# The State of Federated Health Data Networks Globally in 2024

**Michael Briganti, Valerie van Baalen, Eva-maria Didden, Monika Brand**

Johnson & Johnson Innovative Medicine, Research & Development

## Background

Federated Data Networks (FDN) facilitate sharing and analyzing data through a network made up of different organizations and connected through a central hub. Definitions of FDNs vary depending on the data sources included or the technical infrastructure, however there are several core required components. All data partners are in full control of their data and responsible for patient privacy and consent. Patient-level data must remain de-centralized and firewall-protected, and only aggregate information may be shared with other partners directly or through the central hub. Additionally, data is harmonized into a common data model so that a program/script provided by the central hub can be ran by each partner.

FDNs utilize rapidly developing methodology which could benefit from standardized terminology. As a result, this leads to a decrease in discoverability and a lack of transparency. To date there have been no published reviews of the existing health-related FDN landscape. The goal of this poster is to give a high-level overview of the existing FDNs in a rapidly evolving space.

## Methods

We searched PubMed on the keyword “Federated Data Network” for the past 5 years (between April 22, 2024 and April 22, 2019) and identified any publications which described an FDN and included health data. We supplemented these findings by using an internal Large Language Model to identify additional papers based on a similarity score to a subset to key papers identified in the initial search. We screened the identified papers that described FDNs of interest for any other FDNs mentioned or referenced within the publication. Finally, we asked four epidemiology colleagues with FDN research experience if any networks might be missing from our list of found FDNs, as we also wanted to include known FDNs even if they did not have publications.

## Results

We identified 43 FDNs. Many of these FDNs contained data on multiple diseases or were populated based (n = 21). Thirteen FDNs were disease-specific, most commonly dementia (n = 6). Four networks (not included in the 13) were focused on various cancers. In several cases it was not possible to discern the network's disease focus due to lack of publicly available information.

Most FDNs do not publicly state what the underlying source data includes. Of those that do specify, the most common data source is electronic health record (EHR) data (n = 18). Of these, half included data from a secondary source, such as EHR + Insurance Claims or EHR + Biobank (n = 9).

The United States has the largest country specific FDN presence (n = 13). Europe had 15 FDNs, but many were country specific (i.e. UNICANCER's Consore in France) or specific to a particular region (i.e. FederatedHealth among Nordic countries). Three networks were identified in Asia, two in Australia, and zero were identified in South America or Africa.

Several FDNs were global. A global network could be large in scale (TriNetX has 250+ million patients) or smaller (PHederation has 8000+ patients). Networks that might at first glance appear to be smaller in scale than a global network were actually quite large in scale in the United States. As an example, an FDN focused on the Chicago area (CAPriCORN) included 13 million patients across 1000 healthcare sites.

Most FDNs do not specify an underlying CDM. Of those that do, OMOP CDM is the most common. Other FDNs specify they use a custom CDM (i.e. PEDSNet), however PEDSNet CDM isa modified version of the OMOP CDM. A few FDNs used the i2b2 CDM, however they are successors to each other and the decision to use i2b2 may be influenced by legacy reasons.

The network's purpose will influence CDM choice. More recently established FDNs seem to be choosing OMOP as the CDM because of it's prevalence and the availability of a robust community which lends to sustainability of the OMOP CDM going forward. A historical overview and comparison of healthcare CDMs is described in greater detail elsewhere (1).

## Conclusions

By enhancing data sharing collaboration, FDNs may provide new evidence that can lead to improved care for patients due to more efficient collaboration and data generation over time (2). FDNs also allow researchers to explore the differences in treatments that are given across countries (3).

In certain cases (i.e. a rare disease area) a FDN may be necessary when a single data partner would not have enough data on their own to constitute a large enough sample size. In all cases, FDNs should decrease the time needed to obtain required sample size (3). Since networks operate in a federated approach, no patient data leaves the original data owner, which helps address growing concerns surrounding patient privacy and data security.

