



Joining the OHDSI Evidence Network



Agenda

Introduction/recap of last week – Paul (5 min)

Evidence Network Details - Clair (3 min)

- What you need to have in place to join

IRB language and strategies – Ben (5-8 min)

Evidence Network Details, cont - Clair (5 min)

- What you need to do to join

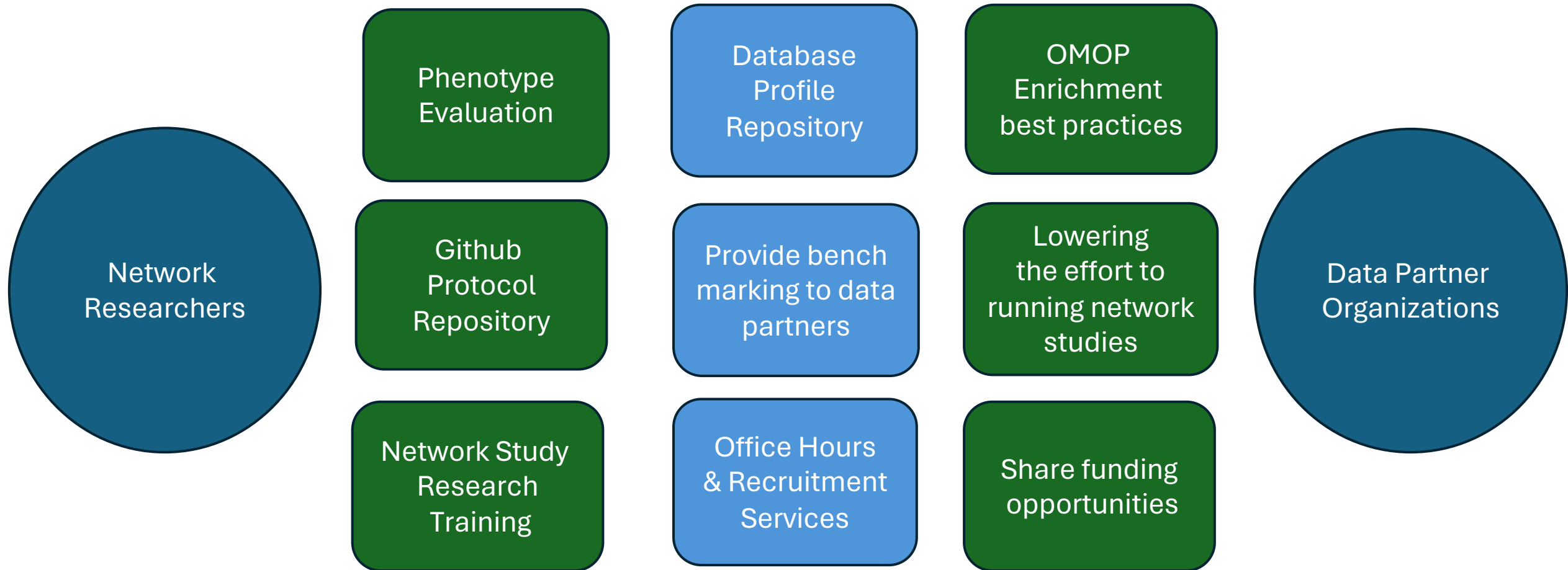
- What will be shared publicly vs held privately

How to join the network demo – Clair (15 min)

Questions (10 min)



Catalyzing activity for generating evidence



Data Partner Organizations (DPO)



An institution



That owns or licenses data



That has been converted to the OMOP CDM v5.3+



Willingness to generate evidence and participate in network studies



Joining the Network as a DPO

What do you need to have in place?

What do you need to do?

What information will be held privately?

What information will be public?

What you need in place

To join the network as a DPO

- Observational health data standardized to the OMOP CDM v5.3 or higher
- Data held in a relational database accessible by the organization
 - List of supported SQL environments here:
<https://ohdsi.github.io/SqlRender/articles/UsingSqlRender.html#translation-to-other-sql-dialects>
- Approval from governance entity (i.e. IRB) to share metadata and concept counts with the OHDSI Coordinating Center (OCC)
 - Note: It is up to each individual DPO as owner or licensee of data to ensure all appropriate governance requirements are followed.
- The ability to run the DbDiagnostics R package against the data



Data Network Protocol – IRB Guidance



1. Define *“OHDSI Study”* or *“OHDSI Network Study”*
 - *Asking for more clarity about what it means to be an “OHDSI Data Partner”*
2. OHDSI Study Protocol Document is a supplemental document – Institution-specific protocol is still required
3. Precisely list every place the data/results will sit.
 - Row-level: OMOP DB Server (managed by IT services)
 - Aggregate-level: platform where DbDiagnostics is run, Columbia Coordinating center, peer-reviewed publication appendix
4. *The degree of sensitive data removal*
 - *Full PHI, Limited-data set, De-identified, Aggressively de-identified*



Data Network Protocol – Study Objectives

Provide a BRIEF statement of your research question and plan:

This study aims to summarize data profile counts and concept distribution patterns across different OHDSI datasets and investigate their impact on study feasibility to facilitate cross-institutional and network studies. The **aggregate statistics** collected will inform the similarities between federated network data partners, what a typical OHDSI dataset looks like, and allow for rapid feasibility assessments for across the network. The concept counts will be generated from our aggressively **de-identified OMOP dataset** of JHM patients. This profile will be kept by the Columbia coordinating center.



