Building the OHDSI Evidence Network – A Global, Open, Federated Collaboration

Clair Blacketer[1,2,4], Haeun Lee[1,8], Benjamin Martijn[1,8], Evanette Burrows[1,4], Patricia Mabry[1,10], Deran McKeen[1,10], Sam Patnoe[1,10], Elizabeth Grossman[1,10], Ben Gerber[1,5], Pantelis Natsiavas[1,6], Aamirah Vadsariya[1,7], Hanieh Razzaghi[1,9], Paul Nagy[1,8]

1. OHDSI Collaborators, Observational Health Data Sciences and Informatics (OHDSI), New York, NY, USA
2. Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, NL
3. Department of Biomedical Informatics, Columbia University, New York, NY, USA
4. Johnson & Johnson, Raritan, NJ, USA
5. Department of Population and Quantitative Health Sciences, UMass Chan Medical School, Boston, MA, USA
6. Institute of Applied Biosciences, Centre for Research & Technology Hellas, Thessaloniki, GR
7. Clinical Informatics Center, University of Texas Southwestern Medical Center, Dallas, TX, USA
8. Johns Hopkins University, Baltimore, MD, USA
9. Applied Clinical Research Center, Children's Hospital of Philadelphia, Philadelphia, PA
10. Health Partners Institute, Bloomington, MN, USA

Background

Federated healthcare data networks have emerged as a powerful solution for enabling real-world evidence (RWE) generation across diverse populations while respecting data privacy and institutional autonomy. Prominent initiatives such as PCORnet and the FDA's Sentinel System have demonstrated the value of shared analytical methods and centralized governance for post-marketing surveillance, pragmatic trials, and comparative effectiveness research(1,2). However, these federated networks present new challenges that can limit their scale. These challenges can include the use of proprietary software, requiring formal contractual participation, being overly dependent on single limited funding opportunities. As the demand for timely, reproducible, and generalizable evidence continues to grow globally, there is an urgent need for flexible, open-source solutions that promote inclusivity and scale. European initiatives such as the European Health Data and Evidence Network (EHDEN) and the Data Analysis and Real World Interrogation Network (DARWIN EU) have seen success in this space and OHDSI is building on that framework through the establishment of an open-source, global, federated network(3,4).

The OHDSI Evidence Network was launched in 2024 to meet this need(5). It offers a fully open-source, international, and federated model for collaborative research using the OMOP Common Data Model. Unlike traditional models, participation in the OHDSI

Evidence Network does not require formal data use agreements or centralized data transfers. Instead, partners retain full control over their data and contribute only summary-level information when they choose to opt into a given study or data characterization effort. This lightweight framework enables broader participation while still supporting high-quality, large-scale evidence generation.

Objective

This abstract describes the development and first-year experience of the OHDSI Evidence Network, focusing on key goals, governance strategies, partner engagement efforts, and lessons learned. Our objective was to build a global, community-driven infrastructure for federated research that is transparent, inclusive, and scalable. We aim to share the principles and practical steps that supported successful implementation, as well as the challenges encountered, and insights gained during the first year.

Methods

The OHDSI Evidence Network was designed to be maximally collaborative while ensuring scientific rigor and adherence to local governance requirements. Participation was voluntary and based on mutual trust, with no formal agreements required. Data partner organizations with OMOP CDM instances were invited to join by expressing interest and running the Database Diagnostics R package locally(6). This tool, developed by the OHDSI community, generated a consistent set of summary statistics, known as a DbProfile, that allows the Coordinating Center to evaluate alignment with all potential studies without access to person-level data.

The project began with the "Save Our Sisyphus" (SOS) pilot in 2023, which tested initial engagement strategies and provided a structured opportunity for partners to explore network participation. The pilot surfaced several key ideas: (1) a desire from partners to learn from others in the community with similar data, (2) the need for clear study protocols to support internal governance reviews, and (3) the benefit of low-burden, transparent communication channels.

Building on these findings, the network adopted a distributed governance model. Each partner retained full authority over local approvals and could opt into any study they were interested in and that Database Diagnostics deemed them a potential fit. To encourage collaboration, the Evidence Network working group was established, as was weekly office hours, a steering committee to guide strategic decisions, and a Data Partner monthly call to share feedback and priorities. Study leads were encouraged to co-develop protocols with interested partners, and partners were invited to co-author resulting publications and participate in protocol reviews.

