV-20	2021-03-29	10.2
ULSM COVID	ULSM COVID-19	Portugal	Hospital + Primary Care	9750	2021-02-11	2021-08-25	5.3.1	1.0	2021-08-27	10.2
PSSJD	PSSJD	Spain	Hospital	601107	2021-10-08	2021-10-08	5.3	1	2022-03-02	10.2
HIC	HIC Data	United Kingdom	Hospital + Lab Results	1910220	2021-01-25	2021-01-27	v5.3.1	v5.0 06-NOV-20	2021-03-22	10.0

ヨーロッパ30か国の193のデータソース, 3億人をカバー

Databases / CPRD AURUM

Fingerprint

Database Dashboard

Hits: 537 Unique Views: 305 Filled: 64.2 %

DATABASE DESCRIPTION

Geographical origin of data

England, United Kingdom

Demographical extent of data collection

- Regional
- Nation-wide
- Other

Database Type

- Inpatient (hospital) care
- Outpatient general practitioner care
- Outpatient specialist care
- Long term/ skill nursing facility care
- Pharmacy care
- Registry
- Biobank
- Cohort

Insurance Status

- Insured
- Uninsured
- Public
- Private
- Other

Socioeconomic status

Data continuity

- Longitudinal
- Cross-sectional

Reasons to enter or exit the database

- The subject enters or leaves the practice of a general practitioner
- The subject enters or leaves the practice of a paediatrician in primary care