As the number of sites in an FDN grows, this may result in a more diverse sample of patients, leading to more generalizable results and stronger evidence generation (3). However, this may have the unintended consequence of including duplicate patients from different data partners.

From our search, some networks were self-described as a “Decentralized Data Network”, “Secure Data Network”, or “Collaborative Data Network.” Due to the vagueness of these terms, future networks meeting FDN criteria would benefit the field by using the term FDN, or even extending to Federated Health Data Network (FHDN) to indicate an FDN specific to health data. FDN technology and methodologies is borrowed from fields outside healthcare, and the specific FHDN may help future researchers to narrow their focus to relevant networks and specific challenges related to health/patient data.

This high-level overview of FDNs describes the diversity present among the networks, while sharing core federated data principles. This landscape may serve as a first step towards understanding what FDNs exist, and highlighting the need for standard FDN terminology to drive the field forward.

## Tables/Figures

**Figure 1. Countries with known FDNs**

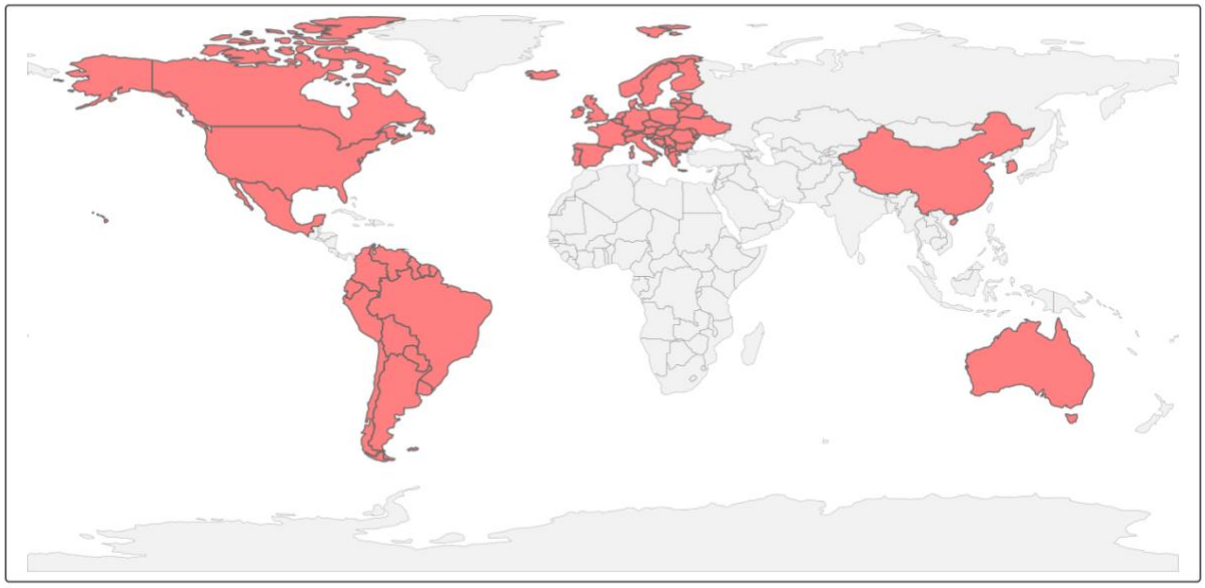

**Figure 2. Countries with known FDNs, by count**

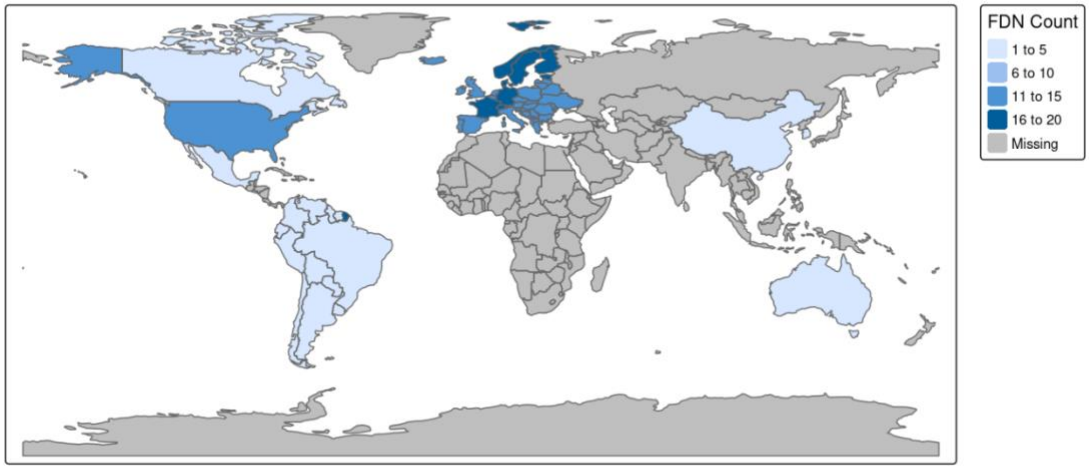

**Table 1. List of Identified FDNs (n = 43)**

| Federated Data Network | Region | Disease (If Listed) | CDM | Website (Or where referenced if no website found) |
|---|---|---|---|---|
| Canadian Distributed Infrastructure for Genomics (CanDIG) | Canada | | | https://www.distributedgenomics.ca/ |
| EHDEN | Europe | Multiple Diseases | OMOP | https://www.ehden.eu/ |
| Accessible Research Commons for Health (ARCH) Clinical Data Research Network | US | Multiple Diseases | | https://www.pcori.org/research-results/2015/accessible-research-commons-health-arch |
| Text Information Extraction System (TIES) Cancer Research Network | US | Cancers | | https://aacrjournals.org/cancerres/article/75/24/5194/606950/A-Federated-Network-for-Translational-Cancer |
| PCORnet | US | Multiple Diseases | PCORnet® CDM | https://pcornet.org/ |
