



# バイオエコノミー市場拡大に向けた バイオデータの連携・利活用の実践

## 医療・ヘルスケア分野における先端事例

IQVIA Solutions Japan G.K.  
Real-World & Analytics Solutions

# 本日のポイント

- 医療・ヘルスケアデータ（RWD）の連携・利活用における新たな手法として、**国際標準データモデル**を活用した**連合解析**の概要と先進事例についてご説明いたします。
- RWD の活用に対して課題となるのは、「**解析の前処理**」と「**プライバシー保護との二律背反**」です。
- 本国においてこれらRWD活用が低迷することは、ライフサイエンス関連市場が**可視化されない**ことにつながり、**国内外からの投資を妨げる一因**にもなっています。  
（新薬開発における”Japan Passing”など）
- 国際標準データモデルの活用は、上記**RWD活用の課題を解決し得る手法**として期待されており、本国においても検討と活用の促進が求められます。



## アジェンダ

- 医療・ヘルスケア分野における  
データ連携・利活用の現状と課題
- データ連携・利活用に関する新たなコンセプト
- データ連携・利活用に関する先端事例
- Q&A



## アジェンダ

- 医療・ヘルスケア分野における  
データ連携・利活用の現状と課題
- データ連携・利活用に関する新たなコンセプト
- データ連携・利活用に関する先端事例
- Q&A

# Real World Data (RWD)

様々な手段により日常的に収集される、  
患者の健康状態および医療行為に関するデータ

RWD are data relating to patient health status and/or the delivery of health care that are routinely collected from a variety of sources.

診療請求に由来するデータ

Medical claims data

デジタルヘルス関連の技術製品  
より生成されるデータ

Data obtained from digital health technologies

電子医療記録に由来するデータ

Electronic health record (EHR) data

製品または疾患登録データ

Product or disease registry data

質問票などから得られる  
健康状態に関連するデータ

Data gathered from other sources that can inform on health status, such as questionnaires

# リアルワールド“データ”とリアルワールド“エビデンス”

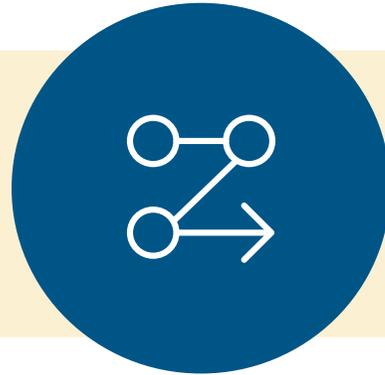
## RWD



様々な手段により日常的に収集される、患者の健康状態および医療行為に関するデータ\*

RWD are data relating to patient health status and/or the delivery of health care that are routinely collected from a variety of sources.\*

## Harmonization



データクレンジング、マッピング、用語集の生成など、形式や構造の標準化プロセス

## Analysis



データマイニング、統計解析、機会学習等の手法を用いたRWDから有用な情報を生成するプロセス

## RWE

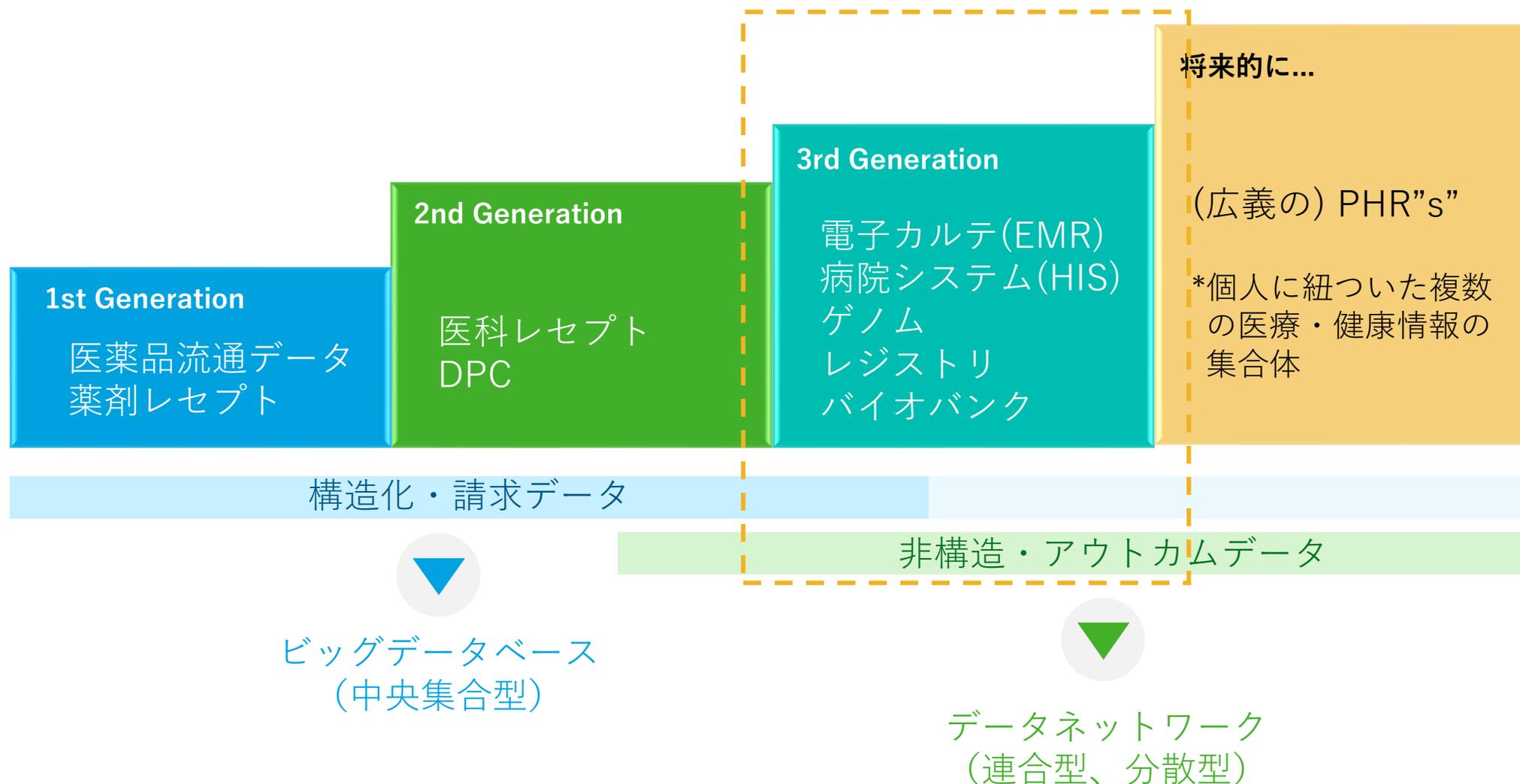


RWDの解析から得られる医薬品・医療機器等の使用実態および潜在的なベネフィット・リスクに関する臨床エビデンス\*

RWE is the clinical evidence about the usage and potential benefits or risks of a medical product derived from analysis of RWD.\*