Demographics

Data Domains Data Provenance Observation Period Visit Concept Browser Meta Data

Number of Patients

409k

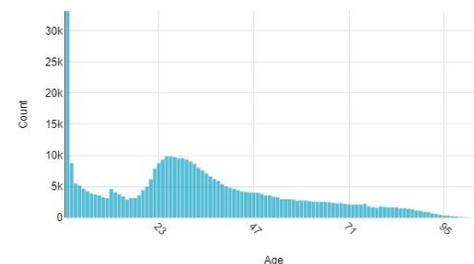
Gender

gender	number of
FEMALE	213k
MALE	196k

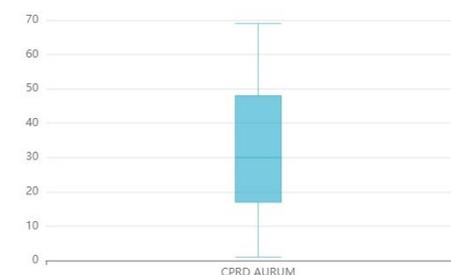
Gender



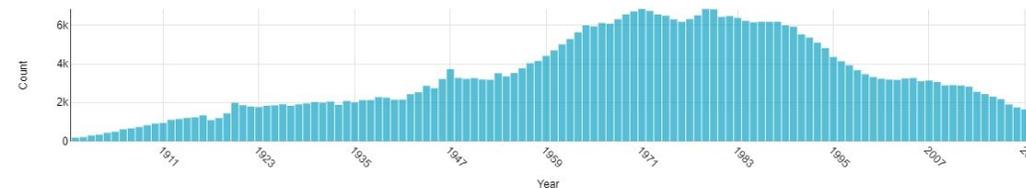
Age at first observation



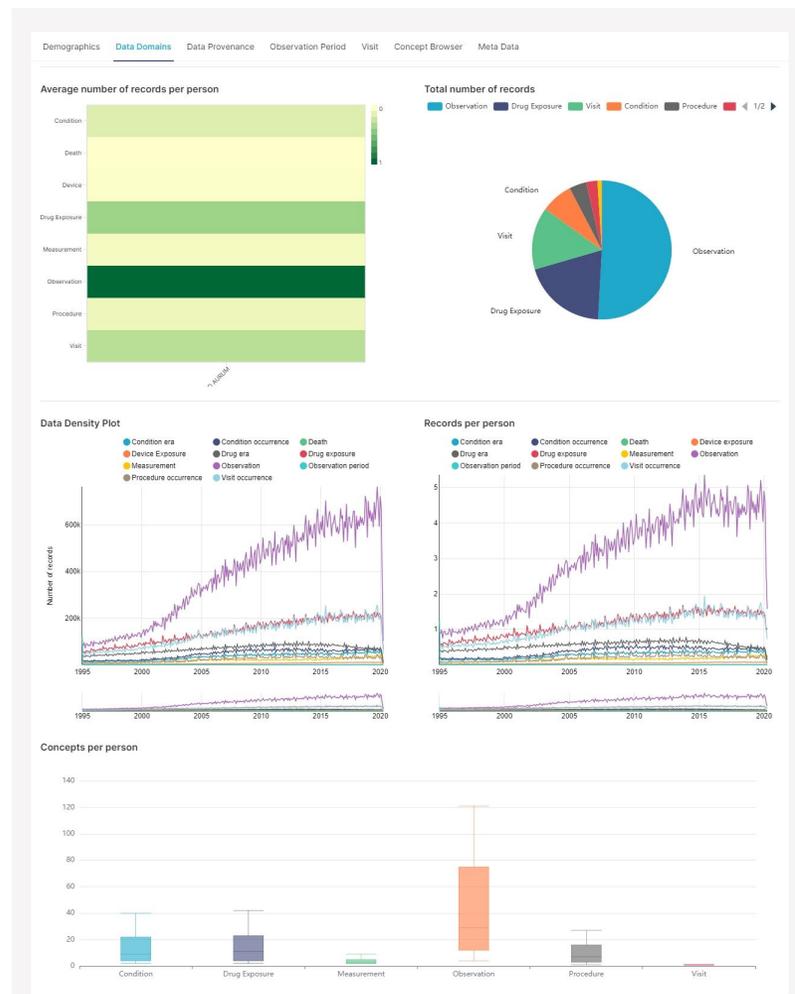
Distribution of age at first observation period



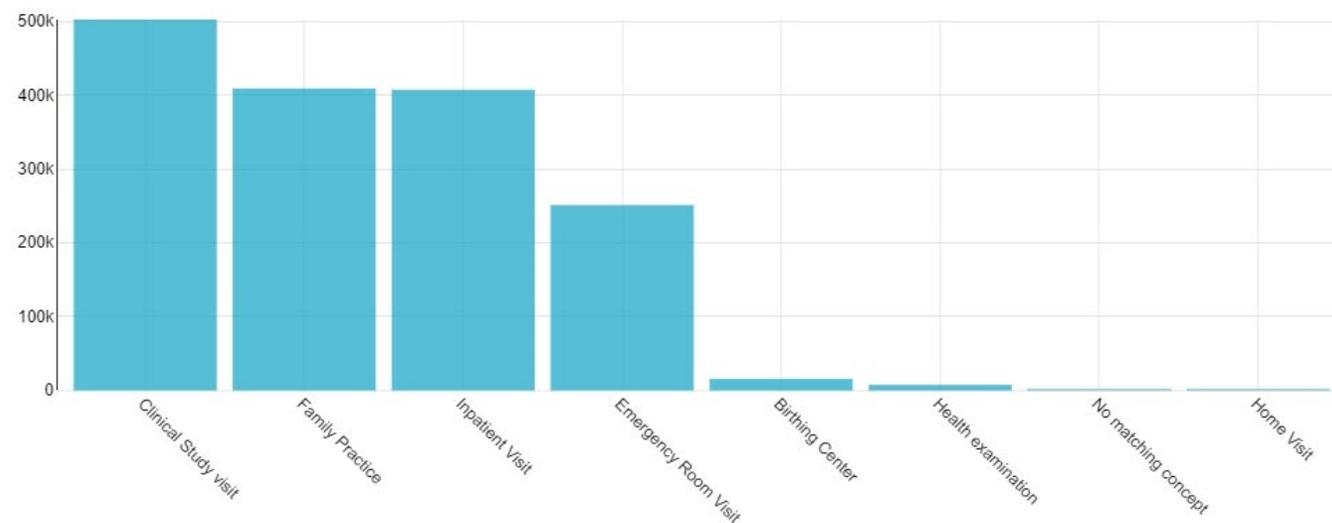
Year of Birth



各DBの情報，患者集団を調査可能



Visit Type Graph



Disclaimer: The presented data can be used for an initial feasibility assessment but should not be used for research without involvement of the data custodian. The number of patients in each database reflects the number of records that were mapped to OMOP CDM. These values may differ from the numbers in the Catalogue.

