

Novo Nordisk: Study R Plots

Novo Nordisk uses R plots to visualize study dataset data to track trends across multiple combined studies, that all belong to the same dataspace, and to track the performance of their diabetic animal model. Having the study datasets defined as part of a dataspace allows for visualizations of multiple studies in a single report. From the combined data grid view, the standard data grid filtering features allow for users to drill down to the specific studies, subjects, and treatments of interest. Since the R reports update to respect the given set of filters on the data grid and to display new data as it is entered into the datasets, there is no need to have duplicate copies of the same code for a predefined set of filter combinations or studies. Instead the reports can be used to explore the dataspace for various different slices of the data.

To view the R reports, the users of the system are able to see a set of links for the reports at the top level, i.e. dataspace project, in the Data Views webpart. In some cases, reports may also be displayed as a webpart on the individual custom study pages. All users on the server have read permissions to the full set of studies in the dataspace, but only certain users are given the permissions to make edits to the data.

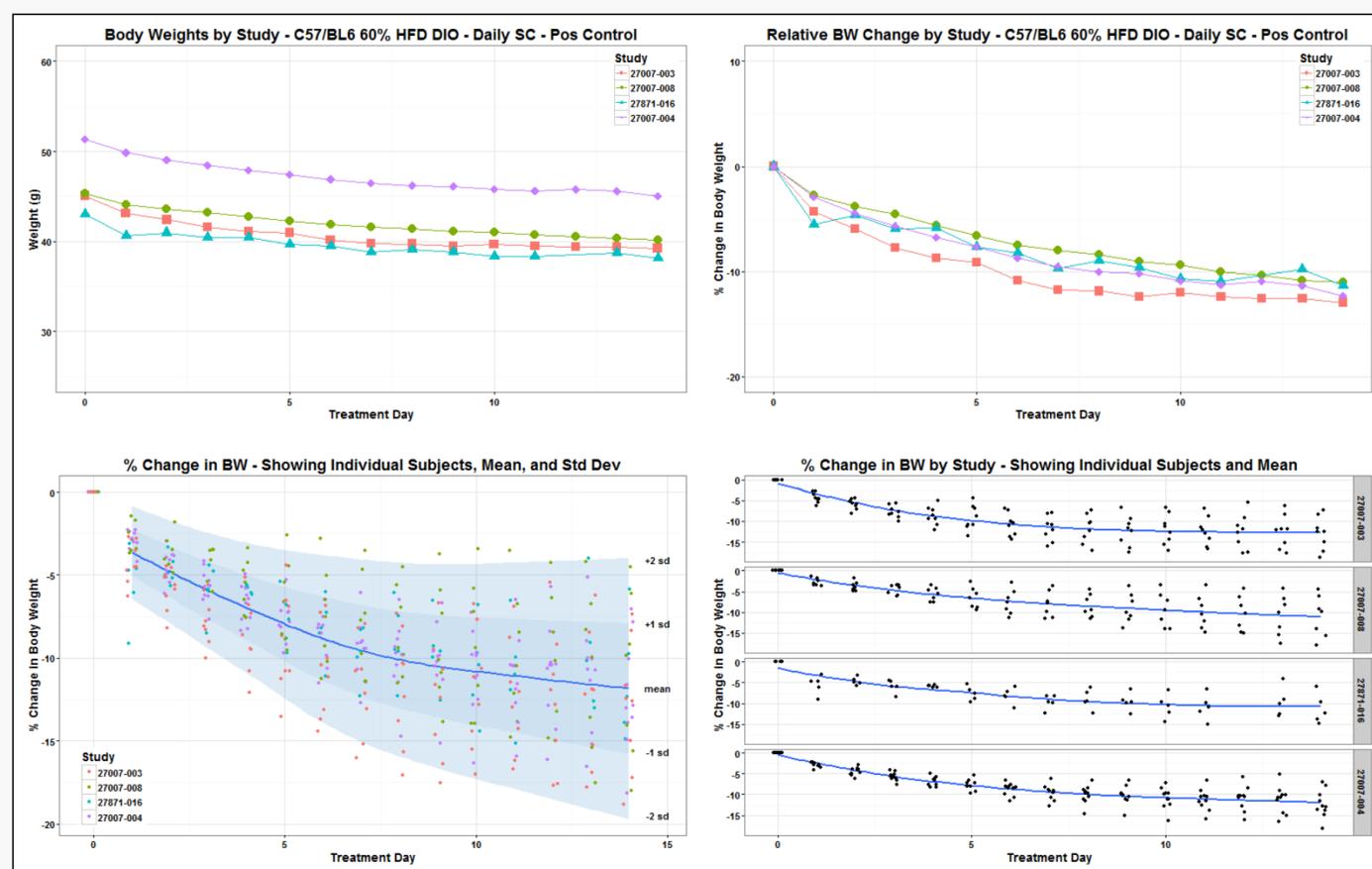
Technologies Used

- **Data source:** common study dataset defined in a dataspace folder, data populated for individual studies
- **Code location:** standard R report attached to a filtered dataset, data collection and transformation via SQL queries
- **Tools/packages:** R packages (ggplot2, plyr, dplyr, gridExtra, survival, Rlabkey)
- **Features:** data grid filtering, real-time plot updates as new data is loaded into studies
- **Accessibility:** from Data Views webpart in parent dataspace project, webpart on custom study page
- **Developer(s):** Jon Rue, Novo Nordisk

Report Examples

Positive Control DIO HFD BW Analysis

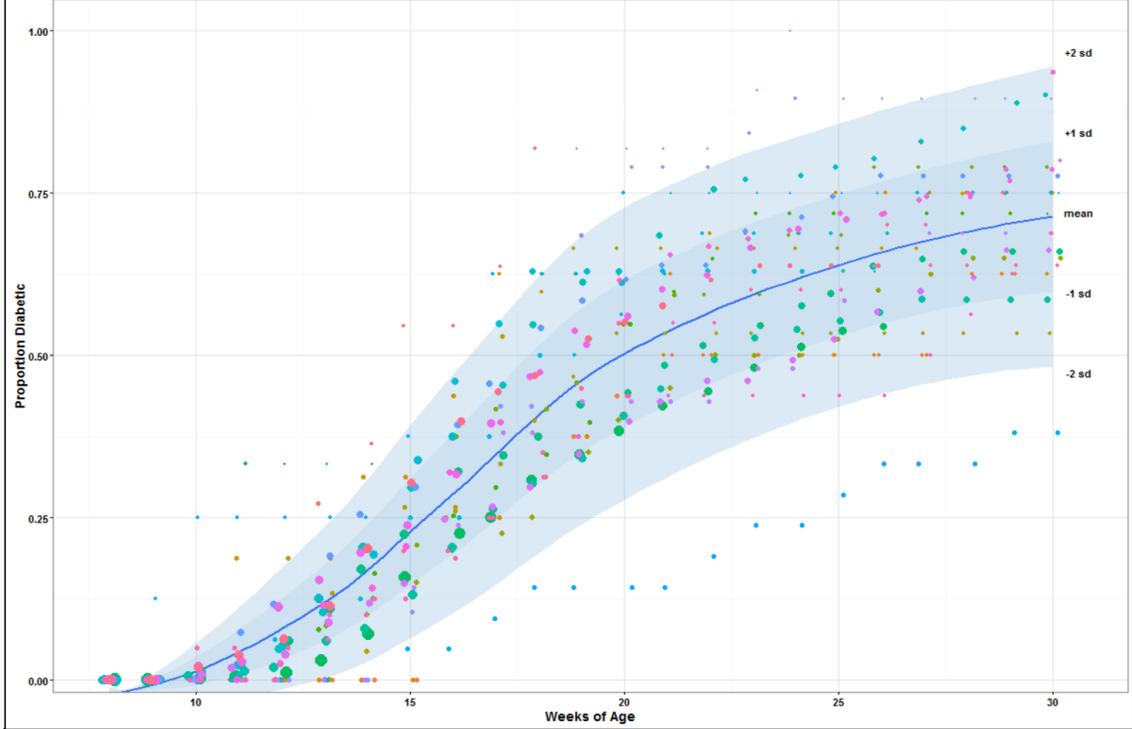
This plot is a series of visualizations that track the performance of a particular positive control compound for a selected set of experimental conditions (C57/BL6 mice on a 60% high fat diet that were enrolled in studies where they received once daily subcutaneous treatments of the compound of interest). The plot shows data across all of the studies in the LabKey system.



