

PHUSE US CONNECT 2021

Cross-study Analysis of SEND Datasets Using an R Package: sendigR*

Presentation AS11

Kevin Snyder, Ph.D.
FDA Center for Drug
Evaluation and Research,
Office Of New Drugs
Silver Spring, MD



William Houser
Principle Scientist
Bristol Myers Squibb
New Brunswick, New Jersey

*Collaboration Work: Kevin Snyder¹, Yousuf Ali², Jesse Anderson¹, William Houser³, Daniel Russo⁴, Bo Larsen⁵, Mark Carfagna⁶

¹United States Food and Drug Administration, Center for Drug Evaluation and Research, White Oak, MD, US; ²Oak Ridge Institute for Science and Education, Oak Ridge, TN, US; ³Bristol Myers Squibb, New Brunswick, NJ, US; ⁴Rutgers University, New Jersey, US; ⁵Novo Nordisk, Copenhagen, Denmark; ⁶Eli Lilly, Indianapolis, IN, US

Disclaimer

This presentation reflects the views of the author and should not be construed to represent FDA's views or policies.

Standard for Exchange of Nonclinical Data (SEND)



- SEND is a data standard created by the Clinical Data Interchange Standards Consortium (CDISC) to enable collection and presentation of nonclinical data in a consistent format.
- In December 2014, the FDA finalized a binding guidance requiring that certain clinical and nonclinical studies be submitted with standardized electronic data, formatted in accordance with the standards recognized in FDA Data Standards Catalog, which currently includes SEND versions 3.0 and 3.1.
 - As of May 2021, > 5,000 SEND datasets have been submitted to CDER.
- The CDISC-SEND data standard has created new opportunities for collaborative development of open-source software solutions to facilitate cross-study analyses of toxicology study data.

FDA/BioCelerate/PHUSE Collaboration



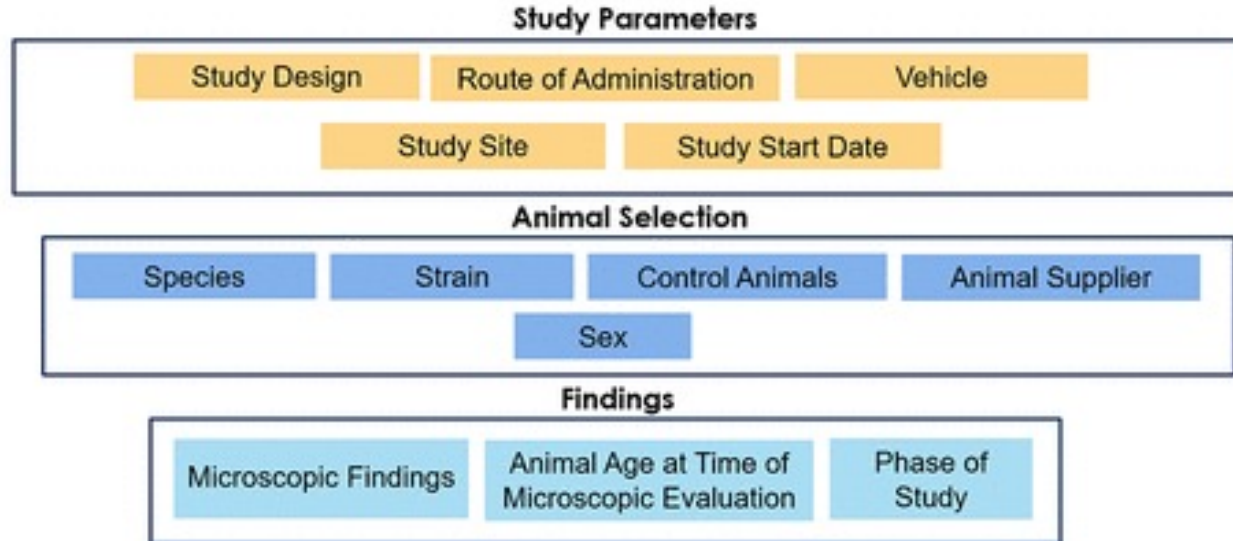
- FDA/CDER has established collaborative partnerships with PHUSE and BioCelerate to develop and publicize novel methods of extracting value from SEND datasets.
- **First publication** focused on identifying challenges and opportunities to improve harmonization of the standard via industry adoption of best practices in SEND dataset creation and/or modification of the SEND data standard in subsequent releases:
 - **SEND harmonization & cross-study analysis: A proposal to better harvest the value from SEND data.** Carfagna MA, Bjerregaard TG, Fukushima T, Houser W, Sloan C, Snyder K, Anderson J, Page T (2020). *Regulatory Toxicology and Pharmacology*, 111, 104542. <https://doi.org/10.1016/j.yrtph.2019.104542>
- **Second publication** explored the feasibility of a practical cross-study analysis use case, i.e., building and querying a historical control database from SEND datasets:
 - **Leveraging the value of CDISC SEND datasets for cross-study analysis: Incidence of microscopic findings in control animals.** Carfagna MA, Anderson J, Eley C, Fukushima T, Horvath J, Houser W, Larsen B, Page T, Russo D, Sloan C, Snyder K, Thompson R, Ullmann G, Whittaker M (2020). *Chemical Research in Toxicology*, 34(2):483-494. <https://doi.org/10.1021/acs.chemrestox.0c00317>