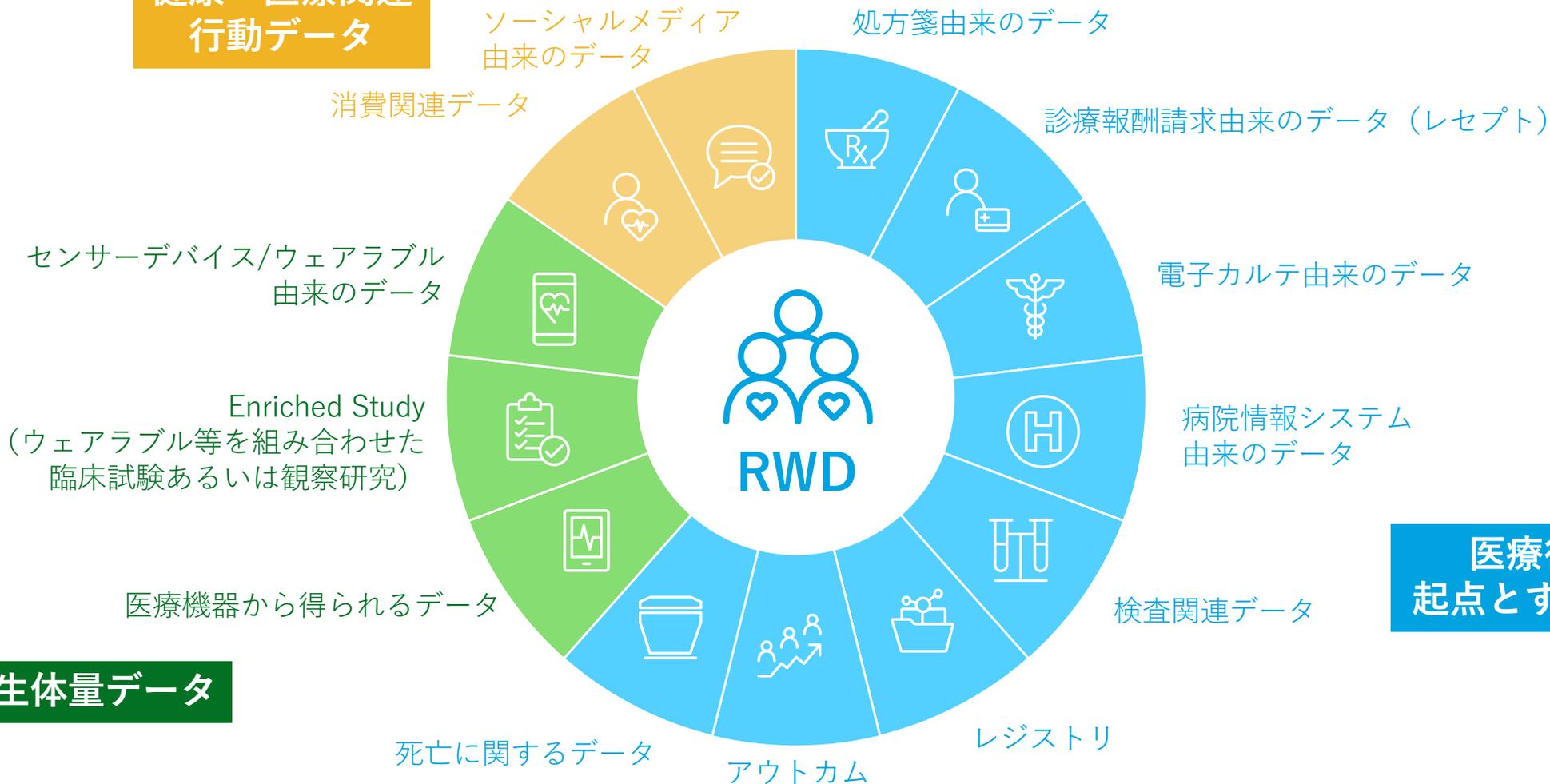
# 一般的なRWDの種類と特徴

第3世代以降の非構造・アウトカムデータの活用が進みつつある



# 多様化するRWDの種類

## 健康・医療関連 行動データ



医療行為を  
起点とするデータ

## 連続生体量データ

## 2つの課題



解析の前処理



プライバシー保護

# 課題 1：解析の前処理 Harmonization

アウトカムRWD活用において最も工数がかかるのは前処理工程

RWD



様々な手段により日常的に収集される、患者の健康状態および医療行為に関するデータ\*

RWD are data relating to patient health status and/or the delivery of health care that are routinely collected from a variety of sources.\*

Harmonization



データクレンジング、マッピング、用語集の生成など、形式や構造の標準化プロセス

Analysis

国内では電子カルテや病院情報システム等の“データ規格”は統一されていないことが多い

テキスト情報や画像情報は構造化処理の工程を経ないと解析が不可能

RWE



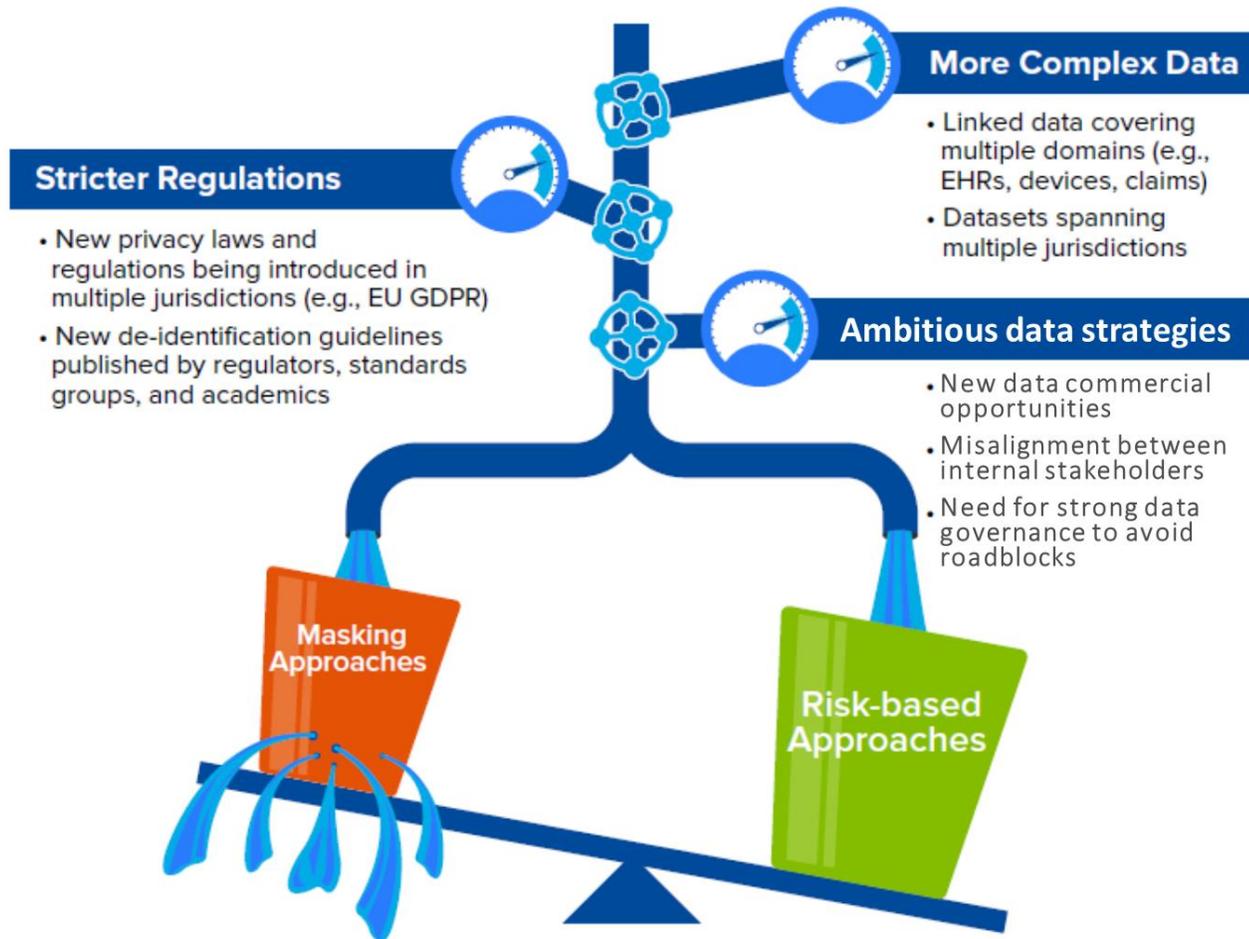
れる等の症例的ナスクにデンス\*

... evidence about the potential benefits or risks of a product derived from analysis of RWD.\*

