FDA U.S. FOOD & DRUG ADMINISTRATION

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

Kevin Snyder, Ph.D.
Associate Director of Nonclinical Informatics
US Food and Drug Administration
Center for Drug Evaluation and Research,
Office Of New Drugs
Silver Spring, MD

Disclaimer

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

SEND: New Opportunities for Cross-Study Analysis of Toxicology Study Data

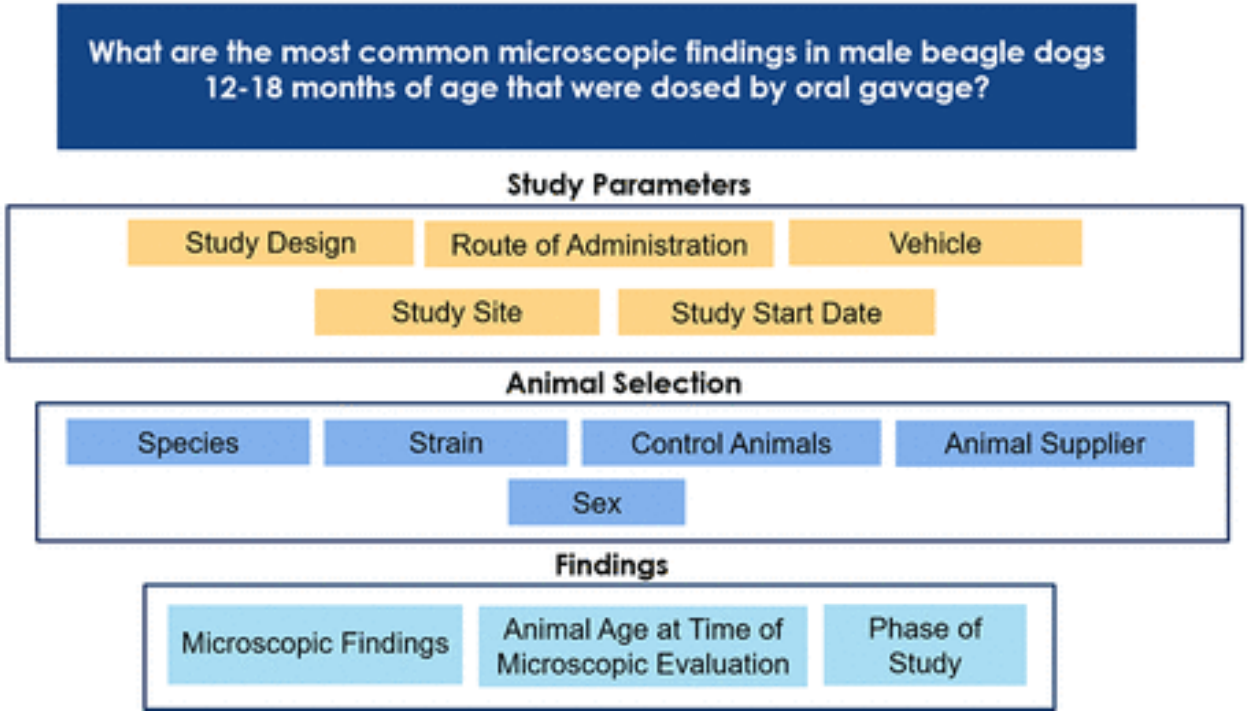
- In December 2014, the FDA finalized a binding guidance requiring that certain clinical and nonclinical studies be submitted with standardized electronic data, formatted in accordance with the standards recognized in FDA Data Standards Catalog.
  - Study types that are currently modeled in SEND include:
    - Single-Dose General Toxicology
    - Repeat-Dose General Toxicology
    - Carcinogenicity
    - Respiratory and Cardiovascular Safety Pharmacology
  - As of October 2021, > 6,000 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