Historical Control Query Design

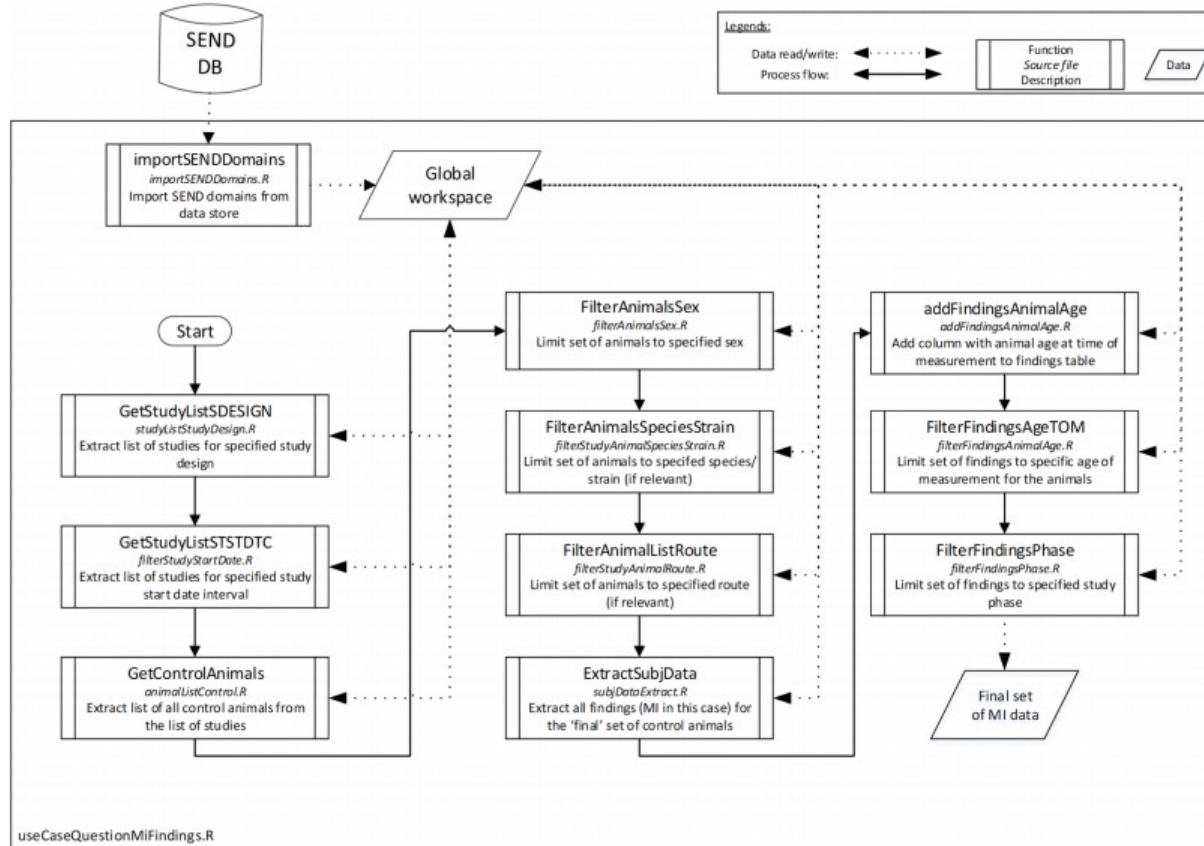


Control Data to Guide Cross Study Analysis Algorithm Development

What are the most common microscopic findings in male beagle dogs 12-18 months of age that were dosed by oral gavage?