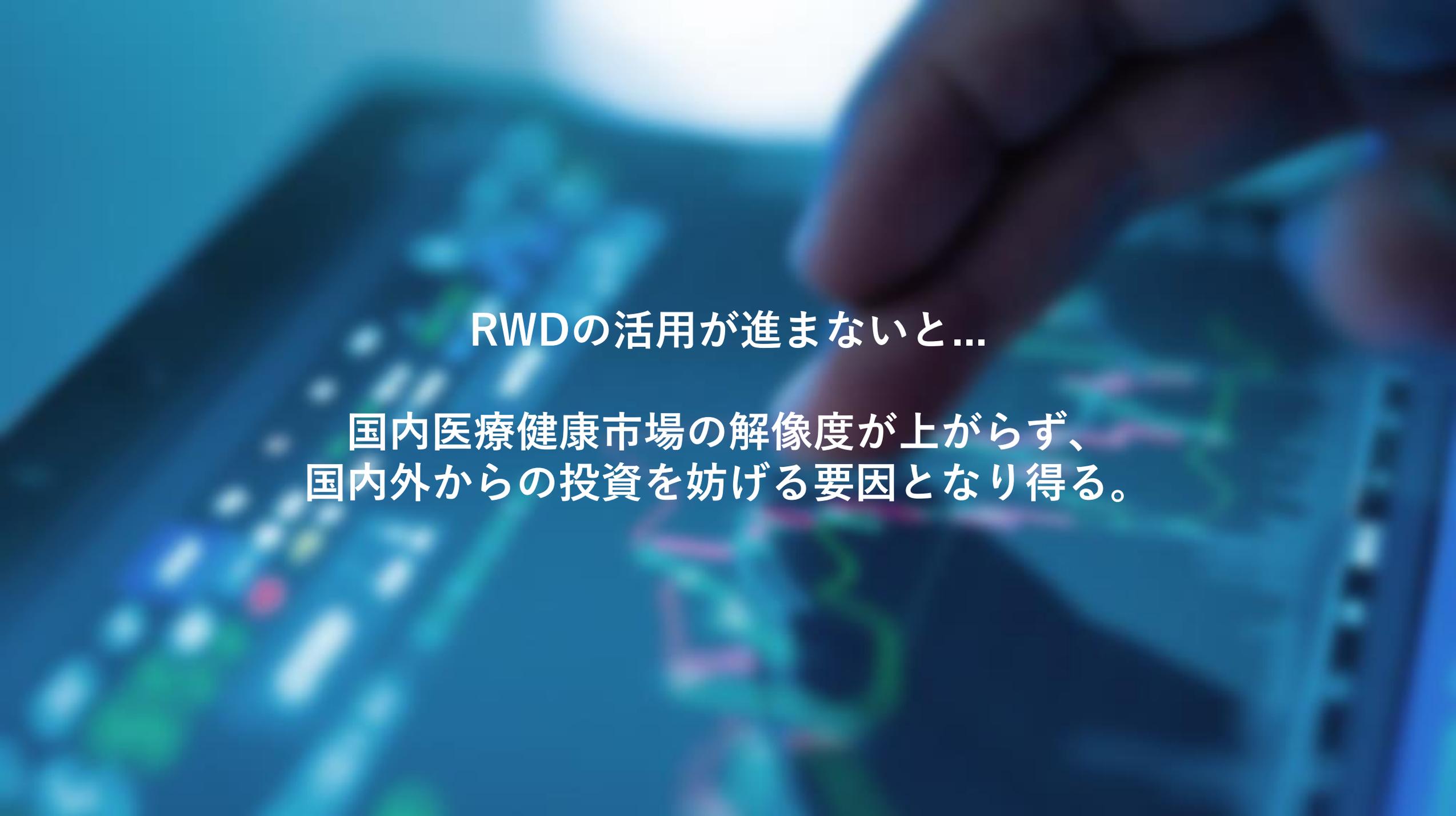
\* Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drugs and Biologics Guidance for Industry

## 課題 2：プライバシー保護と有効利用の二律背反

アウトカムを含む深いデータであるからこそ、個人再特定リスクの考慮が不可欠しかしながら、加工強度を上げるほどデータの有用性も失われる



- ゲノム配列情報など、そもそも匿名化がほぼ不可能なデータも存在
- PHRや希少疾患等の情報は、再特定リスクを考慮して情報を加工すればするほど、本来目的に応じた有用性が失われる

A hand holding a pen pointing at a tablet displaying a colorful data visualization. The background is a blurred image of a hand holding a pen over a tablet with a colorful data visualization.

RWDの活用が進まないと...

国内医療健康市場の解像度が上がらず、  
国内外からの投資を妨げる要因となり得る。



## アジェンダ

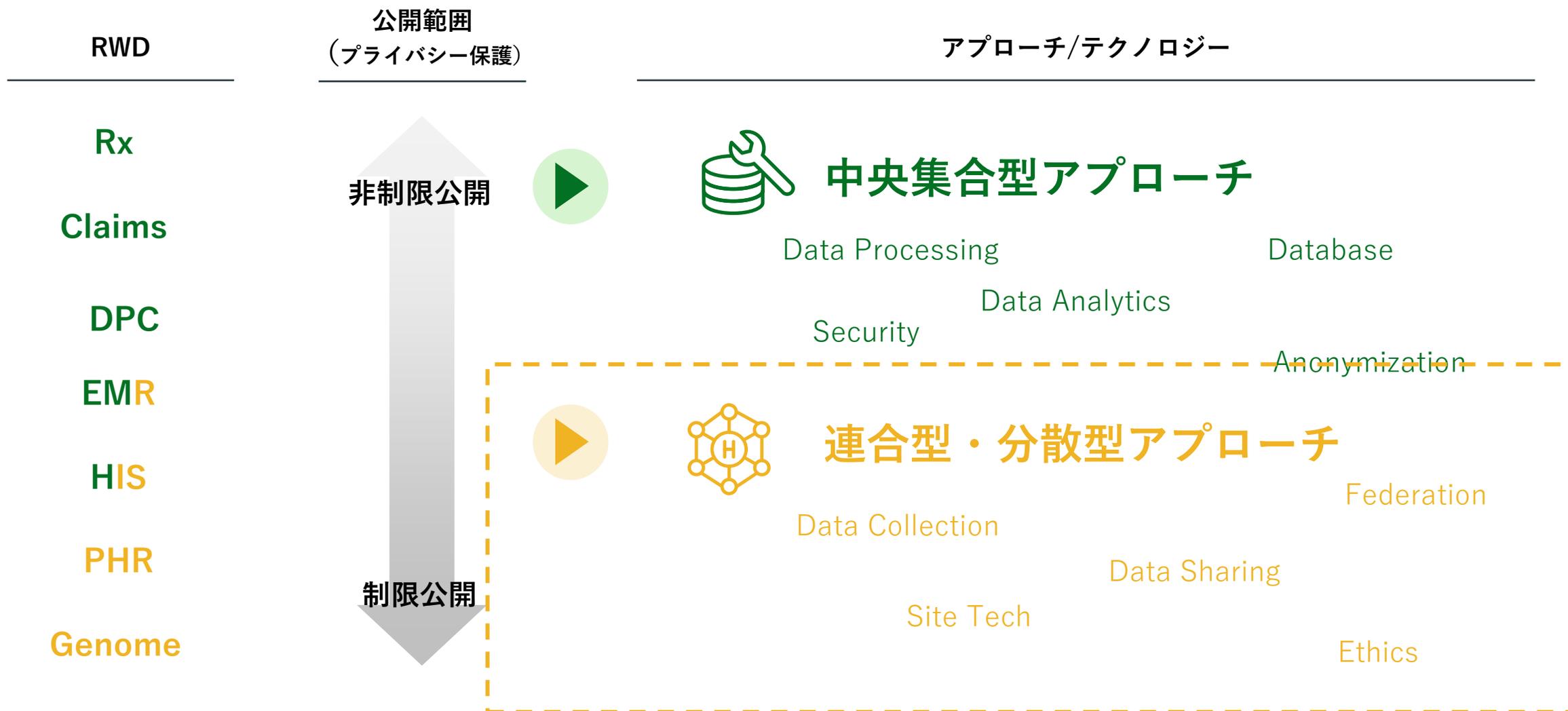
- 医療・ヘルスケア分野におけるデータ連携・利活用の現状と課題
- **データ連携・利活用に関する新たなコンセプト**
- データ連携・利活用に関する先端事例
- Q&A

## 2つの課題の解決策として期待される新たなコンセプト



# RWD活用のアプローチ

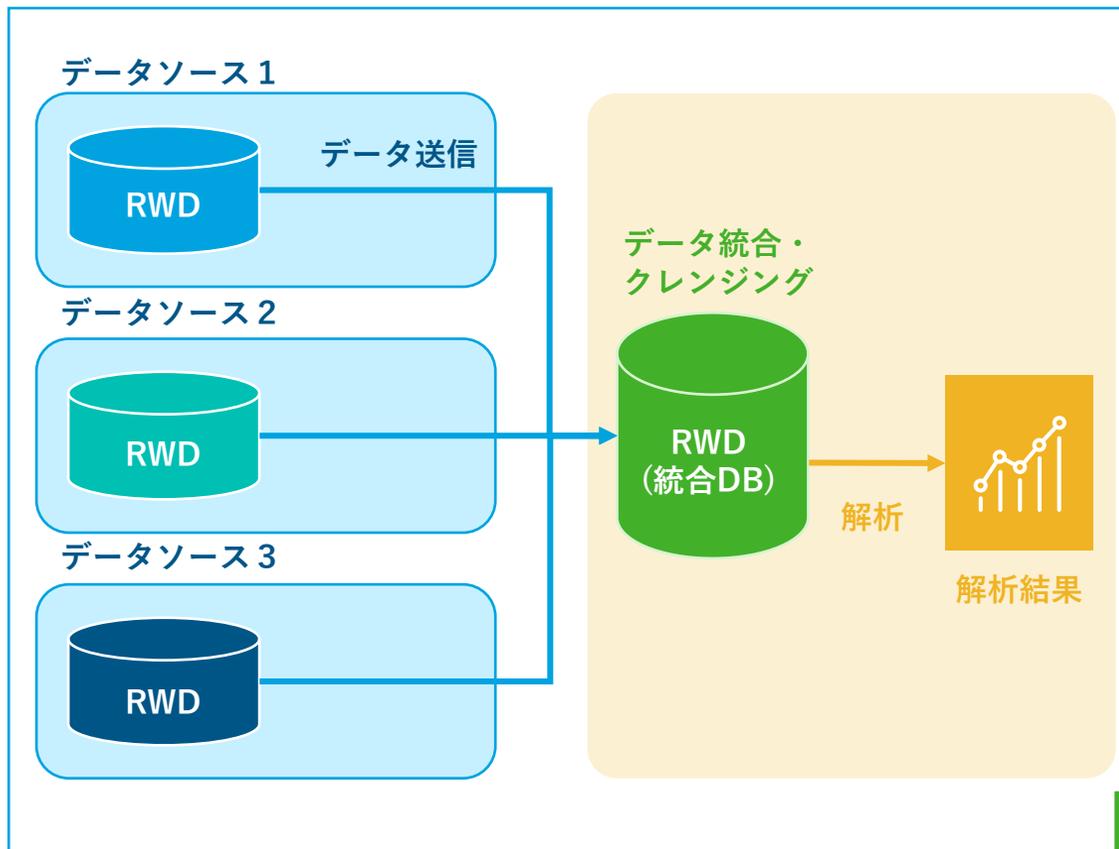
アウトカムを含む深いデータの活用は、従来の“中央集合型アプローチ”から“連合型アプローチ”へ