| Psoriasis and Psoriatic Arthritis Integrated Research Data Network (PIONEER) | Global | Psoriasis and Psoriatic Arthritis | | https://pubmed.ncbi.nlm.nih.gov/22630572/ |
| Harvard's four-site Shared Health Research Information Network (SHRINE) | US | Multiple Diseases | | https://catalyst.harvard.edu/shrine/ |
| Joint Imaging Platform (JIP) of the German Cancer Consortium (DKTK) | Germany | Cancers | | https://jip.dktk.dkfz.de/jiphomepage/ |
| TriNetX | Global | Multiple Diseases | | https://trinetx.com/ |
| UNICANCER's Consore | France | Cancers | OSIRIS | https://pubmed.ncbi.nlm.nih.gov/38397680/ |
| Darwin EU | Europe | Multiple Diseases | OMOP | https://www.darwin-eu.org/ |

| | | | | |
|---|---|---|---|---|
| IBM Watson Health Data Exchange | Global | Multiple Diseases | | https://uk.newsroom.ibm.com/2017-01-11-IBM-Watson-Health-Announces-Collaboration-to-Study-the-Use-of-Blockchain-Technology-for-Secure-Exchange-of-Healthcare-Data |
| ENACT | US | Multiple Diseases | i2b2 | https://www.enact-network.us/ |
| Accrual to Clinical Trials (ACT) Network | US | Multiple Diseases | i2b2 | https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6241502/ |
| HMORN, HMO Research Network | US | Multiple Diseases | | https://www.pcori.org/assets/5-Larson-Slides-HMORN.pdf |
| PEDSnet | US | Multiple Diseases | PEDSnet CDM, based on OMOP | https://pedsnet.org/ |
| Health Data Research UK | UK | Multiple Diseases | | https://www.hdruk.ac.uk/ |
| Scalable Architecture for Federated Translational Inquiries Network (SAFTINet) | US | | | https://pubmed.ncbi.nlm.nih.gov/25848567/ |
| Sentinel | US | Multiple Disease | Sentinel CDM, includes mother-infant linkage table | https://www.sentinelinitiative.org/ |
| Swiss Personalized Health Network (SPHN) | Switzerland | Multiple Diseases | | https://sphn.ch/ |
| FeederNet | South Korea | Multiple Diseases | OMOP | https://www.ohdsi.org/2019-us-symposium-showcase-16/ |
| Big Multiple Sclerosis Data (BMSD) Network | Europe | Multiple Sclerosis | | https://bigmsdata.org/ |
| COINSTAC | Global | | | https://coinstac.org/ |
| EU-ADR | Europe | Multiple Diseases | | https://pubmed.ncbi.nlm.nih.gov/19745234/ |