Results

After one year, the OHDSI Evidence Network includes 26 partner organizations contributing 44 databases across North and South America, Europe, and Asia. The partner institutions comprise academic medical centers, pharmaceutical companies, clinical research organizations, and clinical registries.

Partner feedback indicated that the minimal-burden participation model was a major strength. The lack of required contracts, reliance on existing OHDSI tools, and respect for local data governance processes lowered barriers to entry and accelerated onboarding. Several sites used their Data Diagnostics reports to identify and improve areas of their ETL (Extract-Transform-Load) process, strengthening the quality of their OMOP CDM instance in the process.

However, challenges also emerged. Some partners were initially hesitant to share even aggregate statistics publicly, due to institutional policies, uncertainty about secondary use, local IRB requirements and human subjects protection. To address this, all DbProfiles submitted to the OHDSI Coordinating Center are held securely, with only two key collaborators allowed access to the raw data.

A key success of the Evidence Network was the use of the anonymized, aggregated statistics submitted by the partners to inform person- and record-count values in the publicly available ATLAS instance[1]. This has been extremely useful in multiple contexts to help the community understand the overall depth and breadth of concept representation across the network.

Another challenge was communication. To keep everyone informed, we implemented a tiered notification system: study leads provided brief study descriptions and criteria, which the Coordinating Center then used to notify only those partners whose data characteristics were likely to inform the study. We also implemented weekly office hours to answer any immediate questions of the partners or study leads. This improved responsiveness and relevance.

The network supported ~16 studies during its first year, including a community-led study-a-thon. This event demonstrated the value of the federated approach: feasibility analyses were completed within days across all databases using standardized tools and protocols. Partners appreciated the clear expectations, low overhead, and opportunity to contribute to high-impact research.

[1] https://atlas-demo.ohdsi.org/#/

Conclusions

The first year of the OHDSI Evidence Network demonstrates that a community-led, open-source, and fully federated model can support global real-world evidence generation without requiring formal contracts or centralized data. We learned that:

1. **Trust and Transparency are Essential:** Regular communication, inclusion in decision-making, and clear documentation promote a culture of collaboration.
2. **Low-Burden Participation Encourages Engagement:** A streamlined onboarding process, familiar tools, and respect for data ownership enabled broad participation.
3. **Shared Tools Enable Shared Learning:** Running diagnostics not only helped identify fit-for-use data sources but also supported internal quality improvement.
4. **Decentralized Governance is Feasible:** By letting partners choose when and how to participate, we preserved autonomy while achieving scale.

The OHDSI Evidence Network now serves as a living, evolving example of sustainable collaboration across institutional and national boundaries. As we move into year two, we aim to expand partner participation, produce more informative partner-study matching diagnostics, and support increasingly complex network research. The lessons learned may serve as a blueprint for other open federated initiatives seeking to accelerate responsible, high-quality health research.

References

1. Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, Brown JS. Launching PCORnet, a national patient-centered clinical research network. Journal of the American Medical Informatics Association. 2014;21(4):578–82.
2. Ball R, Robb M, Anderson SA, Dal Pan G. The FDA's sentinel initiative--A comprehensive approach to medical product surveillance. Clin Pharmacol Ther. 2016 Mar;99(3):265–8.
3. Rijnbeek PR, Schuemie MJ, van der Lei J, al et. The European Health Data & Evidence Network (EHDEN): building a federated network to accelerate research. European Journal of Epidemiology. 2020;35(6):613–7.
4. European Medicines Agency. Darwin EU [Internet]. [cited 2024 Sep 26]. Available from: https://www.darwin-eu.org/
5. OHDSI Evidence Network [Internet]. [cited 2025 Jun 17]. Available from: https://ohdsi.github.io/EvidenceNetwork/

6. DbDiagnostics [Internet]. Observational Health Data Sciences and Informatics; 2023 [cited 2023 Apr 19]. Available from: https://github.com/OHDSI/DbDiagnostics