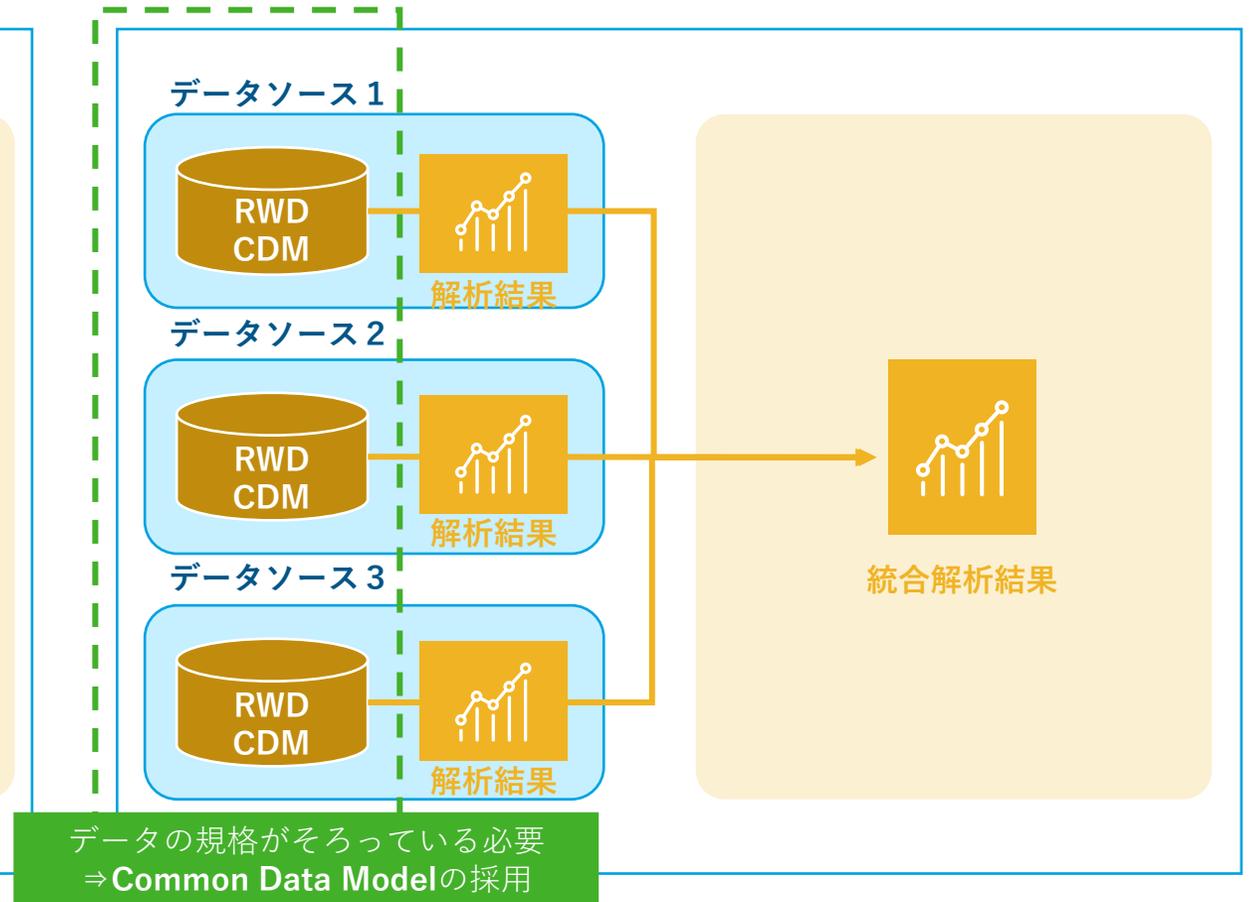
# 新たなRWD活用方法 - 連合解析 (Federated Analysis)

連合解析：データそのものではなく、解析結果を統合することで秘密およびプライバシーに配慮

## 従来型のデータ解析

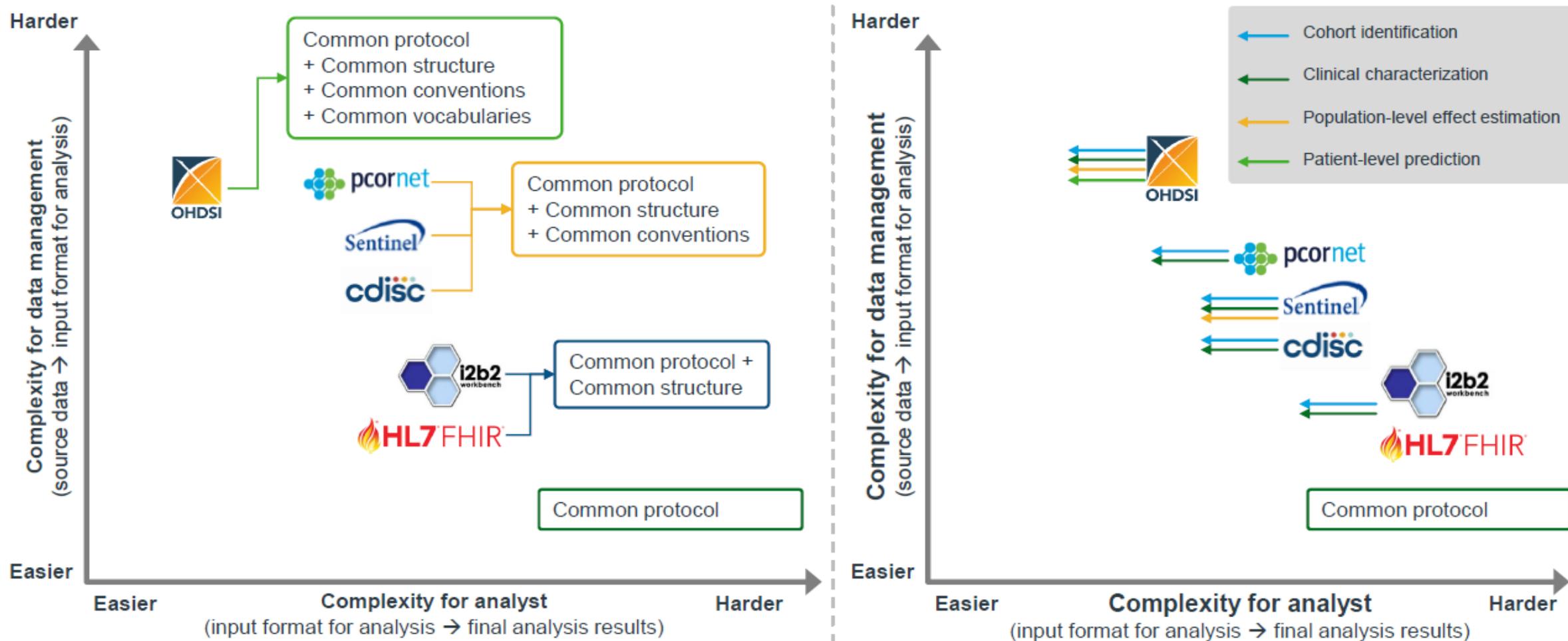


## 連合解析



# 代表的なCommon Data Model (CDM)の比較

「データ管理」と「分析の容易さ」はトレードオフの関係であり目的に応じたバランスが重要





## **OMOP (Observational Medical Outcomes Partnership)**

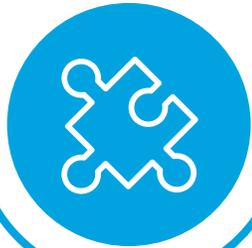
### RWE創出に特化した標準規格

NIH, FDAと米国研究製薬工業協会(PhRMA)の協同プロジェクトであるOMOPにおいて、RWDデータベースを医薬品の安全性調査等に用いることを用意にする目的でOMOP CDMがベース