各DBで取得可能なデータの詳細を調査可能

EHDEN Portal : レコード数を調査

Concept Browser

Show 10 entries

Search diabetes

concept_id	concept_name	domain_id	drc	rc
201826	Type 2 diabetes mellitus	Condition	278300	275100
201820	Diabetes mellitus	Condition	375100	54800
201254	Type 1 diabetes mellitus	Condition	39300	39100
376683	Nonproliferative retinopathy due to diabetes mellitus	Condition	14500	5700
44808385	Pre-diabetes	Condition	4500	4500
4174977	Retinopathy due to diabetes mellitus	Condition	21800	3000
4130162	Insulin treated type 2 diabetes mellitus	Condition	2500	2500
4024659	Gestational diabetes mellitus	Condition	2200	2200
4044391	Neuropathy due to diabetes mellitus	Condition	2300	1700
44809658	High risk of diabetes mellitus	Condition	1600	1600

各項目を検索して該当するレコード数を調査できる ⇒ Feasibility調査

EHDEN Portal : DBを探索



Domain Filter
ENTITY TYPE

Measurement

Concept Filter
CONCEPT

Body mass index x Body mass index (BMI) [Percentile] x Body height x Body weight x

APPLY

The concepts browser shows the total number of records (RC) and total number of records including all descendant (DRC) across the selected data sources. No exact numbers are provided but they are rounded up to the nearest multiple of 100.

Concept Browser

		acronym	2CA-Braga		ABUCASIS		AMYPAD PNHS		AOU- SANLUIGI		AP-HM		AZDDB		AZG		AZMM		AmsterdamUMCdb		AssutaSurgical		BARDENA		BDR		BIFAF
		metric	RC	DRC	RC	DRC	RC	DRC	RC	DRC	RC	DRC	RC	DRC	RC	DRC	RC	DRC	RC	DRC	RC	DRC	RC	DRC	RC	DRC	RC
concept_name	domain_id																										
Body height	Measurement												817k	817k					4.9k	4.9k						704k	704k
Body mass index	Measurement	38.1k	38.1k	7.42M	7.42M	6.1k	6.1k	30.7k	30.7k					66.4k	66.4k	310k	310k							2.35M	2.35M		59.7k
Body mass index (BMI) [Percentile]	Measurement																										
Body weight	Measurement	317k	317k			6.1k	6.1k	30.8k	30.8k	69k	69k	1.82M	1.82M	1.47M	1.47M	464k	464k	5.2k	5.2k	701k	701k				705k	705k	88k

各項目を検索してDB横断的に該当するレコード数を調査できる ⇒ Feasibility調査

アジェンダ

01

セミナー概要

02

EHDENについて

03

社内で調べたこと

- EMAは医薬品の承認後に安全性を評価のための研究をDARWIN EUに委託している
- RWDを使用したNon-interventional study(NIS)に関するガイドライン（draft）を公開しており， data quality frameworkとしてOHDSIの解析ツールの使用を一例として挙げている。

2. Scope

The scope of this reflection paper is the design, conduct and analysis of NIS using RWD to generate RWE for regulatory purposes. The use of RWD in the context of CTs, e.g. to provide an external control arm for a single arm trial or to serve as a data source to recruit participants for a CT, is out of scope of this document.