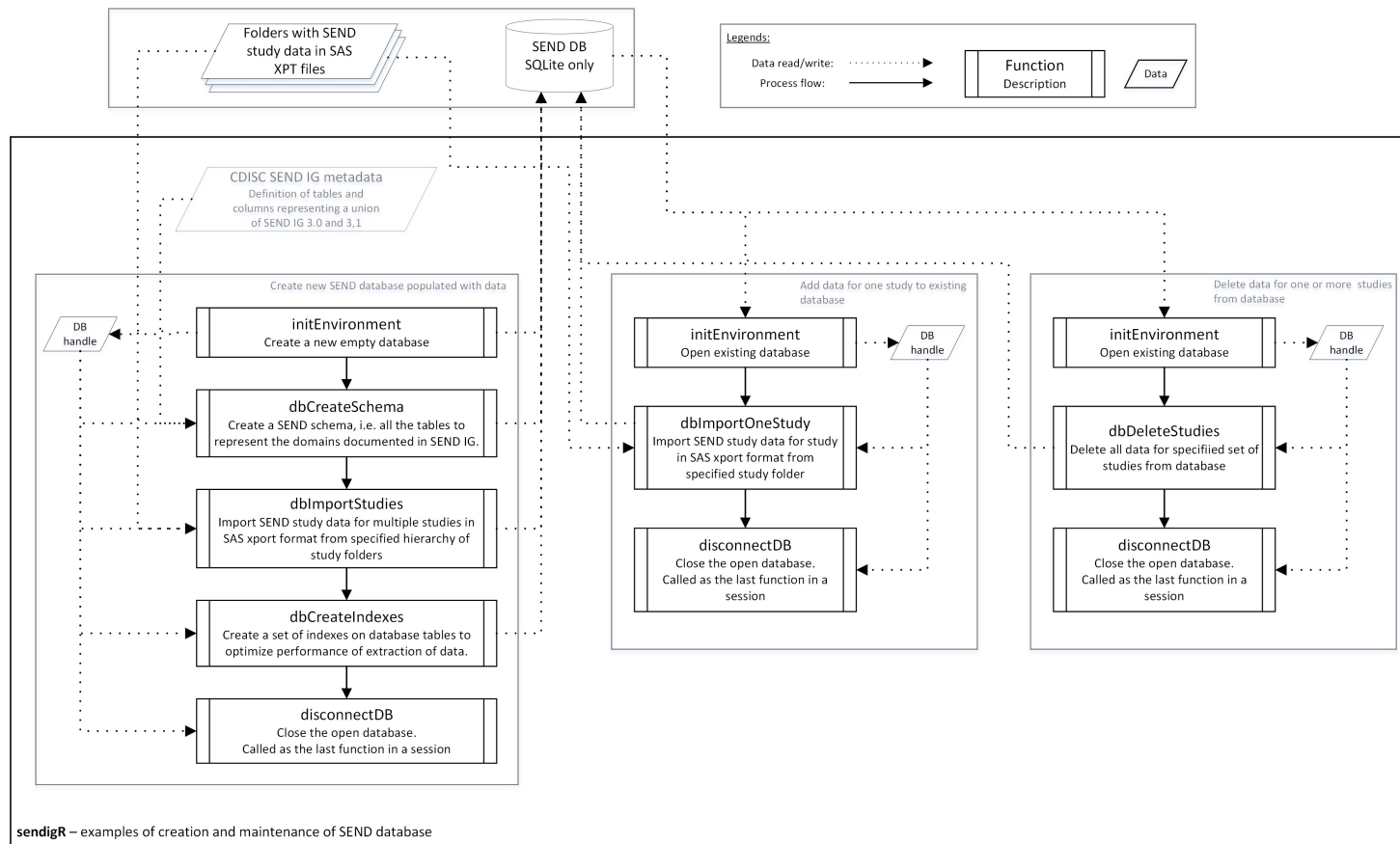
Technical Query Design



R Package: sendigR

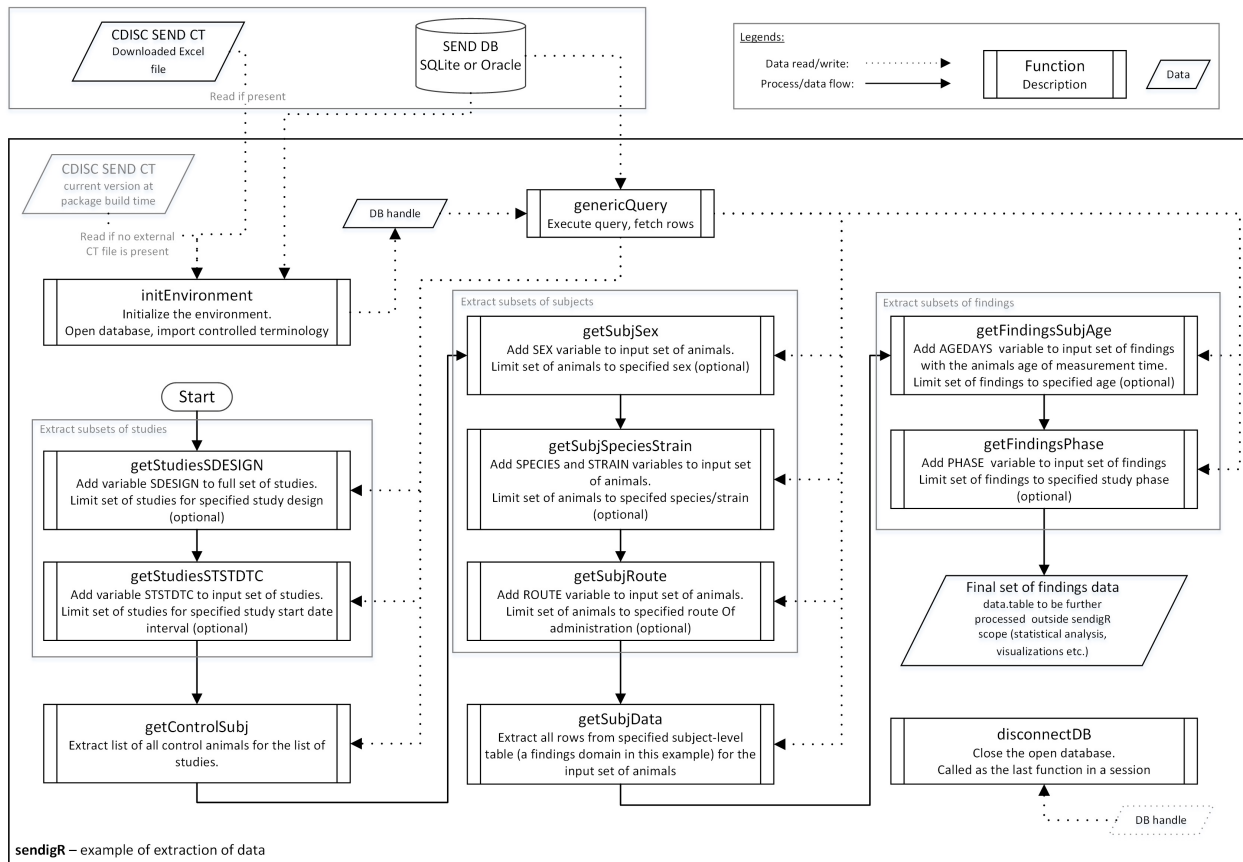
- R scripts from Carfagna et al., 2020 were subsequently enhanced and refactored as an R package called sendigR.
- Functionality of the package is primarily focused on facilitating the targeted extraction of historical control data based on user-specified study and/or animal parameters, e.g. date of study, route of administration of test article, species, animal age, etc.
- Additional functionality will be added to allow users to compare and contrast toxicological profiles of various test articles across studies.
- Source code for sendigR is currently available as a PHUSE GitHub repository:
 - <https://github.com/phuse-org/BioCelerate>

How to use sendigR to Build/Maintain a Database



Functions in the sendigR package can be used to generate a new SQLite database and add/remove studies from an existing SQLite or Oracle database.

How to use sendigR to Query a Database



Functions in the sendigR package can be used to query a database of SEND data to extract findings from control animals that match specific criteria.

R Shiny Web Application

- Users who are not familiar with the R programming language are able to utilize the R Shiny web application to perform cross-study analysis.