# OMOP (Observational Medical Outcomes Partnership)

## RWE創出に特化した標準規格

NIH, FDAと米国研究製薬工業協会(PhRMA)の協同プロジェクトにおいて、RWDを医薬品の安全性調査等に用いることを目的として作られたOMOP CDMがベース



### Data Model

- RWE創出に特化したシンプルな共通データモデルであるOMOP Common Data Model



### Vocabulary

- 標準的な用語コード体系(Standard Vocabulary)が整備されており、各用語には概念コード(Concept ID)が付与

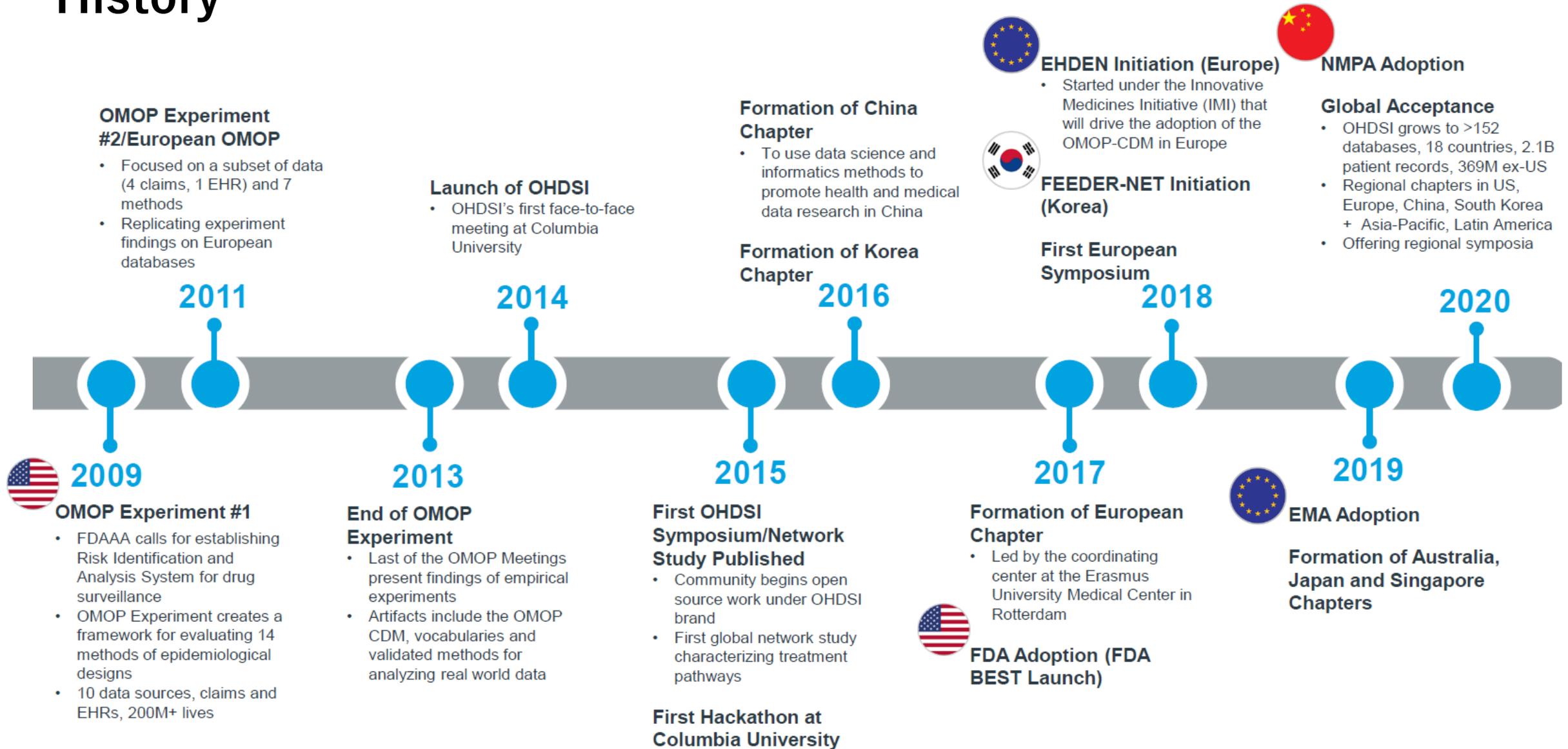


### OHDSI Tool

- OMOPデータ取り扱いに関する様々な共通ツール

Open Source

# History



# OHDSI (The Observational Health Data Sciences and Informatics)

RWEの発展に取り組む、国際的なオープンサイエンスコミュニティ

## OHDSI Mission

To improve health by empowering a community to collaboratively generate the evidence that promotes better health decisions and better care.

## OHDSI Vision

A world in which observational research produces a comprehensive understanding of health and disease.

## OHDSI Values

**Innovation:** Observational research is a field which will benefit greatly from disruptive thinking. We actively seek and encourage fresh methodological approaches in our work.

**Reproducibility:** Accurate, reproducible, and well-calibrated evidence is necessary for health improvement.

**Community:** Everyone is welcome to actively participate in OHDSI, whether you are a patient, a health professional, a researcher, or someone who simply believes in our cause.

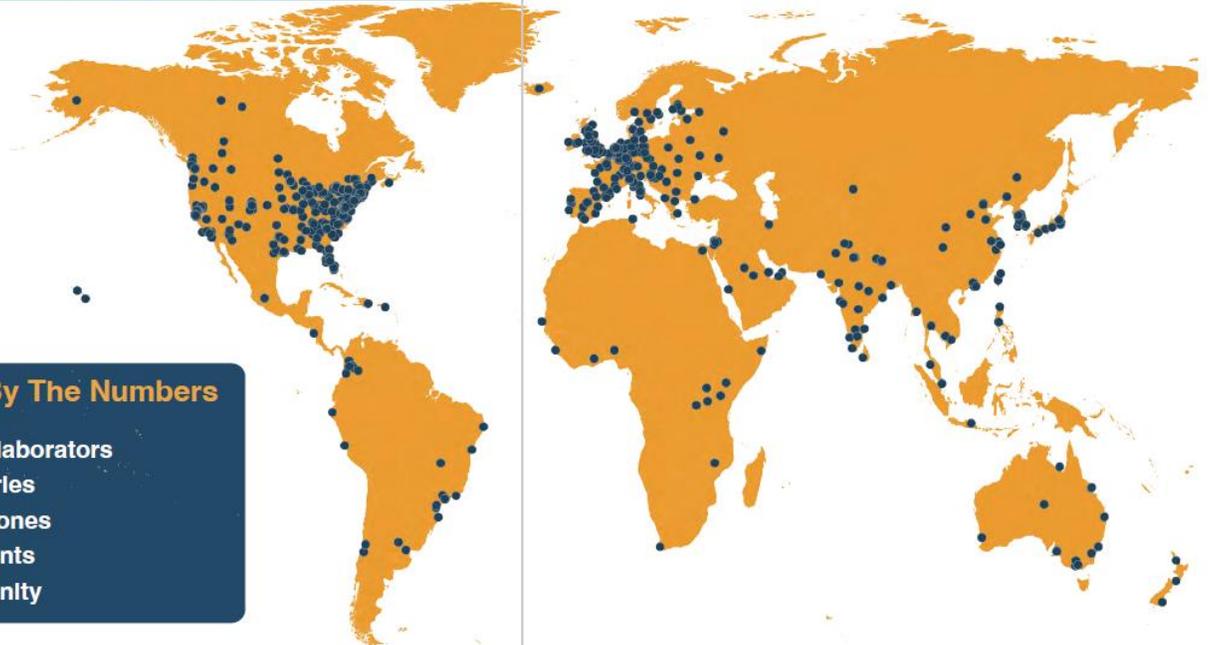
**Collaboration:** We work collectively to prioritize and address the real-world needs of our community's participants.

**Openness:** We strive to make all our community's proceeds open and publicly accessible, including the methods, tools and the evidence that we generate.

**Beneficence:** We seek to protect the rights of individuals and organizations within our community at all times.

## Map of Collaborators

The OHDSI community brings together volunteers from around the world to establish open community data standards, develop open-source software, conduct methodological research, and apply scientific best practices to both answer public health questions and generate reliable clinical evidence.