6.5. Data quality frameworks

Several data quality frameworks provide a set of characteristics determining the fitness-for-use of data (18). While they differ by the number of dimensions and the names given to these dimensions, they do not diverge substantially. It is therefore recommended:

- To use a data quality framework appropriate to the data source, such as the tool developed by the Observational Health Data Science and Informatics (OHDSI) network for databases conforming to the Observational Medical Outcomes Partnership (OMOP) common data model (19) and make the results available to regulators

データ品質レポートを作成

DARWIN EU関連の規制研究に参加する際に提出する

1. Table of content

- 1. Table of content 2
- 2. Execution details 3
 - 2.1. CDM Source Table 3
- 3. Clinical data 4
 - 3.1. Record counts per OMOP CDM table 4
 - 3.2. Data density plots 5
 - 3.3. Distinct concepts per person 7
 - 3.4. Observation Period 7
 - 3.5. Date Range 8
 - 3.6. Day, Month, Year of Birth 9
- 4. Vocabulary mappings 10
 - 4.1. Mapping Completeness 10
 - 4.2. Drug Mappings 10
 - 4.3. Unmapped Codes 10
 - 4.4. Mapped Codes 13
 - 4.5. Source to concept map 17
- 5. Data Quality Dashboard 17
- 6. Drug Exposure Diagnostics 19
- 7. Technical Infrastructure 20
 - 7.1. R packages 20
 - 7.2. System Information 21
 - 7.3. Vocabulary Query Performance 21
 - 7.4. Applied indexes 21
 - 7.5. Achilles Query Performance 22
- 8. Appendix 27
 - 8.1. Vocabulary table counts 27
 - 8.2. Vocabulary concept counts 27

3.1. Record counts per OMOP CDM table

Table 2. The number of records in all clinical data tables

Table	#Records	#Persons	%Persons
measurement	15,072,717	23,517	100%
drug_exposure	1,483,660	22,478	95.6%
procedure_occurrence	1,296,930	23,484	99.9%
visit_detail	1,276,461	23,517	100%
visit_occurrence	1,276,461	23,517	100%
observation	465,207	19,815	84.3%
condition_occurrence	402,721	23,243	98.8%
condition_era	389,210	23,243	98.8%
drug_era	178,000	22,478	95.6%
person	23,517	23,517	100%
observation_period	23,517	23,517	100%
provider	1,181	NA	NA
cdm_source	1	NA	NA
episode_event	0	NA	NA
fact_relationship	0	NA	NA
episode	0	0	0%
cost	0	NA	NA
death	0	0	0%
device_exposure	0	0	0%
dose_era	0	0	0%
note	0	0	0%
payer_plan_period	0	0	0%
specimen	0	0	0%
care_site	0	NA	NA
location	0	NA	NA
note_nlp	0	NA	NA
metadata	0	NA	NA

Query executed in 3.69 seconds

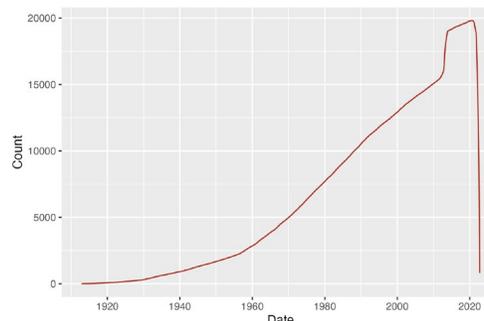


Table 14. All 6 unmapped observations. Counts are rounded up to the nearest hundred. Values with a record count <=5 are omitted.