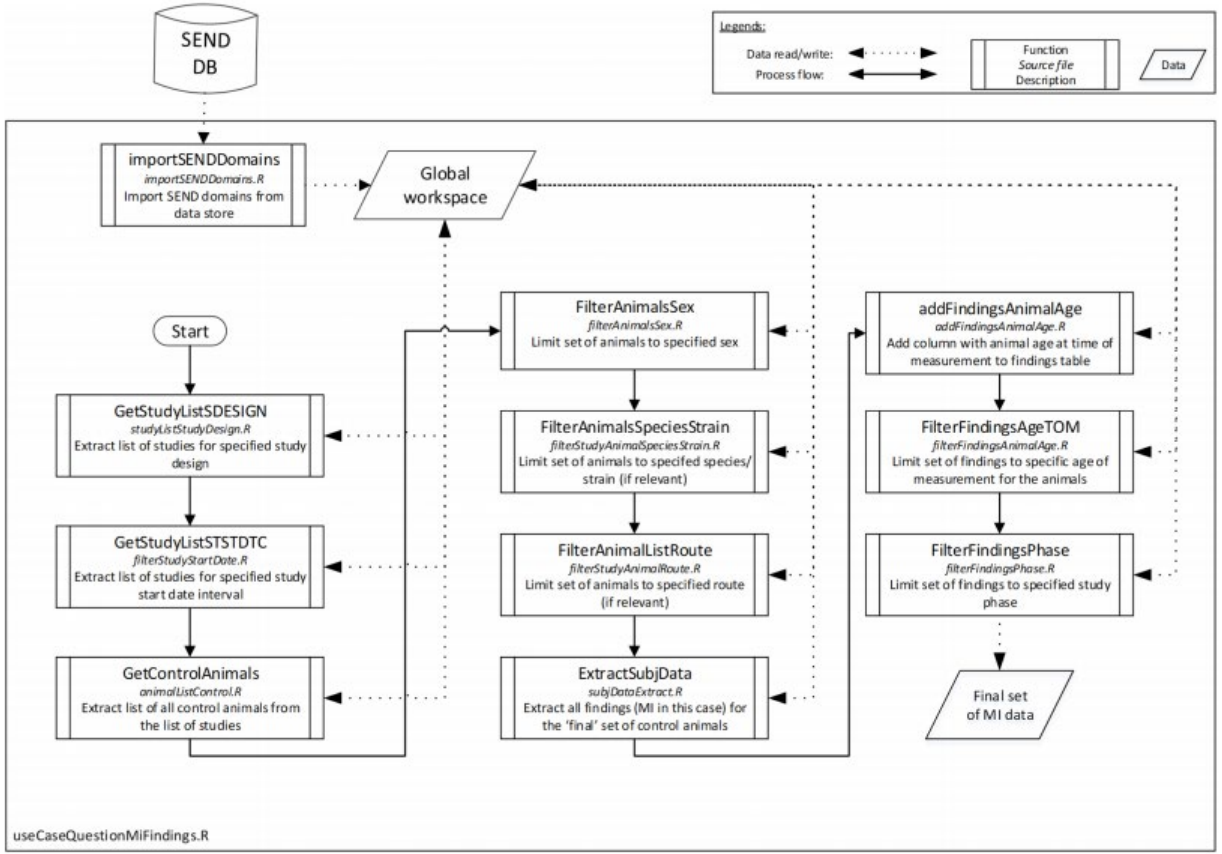
Historical Control Query Design

Control Data to Guide Cross Study Analysis Algorithm Development



(Carfagna et al., 2020)

Technical Query Design



(Carfagna et al., 2020)

Query Result

Organ (MISPEC)	Microscopic Finding (MISTRESC)	Incidence[a]
TESTIS	MATURE	45%
GLAND, PITUITARY	CYST	33%
KIDNEY	MINERALIZATION	23%
THYMUS	ATROPHY	12%
GLAND, PARATHYROID	CYST	9%
TESTIS	HYPOPLASIA	9%
TESTIS	HYPOSPERMATOGENESIS	9%
LUNG	INFILTRATE	8%
THYMUS	CYST	8%
EPIDIDYMIS	CELL DEBRIS	6%

[a]Incidence is reported as the frequency of the finding per total animals (i.e., 102 male dogs).