| | | | | |
|---|---|---|---|---|
| MammoGrid | Europe | Breast Cancer | | https://pubmed.ncbi.nlm.nih.gov/14663986/ |
| FederatedHealth | Norway, Sweden, Denmark, Finland, Estonia | | | https://www.nordicinnovation.org/programs/federatedhealth-nordic-federated-health-data-network |
| Biogrid Australia | Australia | Multiple Diseases | | Biogrid.org.au |
| Medical Informatics in Research and Care in University Medicine | Germany | Multiple Diseases | | https://pubmed.ncbi.nlm.nih.gov/30016814/ |
| HMO Cancer Research Network | US | Cancers | | https://deainfo.nci.nih.gov/advisory/bsa/archive/bsa0611/presentations/Ballard-Barbash.pdf |
| HONEUR | Europe | Hematological | OMOP | https://portal.honeur.org/ |
| The Chicago Area Patient Centered Outcomes Research Network (CAPriCORN) | Chicago Area | Multiple Diseases | PCORnet® CDM | https://www.capricorncdrn.org/ |
| CREDO (Clinical Research & Epidemiology Data Observatory) | China | | | Word of mouth, can't find website or publications |
| PHederation | Global | Pulmonary Hypertension | OMOP | https://portal.phederation.org/ |
| ESFURN | Europe | Study of FGFR in Urothelial Neoplasms | OMOP | https://portal.esfurn.org/ |
| LupusNet | Global | Lupus | OMOP | https://portal.lupusnet.org/ |
| Atropos Health | Global | Multiple Diseases | | https://atroposhealth.com |
| Dementia's Platform UK | UK | Dementia | | www.dementiasplatform.uk |
| Dementia's Platform AU | Australia | Dementia | | Referenced on Dementia's Platform UK website |
| Dementias Platform Korea | South Korea | Dementia | | Referenced on Dementia's Platform UK website |

| | | | | |
|---|---|---|---|---|
| Global Alzheimer's Association Interactive Network | Global | Alzheimer's | | https://www.gaain.org/ |
| Alzheimer's Disease Data Initiative | Global | Alzheimer's | | https://www.alzheimersdata.org/ |
| The EU Joint Programme Neurodegenerative Disease Research | Europe | Neurodegenerative diseases | | https://neurodegenerationresearch.eu/ |
| Opioid Overdose Network | US | | | https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9243402/ |

**References**

1. Weeks J, Pardee R. Learning to Share Health Care Data: A Brief Timeline of Influential Common Data Models and Distributed Health Data Networks in U.S. Health Care Research. eGEMs (Generating Evidence & Methods to improve patient outcomes). 2019 Mar 25;7(1):4.

2. van Baalen V, Didden EM, Rosenberg D, Bardenheuer K, van Speybroeck M, Brand M. Increase transparency and reproducibility of real-world evidence in rare diseases through disease-specific Federated Data Networks. Pharmacoepidemiology and Drug Safety [Internet]. 2024 Apr 1 [cited 2024 Jun 12];33(4):e5778. Available from: https://pubmed.ncbi.nlm.nih.gov/38556812/)

3. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology [Internet]. [cited 2024 Jun 12]. Available from: https://encepp.europa.eu/document/download/266ef169-949a-4330-9c10-99fe40e1e3ea_en?filename=methodologicalGuideAuthors.pdf&prefLang=nl