#	Source Value	#Records	%Records
1	1191 Aspirin	800	0.2%
2	29046 Lisinopril	500	0.1%
3	7984 Penicillin V	500	0.1%
4	10831 Sulfamethoxazole / Trimethoprim	200	0.0%
5	25037 cefдинир	100	0.0%
6	5640 Ibuprofen	100	0.0%

Query executed in 0.01 seconds

OMOP CDM導入例

• EUの状況

- EHDENがSMEを雇い、データベンダー（医療機関等）でのOMOP化をサポート。OMOP化の費用をデータベンダーに対して支援。EHDENはEU政府と欧州製薬団体連合体（EFPIA）による助成金からの資金で運営されている

• アジア

- シンガポールでは保健省と協力し、TRUSTと呼ばれる国家臨床データ共有プラットフォームでOMOP CDMを採用
- 韓国ではThe Federated E-Health Big Data for Evidence Renovation Network (FeederNet) platformが存在し、53病院7000万人（重複あり、2022/3時点）のデータが国家資金援助によりOMOP CDM化された。また、韓国の全人口をカバーするadministrative claim data(HIRA)もOMOPCDM化された

• US

- NIHが主導する精密医療を目指すAll of UsプログラムではOMOP CDMを採用
- 一方で、FDAがリードするSentinel Common data modelも存在する（designed for claims data and being expanded for HER dataでOMOP CDMより拡張性が低いと言われる）

OMOP CDM導入例

- 日本

Yuimediが愛媛大学と共同でHL7 FHIRからOMOP CDMへの変換ツールを開発いたしました

株式会社Yuimedi 2024年5月14日 12時31分



Yuimediは愛媛大学と医療データの国際的な共同利用に向けたデータ標準化の試みとして共同研究を行いました。本共同研究の一部の成果は、2023年11月に行われた医療情報学会において口頭発表を行いました。その結果を踏まえ、さらなる取組も行いました。

愛媛大学は国際共同研究を目的とし、保有する医療データをOMOP CDM化した

<https://prtmes.jp/main/html/rd/p/000000014.000081576.html>

DATASET PROVISION データセット提供

リサーチニーズに応じて、**受診情報・傷病情報・検査情報**等の整理された13のテーブルで、お客さまが分析しやすい形式で抽出し、**匿名加工医療情報**をご提供するサービスです。