## OHDSI By The Numbers

- 3,266 collaborators
- 80 countries
- 21 time zones
- 6 continents
- 1 community

Our community is ALWAYS seeking new collaborators. Do you want to focus on data standards or methodological research? Are you passionate about open-source development or clinical applications? Do you have data that you want to be part of global network studies? Do you want to be part of a global community that truly values the benefits of open science? Add a dot to the map below and JOIN THE JOURNEY!

(出所)OHDSI Our Journey 2022 edition

<https://www.ohdsi.org/wp-content/uploads/2022/10/OHDSI-OurJourney-2022.pdf>



## アジェンダ

- 医療・ヘルスケア分野におけるデータ連携・利活用の現状と課題
- データ連携・利活用に関する新たなコンセプト
- **データ連携・利活用に関する先端事例**
- Q&A



# EHDEN (The European Health Data & Evidence Network)

欧州委員会および欧州製薬団体連合会(EFPIA)によるRWEネットワーク

## A federated network of Data Partners

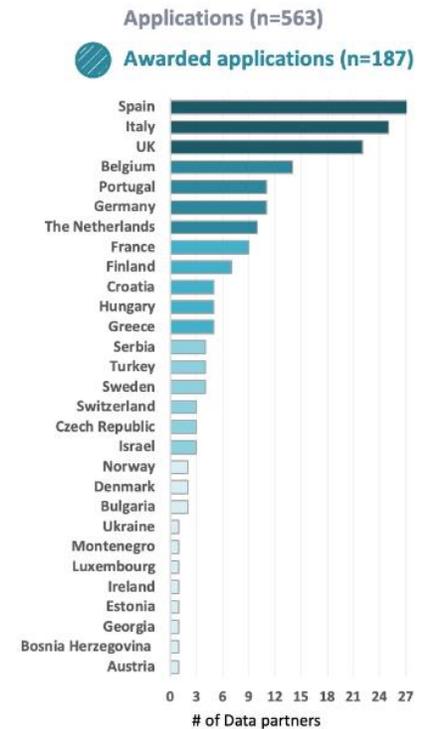
The EHDEN project aims to collaborate with diverse institutions, data sources and data custodians across the EU, with a goal of harmonising source data to the OMOP common data model locally, within a federated network.

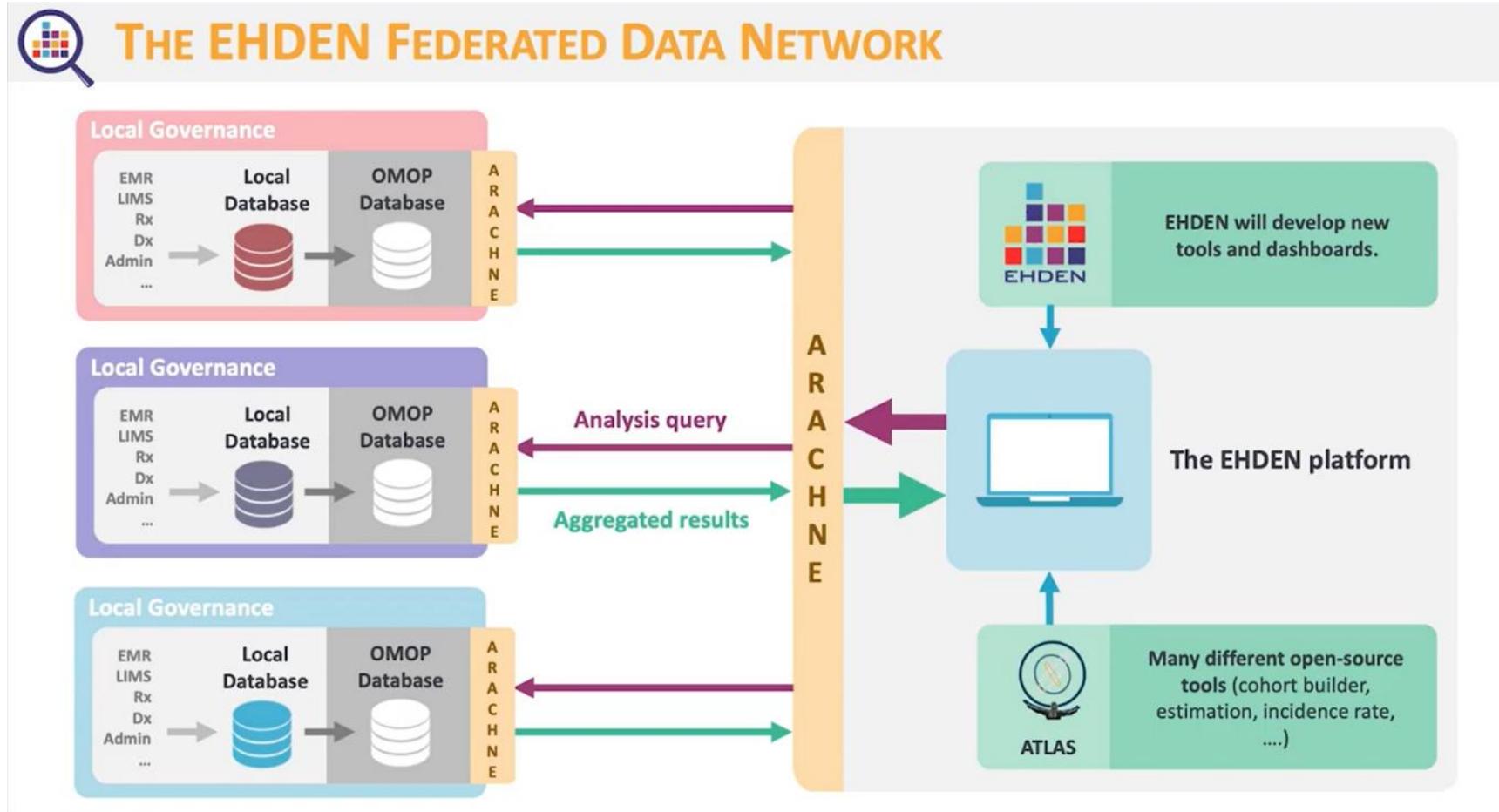
Following the 7 open calls to date we have organised, we currently have 187 Data Partners from 29 different countries which are mapping their data to the OMOP common data model. This includes several EHDEN project partners who have also mapped their data to the OMOP CDM for use in the federated network.

We are looking forward to extending this further via subsequent [open calls](#).



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



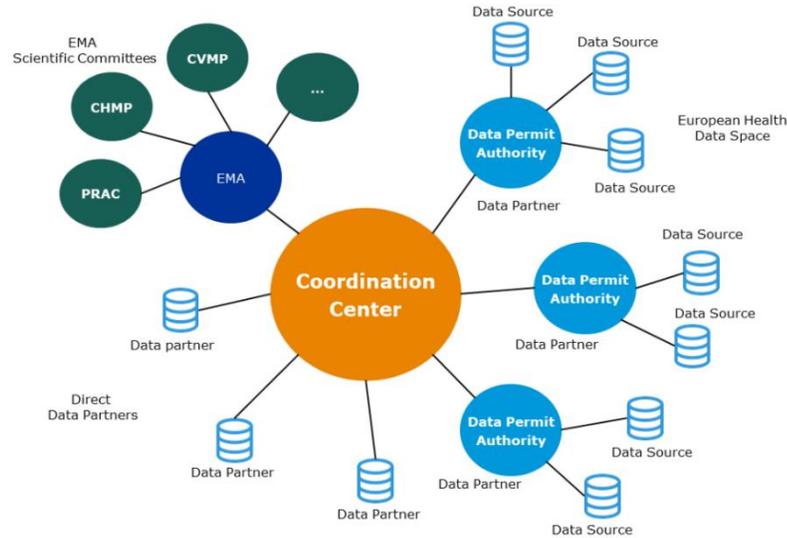




# The DARWIN EU<sup>®</sup> Initiative

EHDS(European Health Data Space)下のEMA主導のRWE利活用プロジェクト

「データ分析・リアルワールド調査ネットワーク調整センター」  
(the Coordination Center for the Data Analysis and Real World Interrogation Network: **DARWIN EU**)