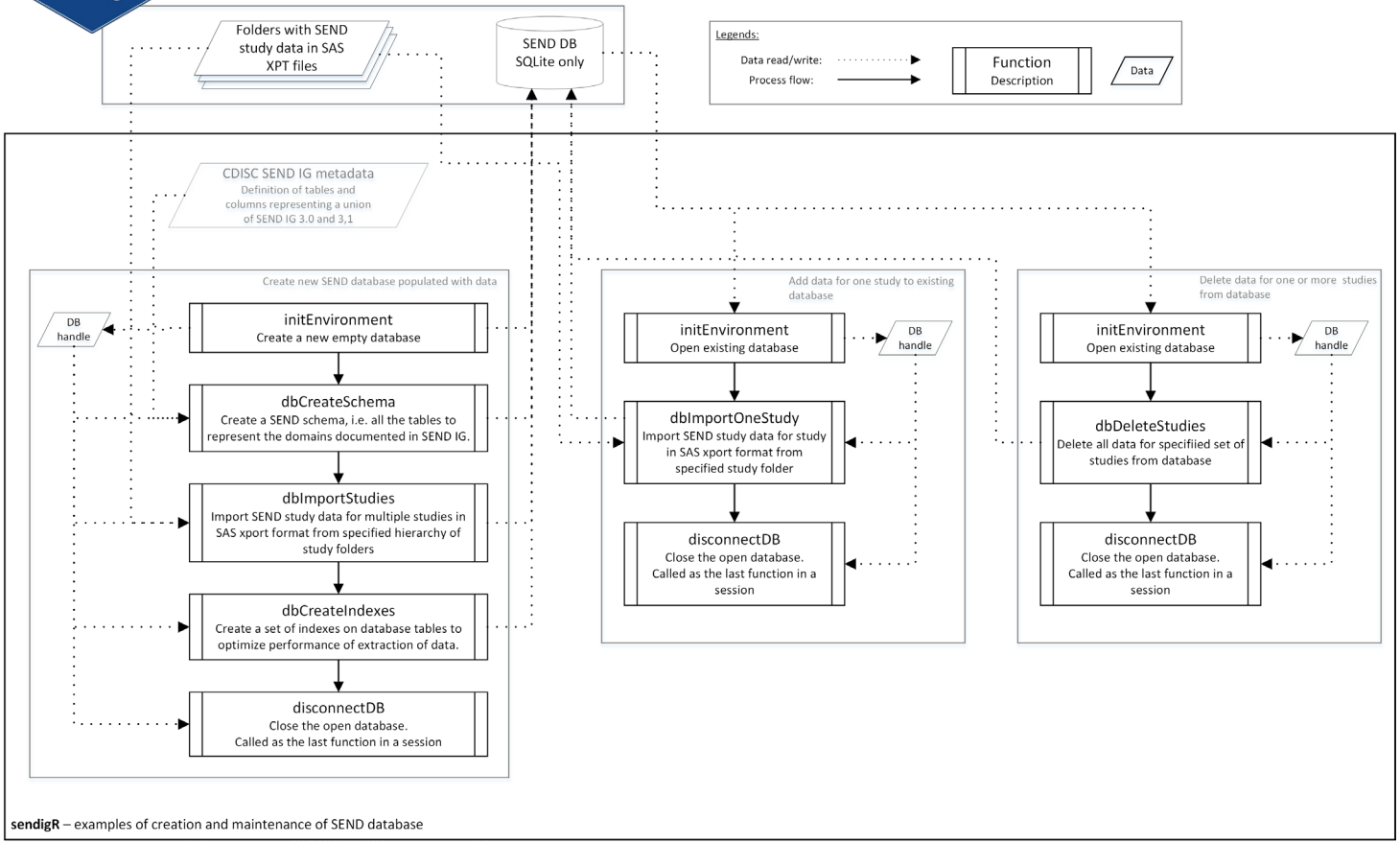
(Carfagna et al., 2020)

