Creating an Analytical Pipeline to Match Biomarker Diagnostics with Clinical Trial Outcomes

**Background**: BXQ-350, a novel protein-lipid small molecule consisting of the ceramide creating protein Saposin C and the lipid Dioleoyl Phosphatidylserine (DOPS), is entering the final stages of Phase 2 clinical trials. With hundreds to tens of thousands of data points per trial, the use of an analytical tool to process the data became a major focus. To meet those ends, Bexion Pharmaceuticals created an in-house analytical platform to perform analysis on these data points. The platform performs data wrangling, processing, calculations, and produces statistical and graphical outputs from multiple different input sources.

**Methods**: The analytical platform was developed using Python v3.13 and open-source modules. Windows Excel files are used for THE input and are formatted in Excel to be properly read. The outputs are statistical Excel files and graphical visualization of the data.

*Analytical Pipeline*

1. Import Excel file and create a pandas dataframe object
2. Merge dataframes based on patient identification if applicable
3. Perform data wrangling on the merged dataframe
	1. Create or update a MySQL table if applicable
4. Process the data and perform calculations
5. Perform QC to determine which changes are within the standard deviation
6. Perform a second round of processing and calculations on data that passed QC
7. Group and separate data for output handling
8. Creation of graphs and output files

**Results:** The analytical platform was created to be agnostic to which type of biomarker is being analyzed allowing for the merging of data and analysis across multiple biomarkers. This allows for correlative analysis to be performed to give a better understanding of both individual and interactive biomarkers. The platform can group patients based on clinical diagnostics in order to tie molecular diagnostics with their clinical counterpoints giving a picture of whether certain biomarkers are able to diagnostically track clinical progression. In addition, statistics are derived for both individual patients and groups to stratify the data.

**Conclusion:** To keep up with the amount of biomarker data being collected from clinical trials an analytical platform was created that is tailored to the needs of analysis. The platform allows for flexible and granular analysis of biomarkers. This includes correlative analysis across different groups of biomarkers to determine if there are possible interactions between them. This includes being able to zoom in on individual patients, groups of patients, or the entire population of the trial. In all, Bexion created a platform to fit the needs to correlate clinical diagnostics with biomarker diagnostics.