Untreated Diabetic Incidence by Study

This plot tracks the historical performance of Novo Nordisk's primary diabetic animal model. This is historically a pretty sensitive animal model so they would like to track stability of the model over time and between studies. This visualization is a Kaplan-Meier survival analysis of diabetes incidence for naïve mice for all studies. It contains the mean trend line for all animals

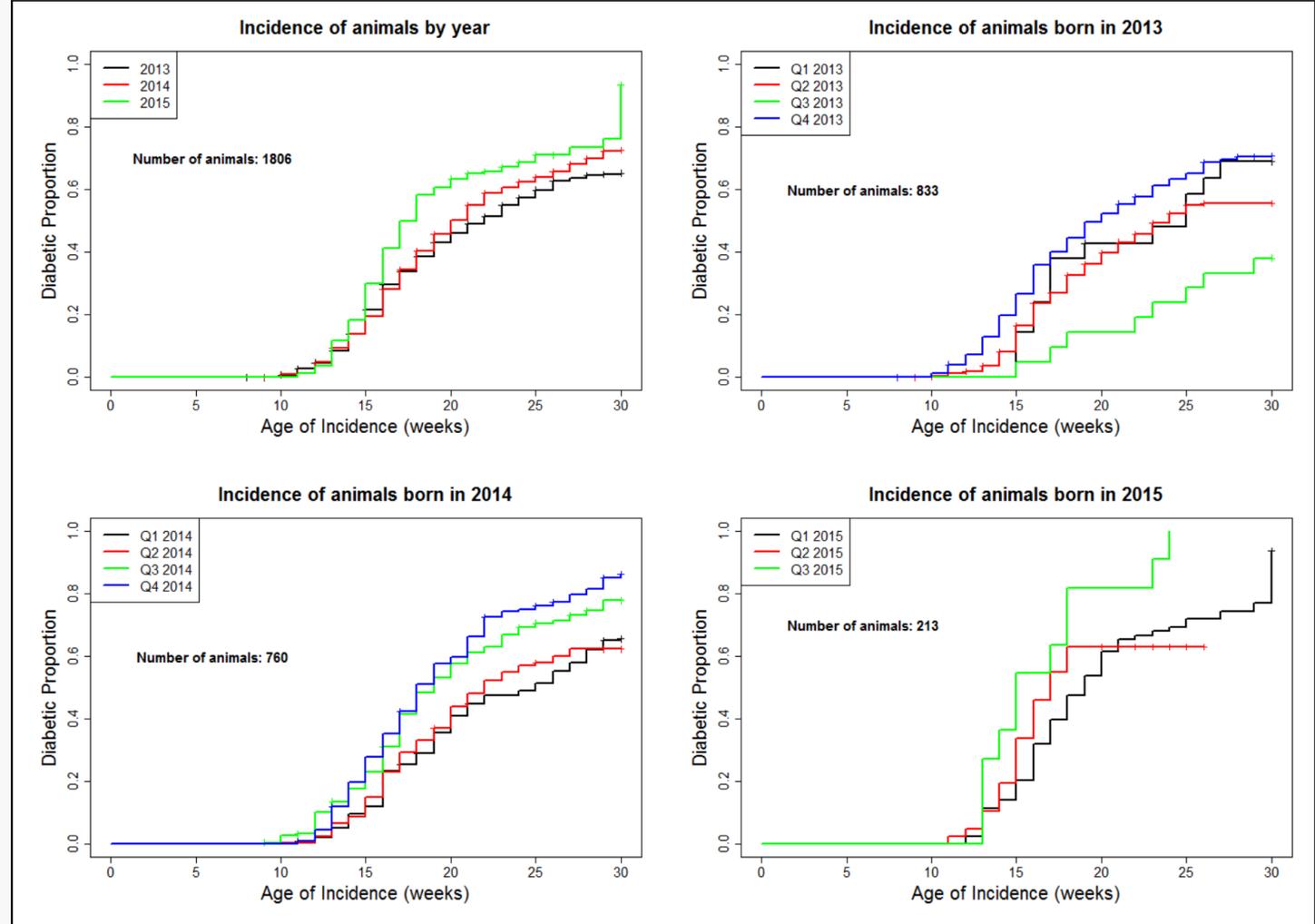
NOD Mouse Model Naive Colony Incidence With Std Dev(to 30 WOA)



observed with shaded regions around the mean to show the expected standard deviations. The colored dots represent cohorts from individual studies and are sized proportionally to the number of animals in the cohort. Like most all of the Novo Nordisk R visualization this is a real time report. As new studies and mice are recorded in the LabKey Server this visualization updates accordingly.

Diabetic Incidence by Year

This plot is built on the same data from the previous example but shows the progression of untreated diabetic incidence over time. The four visualizations show Kaplan-Meier survival plots by year and are broken down by quarter for each year.



Study Incidence with Historical Overlay

This plot is from a study showing the Kaplan-Meier incidence curve for each treatment group in the study with an overlay of the historical untreated colony incidence as a point of reference. This allows Novo Nordisk to see how well a study conforms or deviates from their historical control data.

Kaplan-Meier Incidence for each Treatment Group with Historical Untreated Colony Incidence