R Package: sendigR

FDA

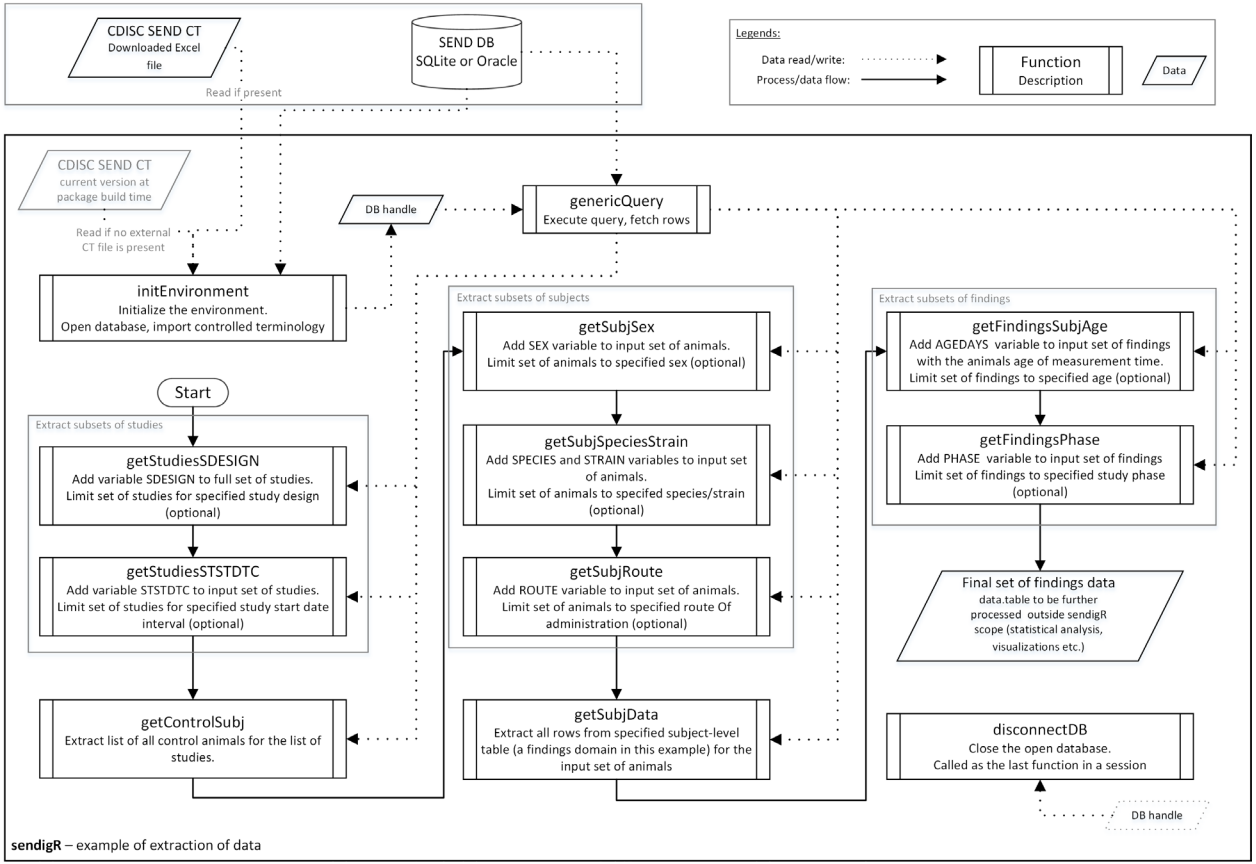
- R scripts from Carfagna et al., 2020 were subsequently enhanced and refactored as an R package called sendigR.
- Functionality of the package is primarily focused on facilitating the targeted extraction of historical control data based on user-specified study and/or animal parameters, e.g. date of 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/sendigR

Usage of 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.

Usage of sendigR to Query a Database

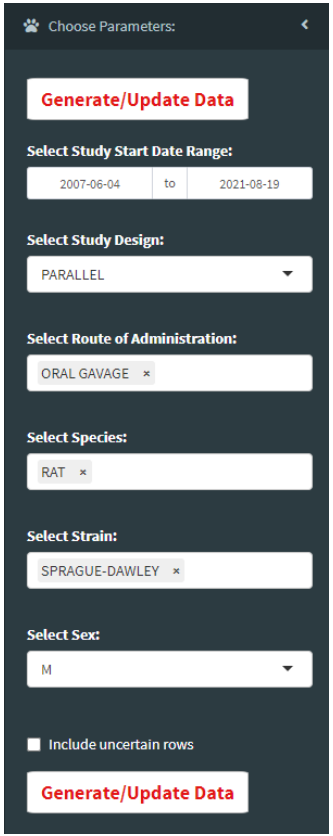


Functions in the sendigR package can be used to query a database of SEND data to extract findings from control animals that match specific criteria.

