# Utilizing sendigR to Increase the Amount of Data and Confidence in a Cross Study Vehicle Effect Analysis by Using Data from Other Sites

**Featured initiative:** SEND Implementation for Cross-Study Analysis

**Featured solution:** sendigR

In August 2022, a cross-organization team consisting of members from BioCelerate, FDA CDER OND and PHUSE published sendigR as a free, open-source tool that can be utilized to interrogate large repositories of SEND formatted data from toxicology studies.

A Member Company utilized sendigR to increase the data pool and increase confidence in an analysis to update vehicle guides. Understanding which, if any, adverse events could be attributed to the vehicle provides better insight into safety of the compound. This can aid with decision making, development plans, potentially decrease time for the compound to move to phase 1 clinical study and minimize adverse events experienced by patients.

In this case, the member company was interested in finding an answer to this type of question: "Are clinical pathology values comparable across naive monkeys in their internal study data and CRO studies." The company already had their internal studies stored and retrievable from their internal data warehouse. However, it was recognized that a larger data pool could increase confidence in the analysis. To increase the data pool, the company wanted to look at data from a Contract Research Organization (CRO) site.

The company identified a list of studies in SEND format and assigned the research question to an undergraduate intern. The intern had Python experience, but R, the language used for sendigR, was new to her. With little training on SEND standard, R and SQLite, she was able to follow the introduction within the sendigR page and GitHUB and begin the analysis. She set up her R Studio and installed the sendigR package from The Comprehensive R Archive Network (CRAN).

The CRO site studies were SEND datasets in .xpt files which were delivered to the member company. Therefore, once sendigR was installed, the intern built a data warehouse with all the CRO studies. Next, she ran queries of the warehouse from an R script using the sendigR package that provided a generic query function. She could now access the CRO data together with the member company data for analysis and plotting.

By using sendigR the intern was able to build a data warehouse for the CRO studies and then search across the datasets in an efficient way to evaluate the comparability of clinical pathology values. The clear and useful instructions available for the SEND standard, R and SQLite allowed the intern to undertake these research questions with minimal training effort.

The ability to efficiently explore these non-routine scientific cross-study questions using a larger data pool leads to quicker development decisions with greater confidence.