ORISE Renewal Report

Appointee: Christopher Cook

Mentor: John Wambaugh

I am involved with several projects centered around EPA’s *Concentration versus Time Database* (CvTdb). The CvTdb is a database of standardized time-course data of chemical compound concentrations measured *in vivo* from hundreds of publications. It serves as a platform for calibrating and validating pharmacokinetic models on a large and diverse set of data, and more broadly, can improve the EPA’s toxicological risk assessments.

Much of my work over the past year has involved developing tools and workflows to aid in the addition of data to the CvTdb. Data are submitted to the CvTdb as Excel template files, which are manually curated by several different people. Around a hundred files needed to be checked for format quality when I first started my position. One of my first tasks involved developing an R Shiny application to help with template quality control. The R Shiny application checks template format, and perhaps more importantly, reconstructs plots reported in the curated publication. I used this tool to go through submitted templates and ‘pass’ or ‘fail’ them. I created the tool as a standalone desktop application, so it would be easy to pass off to curators, so they could check their own templates before submission.

I also assisted with the integration of new data shared with us by Showa Pharmaceutical University, from Tokyo, Japan. I developed an R script that changed the files shared with us by Showa to fit our own format. I also manually curated some of the missing data. This data has since been added to the CvTdb, with much help from folks in the Scientific Computing and Data Curation Division. With the addition of data from Showa Pharmaceutical University, the CvTdb doubled in size.

I lead the programming for an R software package, entitled *invivoPKfit*, that fits 1- and 2-compartment pharmacokinetic models to the data in the CvTdb. These models are used to obtain estimates of key parameters for assessing chemical risk, such as volume of distribution and half-life, and will be reported on the CompTox Chemicals Dashboard (<https://comptox.epa.gov/dashboard/>). I have adapted *invivoPKfit* to fit models to the latest iteration of the CvTdb, which includes the newly added data from Showa Pharmaceutical University. Much of my work involved troubleshooting poor fits, or lack of fits, and updating *invivoPKfit* accordingly.

I presented all aspects of my work as a poster at the North Carolina Society of Toxicology on January 18, 2022. I am presenting an updated version of this poster at the national Society of Toxicology on March 30, 2022. I am working on publishing my work with *invivoPKfit* in summer 2022. I am also working to submit *invivoPKfit* to the Comprehensive R Archive Network (CRAN).