

# Guidance for Communication of the OHDSI Network Study Approach with Institutional Review Boards

<sup>1</sup>Ben Martin, PhD; <sup>1</sup>Mary Grace Bowring, MPH; <sup>1</sup>Paul Nagy, PhD;  
<sup>1</sup>Johns Hopkins School of Medicine, Baltimore, MD, USA.

## Background

The Observational Health Data Science and Informatics (OHDSI) collaborative provides a unique approach, set of open-source tools, and open-science community of support to perform observational health research across distributed networks of real-world data in the form of “network studies.” Due to layers of data security and regulatory requirements around protected health information (PHI), institutional review boards (IRB) have a routine obligation to ensure proper handling of PHI involved in studies using real-world data (RWD). In cases of secondary use of data collected during clinical operations for research purposes and for which consent is not required, studies are considered “exempt research” and do not fall under the Department of Health and Human Services Policy for Human Subjects research per the 2018 amendments to the Code of Federal Regulations (45 CFR §46.104). The revised regulations combined with intra-institutional coordination amongst researchers using secondary RWD present an opportunity to minimize the time between study conception and data analysis.

## Objectives

This abstract aims to provide a guide for clear communication of the federated methods and data governance considerations of OHDSI network studies to local Institutional Review Boards.

## Methods

A review of OHDSI network study documentation, educational media, and researcher experience was refined to provide common language and key points to communicate with local IRB personnel when submitting study protocols for review at potential data partners’ local institutions.

## Results

The key points relevant to local IRB reviewers when participating in a network study are outlined below. Specific excerpts from the Book of OHDSI that are helpful in supporting communication within individual institution IRBs and governance teams are listed in Table 1.

IRB guidance for participating data partners working with an external network study lead:

- The OMOP CDM has mapped patient identifiers and should be considered a limited data set.
- SANT methods are available to conduct randomizing date shifts for each patient while preserving temporal relationship.
- Each participating network site must register the study with their local IRB, if required by institutional governance.
- The study protocol document with rationale, background, objectives, methods, data analysis plan, and measures for protection of human subjects are published publicly on the OHDSI Studies GitHub under the corresponding study folder at <https://github.com/ohdsi-studies>.

- Row level patient data is never shared across separate institutions participating in a network study. Results are generated by aggregating summary statistics calculated at each site individually before combining.
- At each data evaluation stage prior to generation of results (i.e., data diagnostics, phenotype evaluation, and study diagnostics) only aggregate evaluations are combined. Patient-level data is never shared between data partners at any point during the network study process.
- Each network site study team will have to work with their local IT team to install any dependent R packages needed to execute the OHDSI network study packages. Ensure this is permitted by under local data security and governance rules.
- Check with your local governance team to ensure that no agreements need to be executed for sharing the aggregated results.

**Table 1. Language from OHDSI documentation for clear communication with IRBs**

Book of OHDSI Excerpts
“In each study, data remains at the site behind a firewall. No patient-level data pooling occurs across network sites. <b>Only aggregate results are shared.</b> ”
“ <b>OHDSI network studies do not share patient-level data.</b> In other words, patient level data from different sites is never pooled in a central environment. Study packages create results files designed to be aggregate results (e.g. summary statistics, point-estimates, diagnostic plots, etc.) and do not share patient-level information.”
“Each site must perform its own due diligence to ensure the study protocol is approved and authorized to be executed on the local CDM, if required. Data analysts may need to enlist assistance from the local IT team to enable appropriate permissions to run the study.”
“There are a few components that make OHDSI research unique... All documentation, study code and subsequent results are made publicly available on the OHDSI GitHub. Investigators must create and publish a public study protocol detailing the scope and intent of the analysis to be performed. Investigators must create a study package (typically with R or SQL) with code that is CDM compliant. Investigators are encouraged to attend OHDSI Community Calls to promote and recruit collaborators for their OHDSI network study. At the end of the analysis, aggregate study results are made available in the OHDSI GitHub. Where possible, investigators are encouraged to publish study R Shiny Applications to <a href="http://data.ohdsi.org">data.ohdsi.org</a> .”
“Many organizations do not require data sharing agreements be executed between the participating study team members. However, depending on the institutions involved and the data sources, it may be necessary to have more formal data sharing agreements in place and signed by specific study team members. If you are a data owner interested in participating in network studies, you are encouraged to consult your local governance team to understand what policies are in place and must be fulfilled to join OHDSI community studies.”

## Conclusion

This selection of key points and language are to provide a foundation for clear communication within participating network sites that are newer to the unique OHDSI approach to distributed observational research. This abstract may provide a helpful consolidation of information for researcher getting started with OHDSI and a base for developing a more standardized process for dealing with the diversity in IRBs and data governance across the real-world data landscape.