特徴

POINT 1



OMOP CDMに準拠
国際基準に基づいたテーブル構成にて、データセットをご提供いたします。適切なテーブル構成のため、分析業務を効率的に実施頂けます。

POINT 2



適切なデータ抽出
リサーチニーズに応じた適切なデータセットを抽出いたします。ご自身の手で自由にデータを分析頂くことが可能です。

データセットカラム例

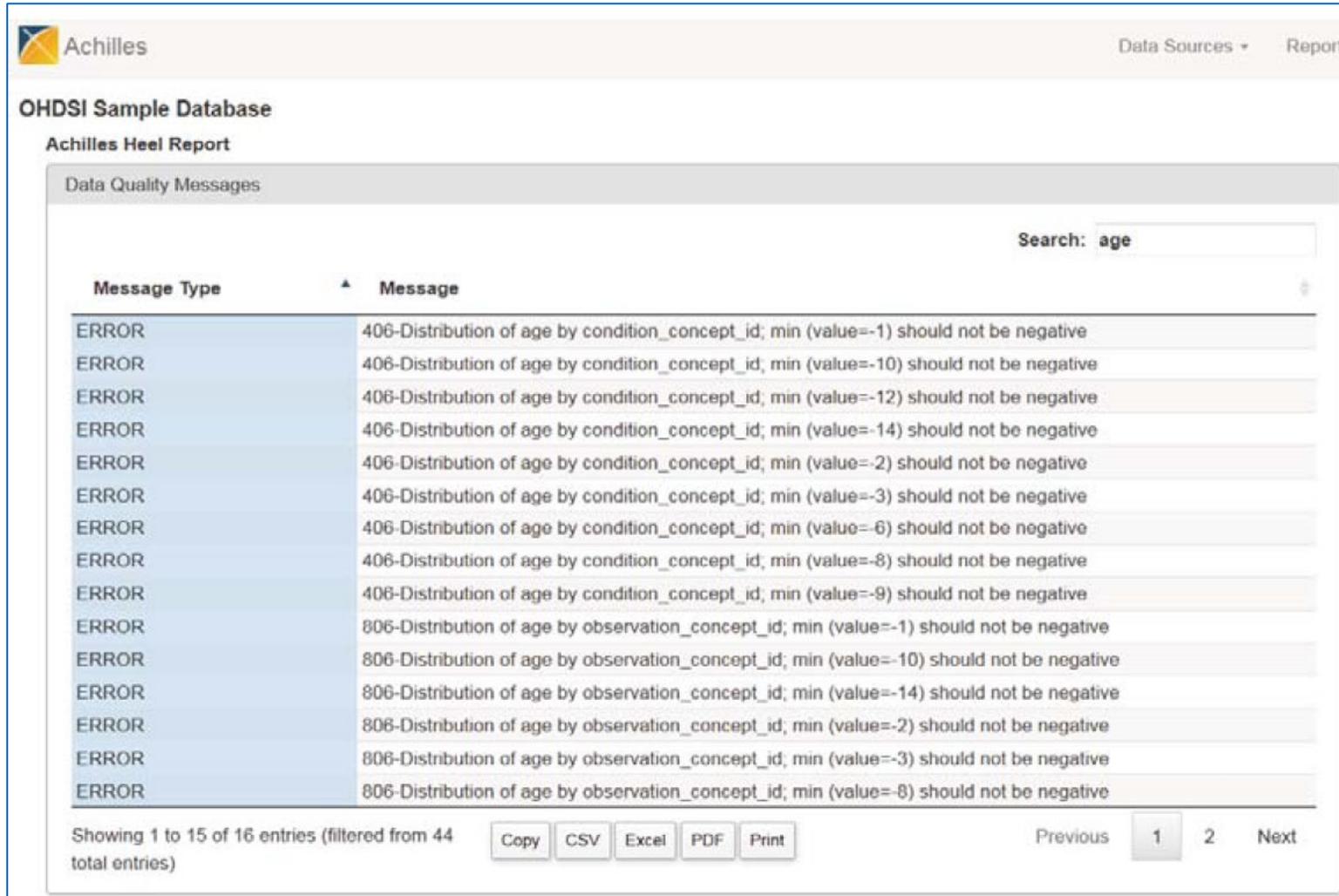
基本情報	観察期間	受診情報	傷病情報	傷病情報詳細	医薬品使用情報
患者 ID 性別 生年 生年月 ...	レコード ID 患者 ID 観察開始日 観察終了日 ...	レコード ID 患者 ID 入院 / 外来 入退日 入院日時 ...	レコード ID 患者 ID 病名管理番号 傷病名 ICD10 コード ...	レコード ID 患者 ID 病名管理番号 傷病名 ICD10 コード ...	レコード ID 患者 ID YJ コード 投与開始日 投与開始日時 ...
医薬品辞書	検査	検体	死亡	関係性	イベント
レコード ID 配合剤区分 ATC コード 一般名 製品名 ...	レコード ID 患者 ID JLAC10 コード 検査項目 分析物名 ...	レコード ID 患者 ID 検体の種類 検体採取日 ...	患者 ID 死亡日 ...	項目 1 テーブル 項目 1 レコード ID 項目 2 テーブル 項目 2 レコード ID ...	レコード ID 患者 ID イベント種別 開始日 終了日 ...

TOPPANは電子カルテデータをOMOP(like) CDMで提供

OMOP化の評価と懸念 (1/2)

- OMOP化したデータから元のデータに遡ることが可能
- OMOPのデータ構造はカスタマイズ可能
 - 疾患特異的なテーブルを作成。オンコロジー，希少疾患など。
 - [Oncology extension](#) (次のOMOP CDMv6.1からリリースされる予定とのこと; githubなどへのリンクも)
 - [OHDSI Oncology WG](#)
 - DIGICORE/DigiONE: FlatironのEU版を目指す (see Mahon, P., Chatzitheofilou, I., Dekker, A. et al. A federated learning system for precision oncology in Europe: DigiONE. Nat Med 30, 334–337 (2024). <https://doi.org/10.1038/s41591-023-02715-8>)
- データ品質の調査は可能
 - 既存のデータのOMOP化による品質調査の結果は報告されているものもある
 - データの質評価のためのツールは存在する
 - DataQualityDashboard (DQD);([github](#), [DQD output example](#); [paper](#)) ※Achillesでは不十分で追加された
 - Achilles ([github](#), [paper](#), [precomputed analysis csv](#))
- 元のデータとOMOP化したデータでRWD研究の結果に差が生じるか不明

