Instructions I

THE STORY

You have now received the data from a clinical trial investigating BPI889. It is a double-blinded, randomized study with 1500 subjects receiving 200 mg of the study compound, i.e., BPI889.

PK-file

The concentration data of BPI889 measured before the dose and at 0.5 h, 1 h, 2 h, 4 h, 6 h, 8 h, 12 h, 24 h, and 48h post-dose are available in the file indicated with PK. Units of concentrations are mg/L. Note that some of the planned samples were not taken and are thus indicated as missing. Concentrations and times are considered continuous data.

Cov-file

You have also received a file with the demographics of the study subjects: sex (M/F), age (yrs), height (cm), weight (kg), body mass index (BMI, kg/m^2), and creatinine clearance (CLcr, mL/min) - a measurement of kidney function. All of these are continuous variables, except sex which is categorical.

In addition, in the same file, you have the categorical, phenotypic information of four drug-metabolizing enzymes: CYP2D6, CYP3A4, CYP2D9, and CYP2C19. Possible phenotypes are poor metabolizers, with a reduced function (0), normal metabolizers (1) and extensive metabolizers, with an increased function (2). These have been collected as the team of BPI889 suspects that one of these enzymes is responsible for eliminating the compound from the body.

Appendices

When investigating the pharmacokinetics of a compound the following indices are usually derived: half-life ($t_{1/2}$), maximum concentration (C_{max}), time of C_{max} (t_{max}). volume of distribution (V_d), area under the concentration-time curve (AUC), and clearance (CL). Details about how to retrieve these indices are given in **Appendix A - Pharmacokinetic Parameters and Measurements (https://uppsala.instructure.com/courses/93701/pages/appendix-a-pharmacokinetic-parameters-and-measurements)** . These indices are derived per individual; not per population.

There are three measures of body composition derivable from demographic information: lean body mass (LBM), total body water (TBW), and body fat percentage (BFP). How to derive these and their units are given in **Appendix B - Measurements of body composition. (https://uppsala.instructure.com/courses/93701/pages/appendix-b-measurements-of-body-composition)** Also, these measures are derived per individual; not per population.

Research question

The team is interested in answering these two main questions:

1. *Is distribution of BPI889 dependent on body composition?*
2. *Is elimination of BPI889 dependent on any of the CYP-enzymes?*

The specific tasks below are investigating these questions.

Specific tasks of the examination project:

1. **Import data files** on your computer as objects in R. Comment on the arguments you set and explain your choices.
2. **Assign appropriate names and classes** to variables. Comment on your choice of classes for the variables.
3. **Combine the PK and COV data** into a single data frame in long format (also known as tidy format).
4. **Calculate Body Fat Percentage (*BFP*),** and as a new variable, categorize it into two groups: above and below 24%.
5. **Calculate PK variables**: C_{max}, t_{max}, $t_{1/2}$, and CL.
6. **Numerically summarize** C_{max}, t_{max}, $t_{1/2}$ and CL: mean, median, standard deviation, and range.
7. **Create a spaghetti plot** to display individual concentrations of BPI889 versus time.
8. **Create a correlation plot** between C_{max}, t_{max}, $t_{1/2}$, and CL.
9. **Create box-and-whisker plots** to display $t_{1/2}$ and CL versus enzymes: CYP2D6, CYP3A4, CYP2C9 and CYP2C19.
10. **Create a scatter plot** to display the correlation between $t_{1/2}$ and BFP, and add a linear regression line.
11. **Perform an ANOVA** on C_{max} and CL across the four individual CYP-enzymes, assuming that the phenotypes of the CYPs are categorical.
12. **Perform a linear regression** of $t_{1/2}$ versus BFP. Report on whether it is statistically significant and in what aspect it is statistically significant.

Throughout the project - PAY ATTENTION TO UNITS

Through these tasks, you show your proficiency in programming so that your code can be assessed based on the grading criteria (**found on this page (https://uppsala.instructure.com/courses/93701/pages/examination-project-in-r-info)**) to give the final grade of the R-part of the course.