R Shiny Web Application

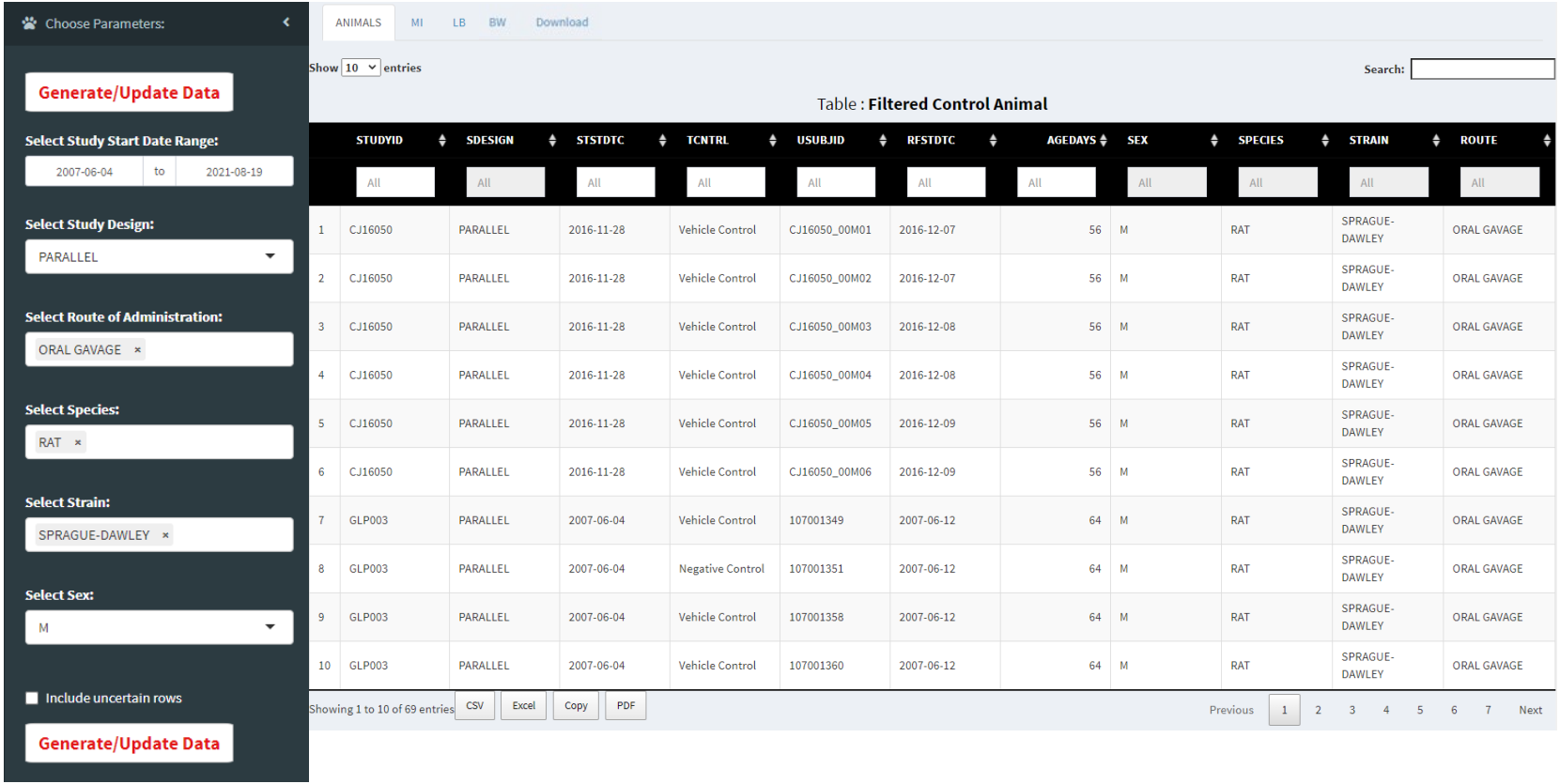
- Users who are not familiar with the R programming language are able to utilize the R Shiny web application to perform cross-study analysis.
- Shiny is an R package developed by RStudio to enable R programmers to easily build interactive web applications (https://shiny.rstudio.com).
- The R Shiny web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