OMOP CDMを用いた、  
EUデータパートナーの  
Federated Network

Figure 1: Overview of DARWIN EU timelines

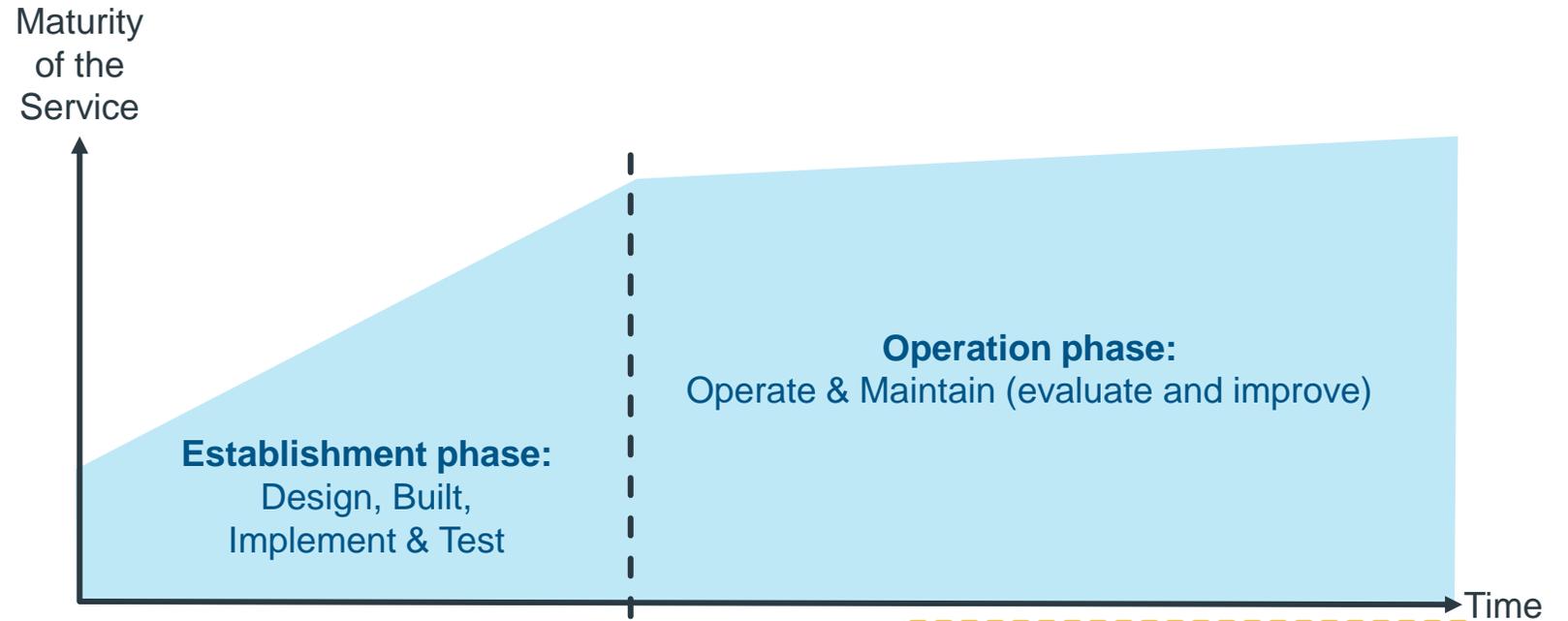


(出所) DARWIN EU <https://www.ema.europa.eu/en/about-us/how-we-work/big-data/data-analysis-real-world-interrogation-network-darwin-eu>



# The DARWIN EU<sup>®</sup> Initiative

- 2フェーズ 5年間のプログラム
  - 2 years: set up the service
  - 3 years: operations
- 5年間のプログラム終了後は更新あるいは再設定、またEHDS (European Health Data Space)との統合も想定
- EMAとの密な連携プログラム



	2021	2022	2023	2024	2025
--	------	------	------	------	------

<b>On Boarded Data Partners</b>	10	20	30	40	
<b># Studies</b>	4	16	61	145	145

(出所) DARWIN EU <https://www.ema.europa.eu/en/about-us/how-we-work/big-data/data-analysis-real-world-interrogation-network-darwin-eu>



# NIH National COVID Cohort Collaborative (N3C)

NIHのCOVID-19コホート研究では、複数データソースをOMOP CDMに変換しモニタリング



National  
COVID  
Cohort  
Collaborative

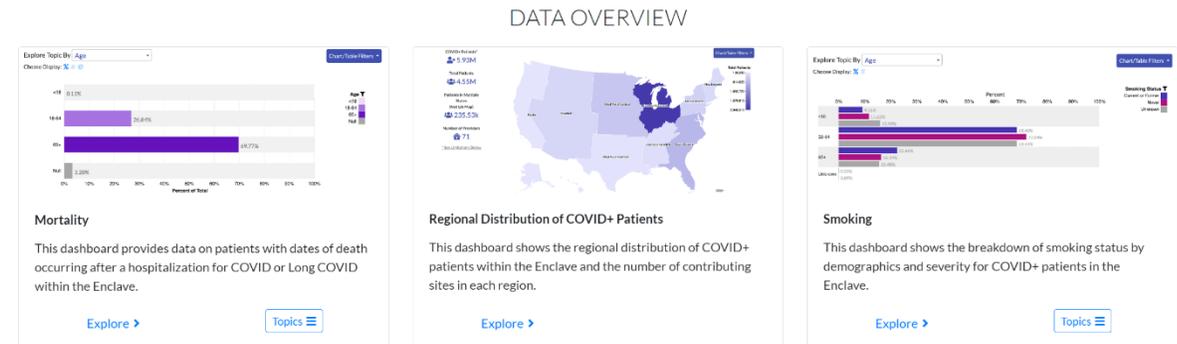
The N3C Data Enclave represents one of the largest secure collections of harmonized clinical health data in the United States.

-  **Sites:** 77
-  **Lab Results:** 10.8 billion
-  **Persons:** 18.1 million
-  **Medication Records:** 3.6 billion
-  **COVID+ Cases:** 7,123,277
-  **Procedures:** .9 billion
-  **# of Rows:** 22.7 billion
-  **Visits:** 1.2 billion
-  **Clinical Observations:** 2.0 billion

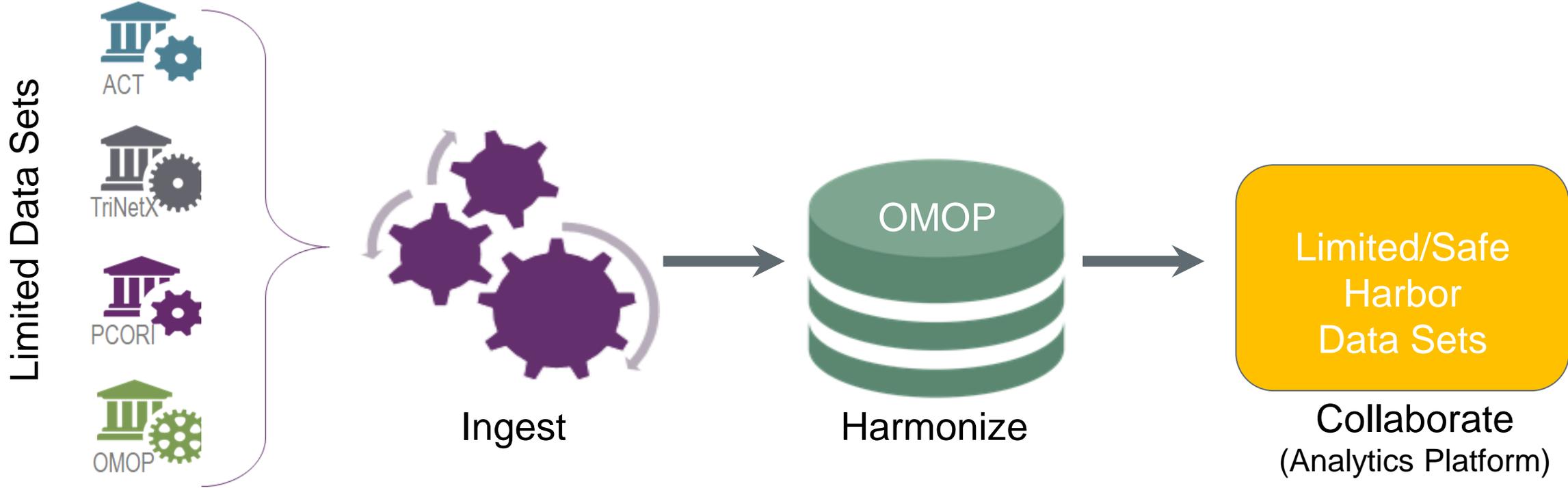
Data as of February 23, 2023

**Overview:** Consortia of distributed clinical data networks (PCORnet, OHDSI, ACT/i2b2, TriNetX)

**Goal:** Improve the efficiency and accessibility of analyses with COVID-19 clinical data, expand ability to analyze and understand COVID, and demonstrate a novel approach for collaborative pandemic data sharing



