PP13 SENDing Toxicology Study Data Analysis into the 21st Century with a New R Package: sendigR

Cm Sabbir Ahmed[1,2], Md Yousuf Ali[1,2], S. Butler[1,2], M. DeNieu[4], W. Houser[3], B. Paisley[5], M. Rosentreter[6], K. Snyder[1], W. Wang (Wenxian.wang@bms.com)[3]
US Food & Drug Administration, Silver Spring, MD[1]*; [2]Oak Ridge Institute for Science and Education, Oak Ridge, TN; [3]Bristol Myers Squibb, New Brunswick, NJ; [4]Labcorp Drug Development, Madison, WI; [5]Eli Lilly & Co., Indianapolis, IN; [6]Bayer AG, Wuppertal, Germany,

Introduction

The CDISC-SEND data standard created new opportunities for collaborative development of open-source software solutions to facilitate cross-study analyses of toxicology study data. A public private partnership between BioCelerate and FDA/CDER works to develop and publicize novel methods of extracting value from SEND datasets. In collaboration with PHUSE, an R package, *sendigR*, has been developed to enable harmonization and visualization of SEND datasets for cross-study analysis.

Currently, the 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. Additional functionality will be added to allow users to compare and contrast toxicological profiles of various test articles across studies.

sendigR Components

sendigR includes:

1) a Python package to harmonize SEND datasets using SEND CT and extensible terminology,
2) an R script to construct a relational database from any collection of SEND datasets,
3) an R Shiny application with a graphical user interface, allowing users who are not familiar with programming to perform cross-study analyses.

sendigR enables experienced R programmers to integrate the package functions into their own custom scripts/packages enabling continued improvements to the functionality of these tools.

SEND Transport Files Cleaner

Python Package xptcleaner is integrated with the sendigR package to standardize SEND Transport files using CDISC controlled terminologies. It can be used to harmonize datasets before importing to sendigR database. Experienced Python programmers, on the other hand, will be able to contribute to the xptcleaner project to allow more variables to be harmonized.

Controlled Terminology lists currently supported:

- Sex
- Strain/Substrain
- Species
- SEND Severity
- Route of Admin. Response
- Standardized Disposition Term
- Specimen
- Anatomical Location
- Non-Neoplastic Finding Type
- SEND Control Type

Methodology

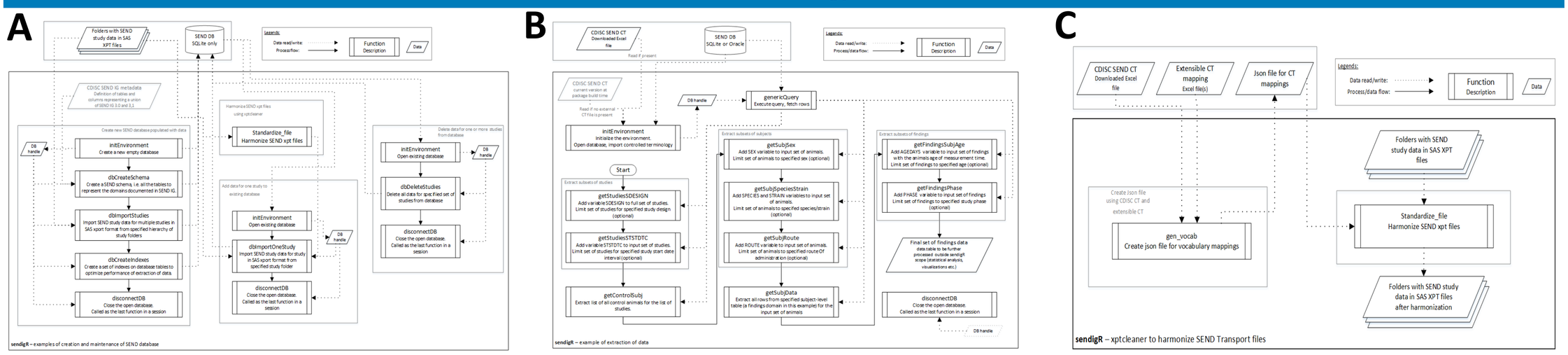
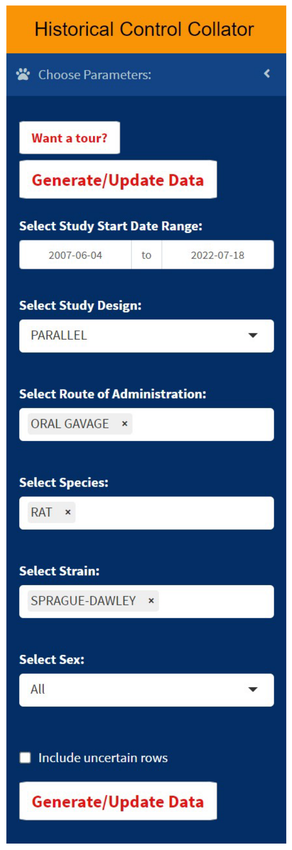
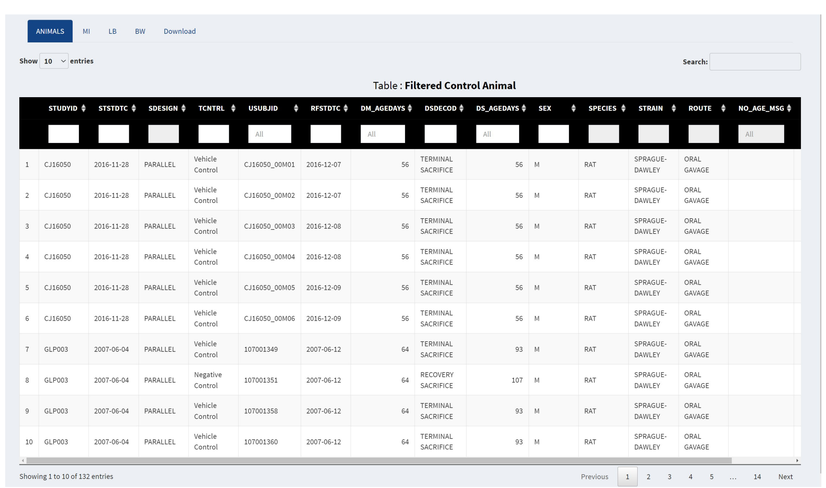
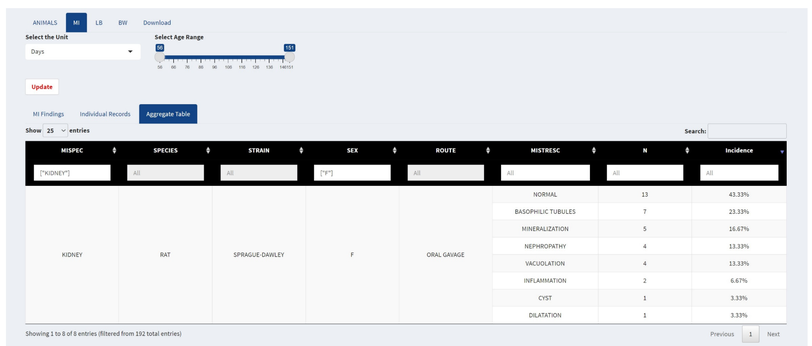