R Shiny Web Application

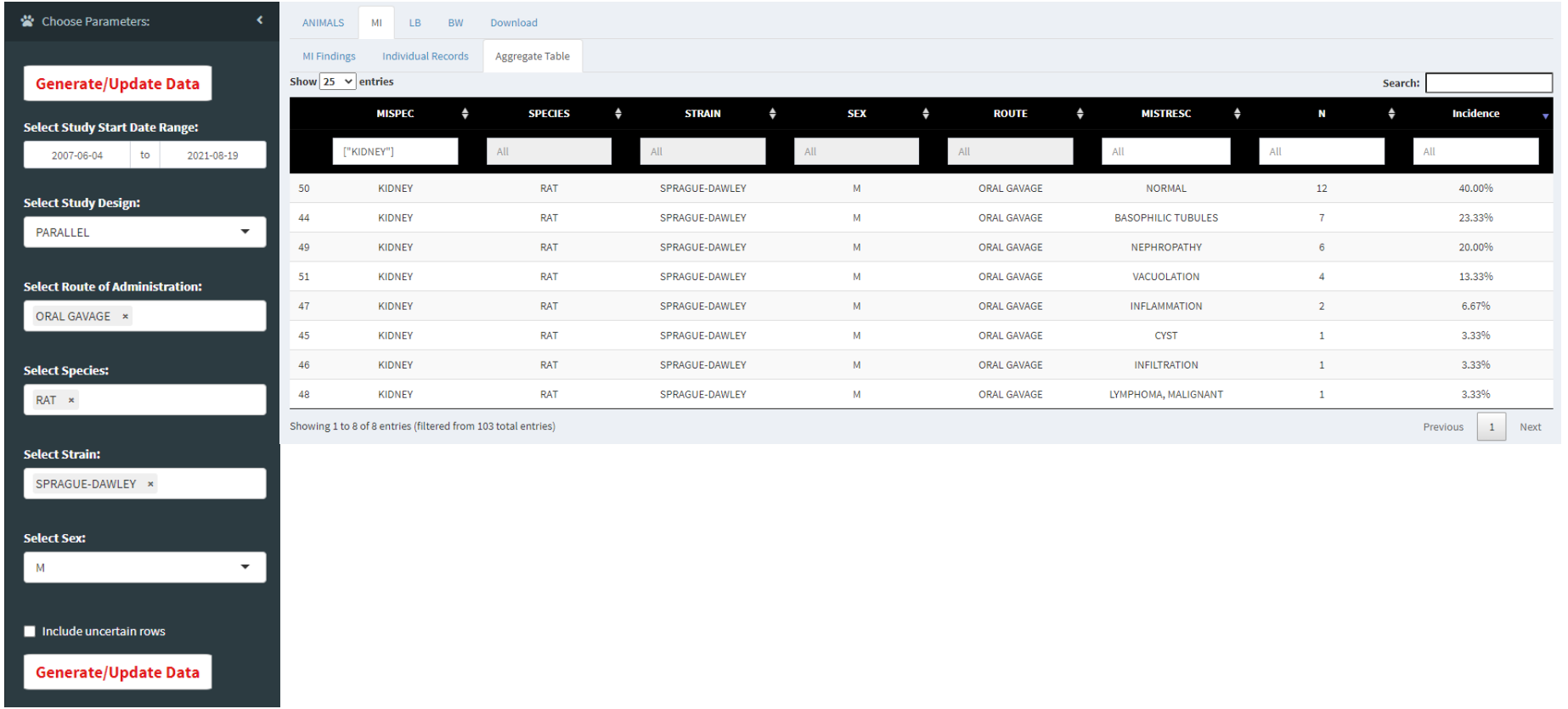


- Users can customize the selection of control animals by study design/date, route of administration, species, strain, and sex.
- Line listings are generated describing the characteristics of each control animal selected.
- Historical control distributions of toxicology study endpoints, i.e. histopathology (MI), clinical pathology (LB), and body weights (BW), are aggregated and displayed in an interactive table.