- Shiny is an R package developed by RStudio to enable R programmers to easily build interactive web applications (<https://shiny.rstudio.com>).
- The R Shiny web application is currently hosted on a public R Shiny server provided by Rstudio:
 - <https://phuse-nonclinical-scripts.shinyapps.io/sendigR>



Software Demonstration



Conclusion

- The **sendigR** package is a free open-source software solution that will enable data scientists and toxicologists to interrogate large repositories of SEND-formatted toxicology study data.
- Current functionality facilitates historical control data analysis, but functionality will be expanded to cover additional cross-study analysis use cases, e.g., assessment of toxicological profiles, off-target toxicity, etc.
- This is an open, collaborative project. Feel free to try out **sendigR** and contact Kevin.Snyder@fda.hhs.gov if you are interested in providing constructive feedback or getting involved as a contributor.

Acknowledgments

- FDA
 - Kevin Snyder (CDER/OND)
 - Jesse Anderson (CDER/OCS)
 - Yousuf Ali (ORISE)
 - Vaishnavi Methuku (ORISE)
- BioCelerate
 - Mark Carfagna (Eli Lilly)
 - Brianna Paisley (Eli Lilly)
 - Bo Larsen (Novo Nordisk)
 - Bill Houser (Bristol Myers Squibb)
- PHUSE
 - Daniel Russo (Rutgers University)
 - Michael Rosentreter (Bayer)

Questions