Figure 1. Schematic diagrams depicting how sendigR can be used to build, maintain, and query a SQLite database of SEND data. (A) Functions in the sendigR package can be used to generate a new SQLite database and add/remove studies from an existing SQLite or Oracle database. (B) Functions in the sendigR package can be used to query a database of SEND data to extract findings from control animals that match specific criteria. (C) Functions in the sendigR package can be used to harmonize SEND Transport files using CDISC and extensible CTs.

Results







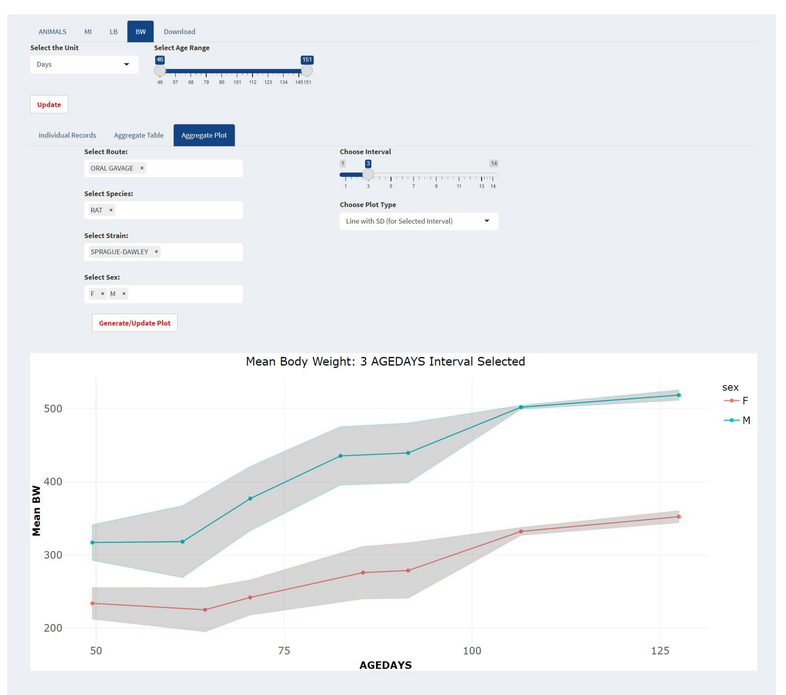


Figure 2. Screenshots from the sendigR R Shiny web application. (D) Users can customize the selection of control animals by study design/date, route of administration, species, strain, and sex. (E) Line listings are generated describing the characteristics of each control animal selected. (F) Historical control distributions of toxicology study endpoints, i.e. histopathology (MI), clinical pathology (LB), and body weights (BW), are aggregated and displayed in an interactive table. (G) Age days vs Mean BW plot can be generated for Body Weight (BW) domain.

Conclusion

The sendigR package will provide data scientists and toxicologists with a free, open-source toolbox that can be utilized to interrogate large repositories of SEND-formatted toxicology study data.

Relevant Publications/Links

- A demo version of the R Shiny application is available and publicly hosted on https://phuse-org.shinyapps.io/sendigR.
- The stable version of this sendigR package is on the Comprehensive R Archive Network (CRAN - Package sendigR (r-project.org)) while the development version is on https://github.com/phuse-org/sendigR.
- 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). Leveraging the value of CDISC SEND datasets for cross-study analysis: Incidence of microscopic findings in control animals. *Chemical Research in Toxicology*, 34(2):483-494. https://doi.org/10.1021/acs.chemrestox.0c00317

*This Poster reflects the views of the authors and should not be construed to represent FDA's views or policies.