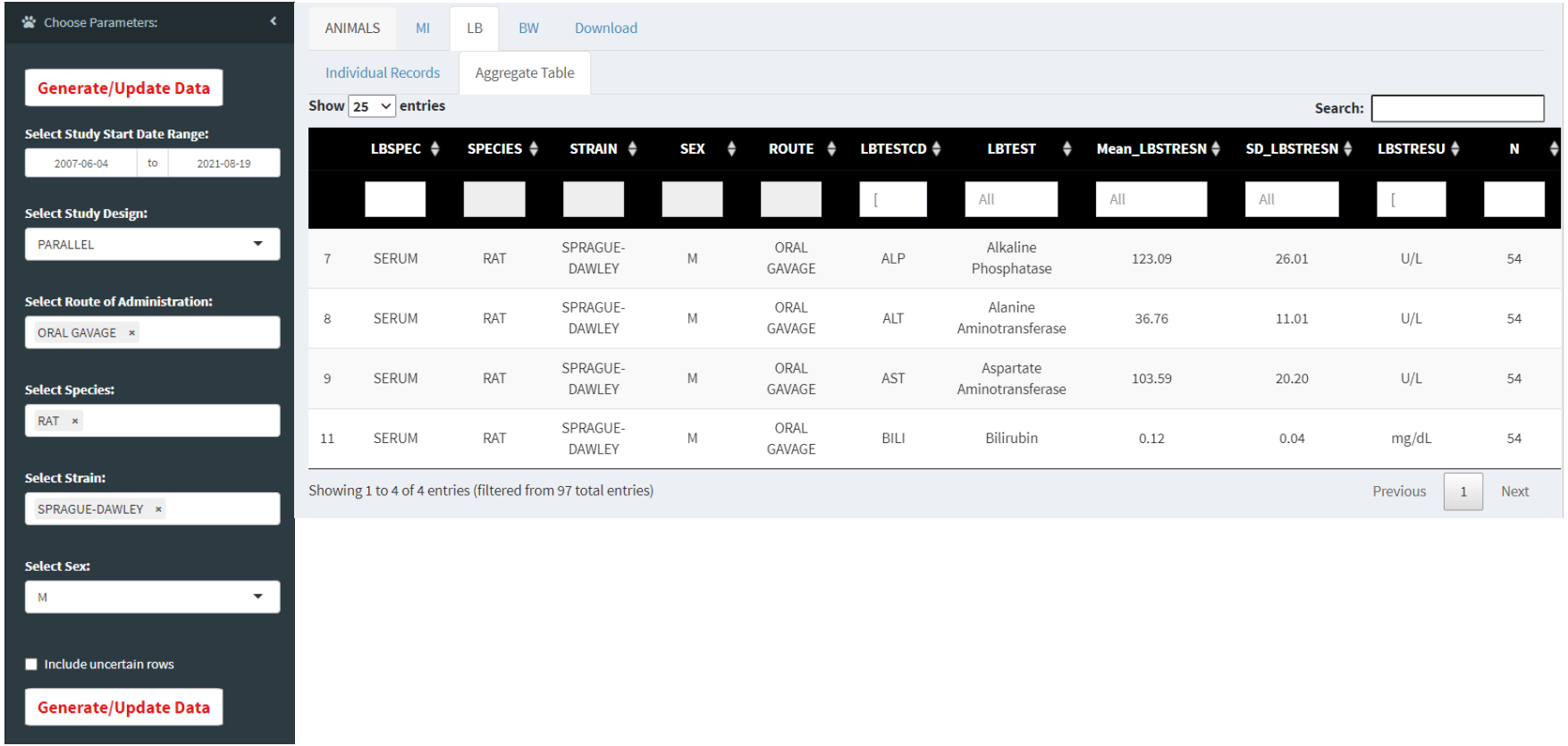
R Shiny Web Application



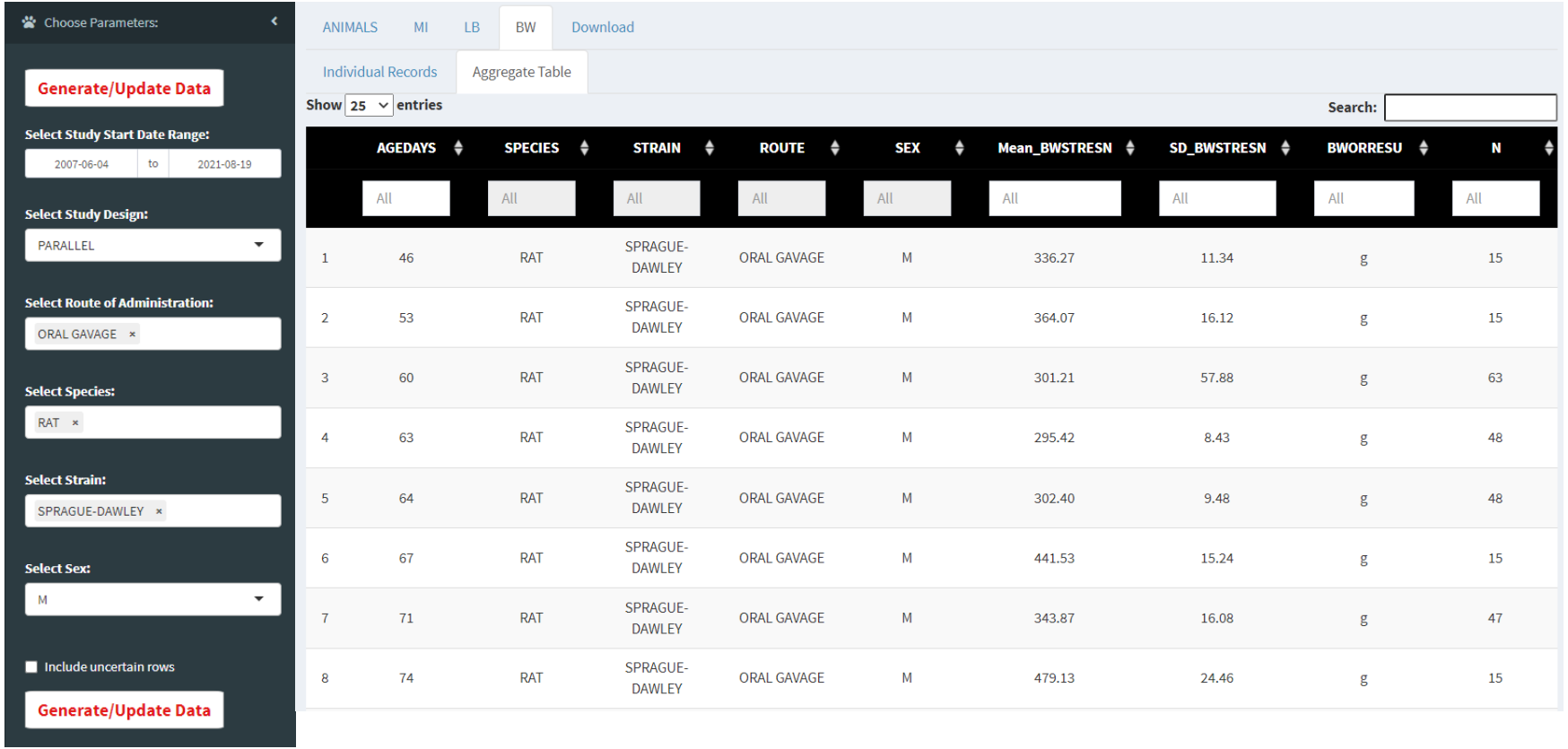
R Shiny Web Application



R Shiny Web Application



R Shiny Web Application



	AGEDAYS	SPECIES	STRAIN	ROUTE	SEX	Mean_BWSTRESN	SD_BWSTRESN	BWORRESU	N
	All	All	All	All	All	All	All	All	All
1	46	RAT	SPRAGUE-DAWLEY	ORAL GAVAGE	M	336.27	11.34	g	15
2	53	RAT	SPRAGUE-DAWLEY	ORAL GAVAGE	M	364.07	16.12	g	15
3	60	RAT	SPRAGUE-DAWLEY	ORAL GAVAGE	M	301.21	57.88	g	63
4	63	RAT	SPRAGUE-DAWLEY	ORAL GAVAGE	M	295.42	8.43	g	48
5	64	RAT	SPRAGUE-DAWLEY	ORAL GAVAGE	M	302.40	9.48	g	48
6	67	RAT	SPRAGUE-DAWLEY	ORAL GAVAGE	M	441.53	15.24	g	15
7	71	RAT	SPRAGUE-DAWLEY	ORAL GAVAGE	M	343.87	16.08	g	47
8	74	RAT	SPRAGUE-DAWLEY	ORAL GAVAGE	M	479.13	24.46	g	15

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)
- PHUSE
  - Daniel Russo (Rutgers University)
  - Michael Rosentreter (Bayer)
- BioCelerate
  - Mark Carfagna (Eli Lilly)
  - Brianna Paisley (Eli Lilly)
  - Bo Larsen (Novo Nordisk)
  - Bill Houser (Bristol Myers Squibb)
  - Wenxian Wang (Bristol Myers Squibb)