Data Network Protocol – Methods

Describe the study design and the database(s) that will be utilized, including the specific sources from which you will collect data or samples:

This study will not involve consenting/recruitment/remuneration/in-person activities as it will only perform a **secondary analysis** of existing clinical data. **No chart review** will be performed at JHU. Upon IRB approval we plan to execute pre-defined queries and **summary statistics** using the DbDiagnostics R package, available on GitHub (<https://github.com/ohdsi/DbDiagnostics>), against local databases to generate summary statistics. The study team will also analyze aggregate-level descriptive information about the data and develop network-based characterizations.



Data Network Protocol – Dissemination

Please describe any plans for data dissemination, including but not limited to, any plans to deposit data in external or internal repositories or share individual (person)-level data for publication purposes.

This study has **no plans nor requirements for sharing individual-level data**. Any data shared/disseminated will only consist of aggregate or **summary-level data and results**. Specific summary statistics that will be generated and disseminated in this analysis are included in Section 4, Study Statistics (below). The summary statistics generated will be kept by the OHDSI Coordinating Center for reuse to support network research.

What you need to do

To join the network as a DPO

- Run the [DbDiagnostics package](#) `executeDbProfile` function to generate metadata and high-level concept counts about each data source submitted to the network
 - The aggregate information gathered by the package is listed here:
<https://ohdsi.github.io/DbDiagnostics/articles/SummaryStatistics.html>
 - If the [Achilles](#) package was run previously and the results stored this step will take approximately 15-30 minutes, depending on the environment
 - If the Achilles package was not run previously or if the results were not stored this step will take approximately 1-8 hours, depending on the environment.
- Send the resulting information to the OCC via SFTP. Please contact evidencenetwork@ohdsi.org for the key file when you are ready

What information will be held privately...

Data source and site-specific metadata will be held securely at the OHDSI Coordinating Center at Columbia University and **will not** be shared openly.

	cdm_source_name character varying (255) 🔒	statistic text 🔒	concept_id character varying (255) 🔒	person_count integer 🔒
1	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	28060	10
2	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	77074	10
3	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	81151	10
4	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	132797	10
5	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	201826	10
6	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	254761	60
7	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	255848	20
8	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	257012	20
9	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	260139	40
10	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	312437	10
11	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	313217	10
12	Synthea_v54_OHDSI_Example	Number of persons with at least one condition occurrence, by condition_concept_id	314754	10

... vs What will be shared publicly

Only aggregate concept counts and the total number of data sources with a record of each concept will be shared as a public resource.

	statistic text	concept_id character varying (255)	total_person_count numeric	num_data_sources bigint
1	Number of data sources with at least one condition occurrence, by condition_concept...	0	118600921	26
2	Number of data sources with at least one condition occurrence, by condition_concept...	132238	1651	13
3	Number of data sources with at least one condition occurrence, by condition_concept...	132258	6710	20
4	Number of data sources with at least one condition occurrence, by condition_concept...	132277	573790	20
5	Number of data sources with at least one condition occurrence, by condition_concept...	132321	26210	19
6	Number of data sources with at least one condition occurrence, by condition_concept...	132333	88330	20
7	Number of data sources with at least one condition occurrence, by condition_concept...	132342	342294	28
8	Number of data sources with at least one condition occurrence, by condition_concept...	132344	1460714	33
9	Number of data sources with at least one condition occurrence, by condition_concept...	132356	198620	18
10	Number of data sources with at least one condition occurrence, by condition_concept...	132391	414550	21
11	Number of data sources with at least one condition occurrence, by condition_concept...	132392	47668	27
12	Number of data sources with at least one condition occurrence, by condition_concept...	132393	2106730	27
13	Number of data sources with at least one condition occurrence, by condition_concept...	132397	1310	3



Let's Generate a DbProfile!





What happens after you send your DbProfile to the OCC?

1. Your organization will receive an @ohdsi.org account i.e. VA@ohdsi.org to be used to notify you of potential network studies and other internal communications
2. Your organization will be listed as an OHDSI Evidence Network DPO on the OHDSI.org website
3. You will receive a report from the OCC putting your data source in the context of network*

*Once there are enough participating data partner organizations



Questions?

We are here to help!

We are hosting office hours every **Friday from 9am-10am EST** in the Evidence Network teams channel. Fill out this [form](#) and choose 'Evidence Network' to join.

Email us at evidencenetwork@ohdsi.org