# Implementing a National COVID-19 Surveillance Program



1

Data partnership & governance

2

Phenotype & Data acquisition

3

Data ingest & harmonization

4

Collaborative analytics & FAIR Sharing/Credit





# FEEDER-NET (Federated E-health Big Data for Evidence Renovation Network)

OMOP CDMを用いた韓国のRWEネットワーク

## Introduction

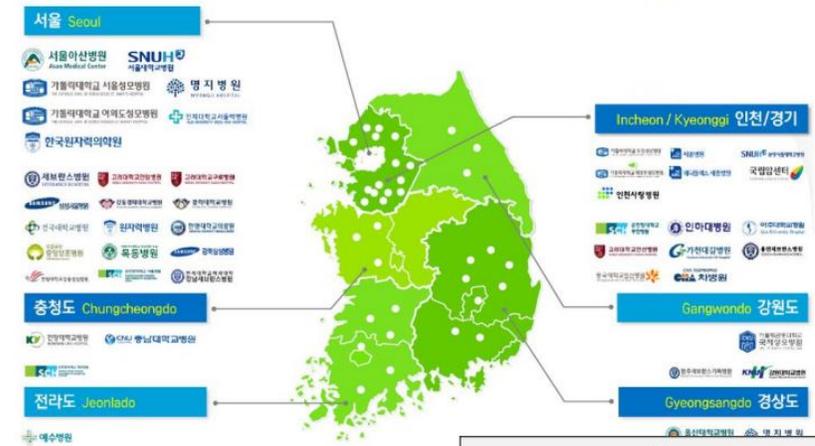
- The Federated E-Health Big Data for Evidence Renovation Network (FEEDER-NET) project was initiated on 2018 (for 3 years) with \$10 million budget from the Ministry of Trade, Industry & Energy (MOTIE) of South Korea.
- On 2019, new projects, named FEEDER-NET+, for incorporating more hospitals and advancing/sophisticating the FEEDER-NET platform will start April 2019 for 4 years

## FEEDER-NET+ Goals

- To build a CDM-based bio-health big data network
- To build a bio-health evidence sharing platform
- To foster collaborative research and encourage industry to develop business services
- **Ultimately to create a sustainable bio-health big data ecosystem for research and industry**

## FEEDER-NET+ Data Network

- 61 hospitals and data partners joined into the FEEDER-NET+ all over the country. (including about 70% of tertiary hospitals)
- 12 hospitals' EMRs were converted to CDM
- We have bi-monthly data partner's summit meeting since March 2017.



Summary of Data Network	
Hospitals	Tertiary hospital: 27 General hospital: 34; Clinic: 1
# of Pts	97,367,192

(出所) FEEDER-NET+ (Federated E-health Big Data for Evidence Renovation Network) in Korea [https://www.ohdsi.org/wp-content/uploads/2019/09/Seongwon-Lee\\_FEEDER-NET-in-Korea\\_poster2019-1.pdf](https://www.ohdsi.org/wp-content/uploads/2019/09/Seongwon-Lee_FEEDER-NET-in-Korea_poster2019-1.pdf)

# Oncology RWD Network

次頁以降詳述

## HARMONY Alliance

- 血液腫瘍
- 19か国、120+団体
- 10万人以上の患者レコードがOMOP CDM標準化(23年6月時点)
- 2017-2023 (IMI, BD4BO)

## ASH RESEARCH COLLABORATIVE\*1

- 血液腫瘍 (SCD, MM)
- SDC: 500例、MM: 1,000例程度がNovartisとAmgenより登録(19年12月時点)
- 2018-

## PIONEER

- 前立腺がん
- 9か国、32団体
- 21 DBがOMOP CDM標準化(22年9月時点)
- 2017-2023 (IMI, BD4BO)

## DigiCore

- オンコロジー
- 16か国、36施設 +2 cancer networks
- 2021-
- 2022-ネットワークの取組み(DigiONE)

## CAREL\*2

- オンコロジー
- Federated Networkの構築を目指した、檀国大学校とカトリック大学校によるパイロット研究およびツール開発

Big Data Platform (従来型)

Federated Network

\*1 Wood et al., Blood Adv (2021) 5 (23): 5429–5438. <https://doi.org/10.1182/bloodadvances.2021005902>

\*2 Park et al., Technol Cancer Res Treat . 2023; 22: 15330338221149262.. <https://doi.org/10.1177/15330338221149262>

# DigiONE (the DIGital Oncology Network for Europe)

EUにおける最先端の連合型研究基盤構築の取組

<https://doi.org/10.1038/s41591-023-02715-8>

## A federated learning system for precision oncology in Europe: DigiONE

Piers Mahon, Ismini Chatzitheofilou, Andre Dekker, Xosé Fernández, Geoff Hall, Aslaug Helland, Alberto Traverso, Cedric Van Marcke, Janne Vehreschild, Gennaro Ciliberto & Giovanni Tonon

Check for updates

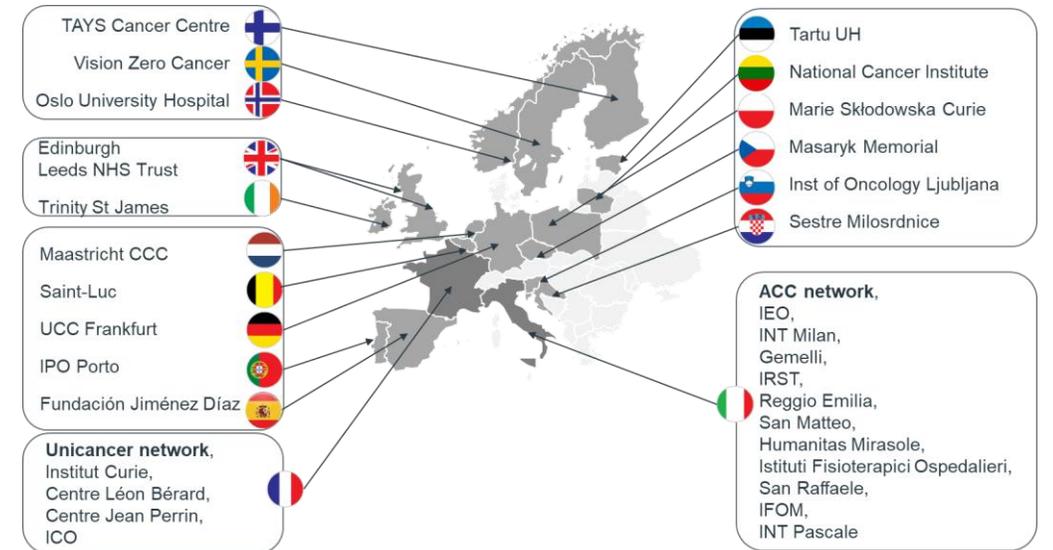
DigiONE is a pilot European learning health system in precision oncology that aims to identify optimal cancer treatments by learning from every patient, not just those in trials, through privacy-preserving interrogation of their standardized routine electronic health records.



Every cancer is different, depending on its particular molecular subtype. Cancer treatments will only be effective on certain subtypes. The optimal matching between patients and treatments will require granu-

explore new indications for approved targeted and immunotherapies

## Members



• Mahon et al., Nature Medicine 2024, 30, 334–337 <https://doi.org/10.1038/s41591-023-02715-8>

# DigiONE

2022年よりDigiCore6施設でFederated Network Studyを目的としたプログラム(DigiONE)を開始

## DigiONE (Digital Oncology Network for Europe)



1. 汎用性の高い、最小限のオンコロジー研究用データセット  
**Minimal Essential Description of Cancer: MEDOC**の定義  
(フランスのOSIRISとCancer OMOPがベース)
2. GDPRに準拠し、相互運用可能で、質の高い**MEDOCデータセット**を  
**6施設で整備**
3. 適切なデータガバナンス、個人情報保護、契約に基づいて、**上記6施設間でのFederated Network Study**(各施設における解析結果を統合したコホート研究)を実施
4. 日常診療における遺伝子診断等のデータの追加  
(PDFファイルを変換)
5. POCとしての研究成果の発表

# まとめ

- 医療・ヘルスケアデータ（RWD）の連携・利活用における新たな手法として、**国際標準データモデル（RWD CDM）**を活用した**連合解析**の概要と先進事例についてご説明いたしました。
- RWD CDMの活用は「**解析の効率化**」と「**プライバシー保護**」の課題解決に寄与することが期待されており、諸外国においても先行事例が増えてきています。
- 本国が今後**バイオ関連産業の育成**において**国際競争力を獲得**していく目的においても、国内外に分散しているデータを**効率的、効果的に連携**していくことにより、これまで以上の“ビッグデータ”の活用を推進していくことが不可欠であると考えます。



## アジェンダ

- 医療・ヘルスケア分野におけるデータ連携・利活用の現状と課題
- データ連携・利活用に関する新たなコンセプト
- データ連携・利活用に関する先端事例
- Q&A