Achilles Heel のデータチェック結果の例



The screenshot shows the Achilles Heel Report interface for the OHDSI Sample Database. It displays a list of Data Quality Messages filtered by the search term 'age'. The messages are categorized by Message Type (ERROR) and Message content. The interface includes a search bar, a table of results, and navigation controls at the bottom.

Message Type	Message
ERROR	406-Distribution of age by condition_concept_id; min (value=-1) should not be negative
ERROR	406-Distribution of age by condition_concept_id; min (value=-10) should not be negative
ERROR	406-Distribution of age by condition_concept_id; min (value=-12) should not be negative
ERROR	406-Distribution of age by condition_concept_id; min (value=-14) should not be negative
ERROR	406-Distribution of age by condition_concept_id; min (value=-2) should not be negative
ERROR	406-Distribution of age by condition_concept_id; min (value=-3) should not be negative
ERROR	406-Distribution of age by condition_concept_id; min (value=-6) should not be negative
ERROR	406-Distribution of age by condition_concept_id; min (value=-8) should not be negative
ERROR	406-Distribution of age by condition_concept_id; min (value=-9) should not be negative
ERROR	806-Distribution of age by observation_concept_id; min (value=-1) should not be negative
ERROR	806-Distribution of age by observation_concept_id; min (value=-10) should not be negative
ERROR	806-Distribution of age by observation_concept_id; min (value=-14) should not be negative
ERROR	806-Distribution of age by observation_concept_id; min (value=-2) should not be negative
ERROR	806-Distribution of age by observation_concept_id; min (value=-3) should not be negative
ERROR	806-Distribution of age by observation_concept_id; min (value=-8) should not be negative

Showing 1 to 15 of 16 entries (filtered from 44 total entries) Copy CSV Excel PDF Print Previous 1 2 Next

OMOP化の評価と懸念 (2/2)

- Cohortの特性を確認し研究を断念することもある

- 例：例えば一型糖尿病を特定しなかったが糖尿病という一般的な病名しか使われていないといったことが明らかとなって利用を断念するなど。

- Cohortの特性確認ツールは色々ある

- {CohortCharacteristics} ([github](#))
- {OMOPSketch}：各コンセプトの数などを確認するパッケージ ([github](#))([shiny](#))
- EHDEN Portal：EHDENの提供するOMOP mapped databasesがまとまっているダッシュボードでコンセプト数の確認

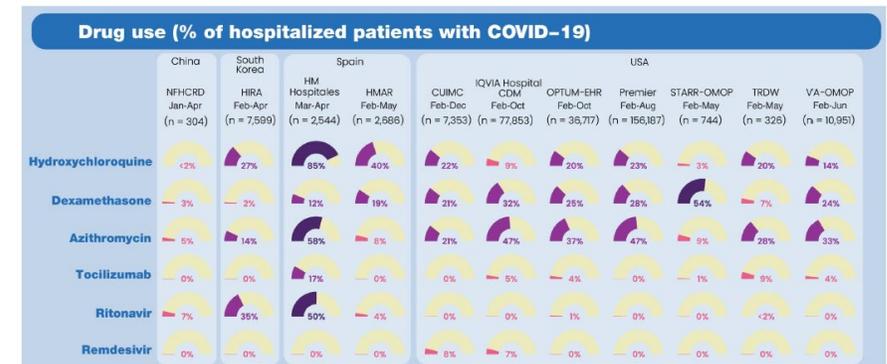
複数DB間の解析方法

- DBは結合せずに解析する

- データを外部に出さず分析できるのがOMOP CDMの利点
- DB間で異なる点が多くある（ヘルスケアシステム、保険償還ポリシー、カバーしている集団特性や、データ収集、コーディング、記録における異なるプラクティス）

- DB間で結果が異なるときの解釈

- EMAによるRWDを使用したNISに関するガイドライン
 - DB間で異なる点を考慮する（集団特性等）。[EMA RWD catalogues](#)等を参考にDB間の特徴を比較
 - プライマリケアデータvs病院データ, 同じ臨床疾患分類システムごと等の感度分析を実施する（先行研究：<https://onlinelibrary.wiley.com/doi/10.1002/pds.5631>）
 - フォレストプロット等で可視化



ネットワーク研究の守秘義務

- EHDENにおけるネットワーク研究では、実際に製薬企業からの依頼による研究もあり、情報の開示範囲を相談することは可能（解析者がミーティングに参加できない、資料も部分的にしか共有されない等）

ご清聴ありがとうございました！

オックスフォードで
Lovelyな天気と
サマースクールを
楽しんできてください

- [Software Tools – OHDSI](#) (いろいろなツールのDemoが見れる)
- Reich C, Ryan PB, Stang PE, Rocca M. Evaluation of alternative standardized terminologies for medical conditions within a network of observational healthcare databases. J Biomed Inform. 2012 Aug;45(4):689-96. doi: 10.1016/j.jbi.2012.05.002. Epub 2012 Jun 7. PMID: 22683994. (語彙マッピングについての論文)
- [HADES \(Rパッケージ群\) のgithub](#)
 - Schuemie M, Reps J, Black A, et al. [Health-Analytics Data to Evidence Suite \(HADES\): Open-Source Software for Observational Research](#) Studies in health technology and informatics. 2024 Jan 25;310:966-970.
 - Schuemie MJ, Cepeda MS, Suchard MA, et al. [How Confident Are We About Observational Findings in Health Care: A Benchmark Study](#) Harvard Data Science Review, 2(1).
- [EHDEN Academy](#) (英語に抵抗がなければかなりガッツリ学べるE-learning)

(参考) EMA RWD catalogues

Clinical Practice Research Datalink

First published: 01/02/2024 Last updated: 08/07/2024

Data source Birth registry Hospital discharge records Other Primary care medical records

Download

Administrative details Data elements collected Quantitative descriptors Data flows and management

Page content

- Administrative details
- Contact details
- Data source regions and languages
- Data source establishment

Administrative details

PURI	https://redirect.ema.europa.eu/resource/30008
Data source ID	30008
Name of data source	Clinical Practice Research Datalink
Data source acronym	CPRD
Data holder	Medicines and Healthcare products Regulatory Agency (MHRA)
Data source type	Birth registry Hospital discharge records Other Primary care medical records
Data source type, other	Variety of linked datasets are available.
Main financial support	Funding by own institution Funding from industry or contract research
Care setting	Hospital inpatient care Hospital outpatient care Other Primary care – GP, community pharmacist level

項目が多くDB比較に有用

tails Data elements collected Quantitative descriptors Data flows and management

The data source contains the following information

Disease information	Yes
Does the data source collect information with a focus on a specific disease? This might be a patient registry or other similar initiatives.	
Rare diseases	Yes
Are rare diseases captured? In the European Union a rare disease is one that affects no more than 5 people in 10,000.	
Pregnancy and/or neonates	Yes
Does the data source collect information on pregnant women and/or neonatal subpopulation (under 28 days of age)?	
Hospital admission and/or discharge	Yes
ICU admission	Yes
Is information on intensive care unit admission available?	
Cause of death	Captured
Cause of death vocabulary	ICD-10 Other
Cause of death vocabulary, other	Medcode ID's
Prescriptions of medicines	Captured
Prescriptions vocabulary	other
Prescriptions vocabulary, other	Prodcodex ID's
Dispensing of medicines	Captured