Team Project

Start Assignment

Due Dec 7 by 5p.m. **Points** 100 **Submitting** a file upload

BTC1877H – Data Science in Health Part II

Fall Term 2022

Team Project

General Information

In this project you will use a dataset describing blood transfusions and outcomes of lung transplant patients. You will use this data in order to answer the following two main general questions:

1. What are the characteristics of patients that require transfusions and what are the factors influencing the need and amount of transfusions? Note here that transfusion of more than 10 RBC units is considered “massive”.
2. In conjunction with the above, what is the impact of transfusion on patient outcomes, including mortality?

The dataset, found **here (https://q.utoronto.ca/courses/267919/files/23154119?wrap=1)** **(https://q.utoronto.ca/courses/267919/files/23154119/download?download_frd=1)** , contains a number of variables that are organized in sections and color coded in an excel datasheet as follows:

Orange - patient demographic data
Green - underlying respiratory diagnosis + important intraoperative descriptions (first lung transplant vs repeat, use of ExVIVO perfusion, whether the patient was on ECLS/ECMO prior to transplant)

Red – preoperative bloodwork (hemoglobin, platelets, coagulation parameters)
Teal - ECMO/No ECMO (whether patient’s transplant was performed with ECMO support or not). The second column highlighted includes Cardiopulmonary Bypass, which is an infrequently used and more complex form of ECLS and it is known to be associated with more bleeding.
Yellow - blood product transfusion data.
Blue - survival and ICU LOS data.

For this project, you will need to make a thorough investigation in the relevant literature and use it to guide your refinement of the research questions mentioned above, as well as to guide your analytical strategy (which predictors to consider, what outcomes etc.)

Format of the report to be submitted

The format of the report will be that of a scientific paper. It needs to include the following sections: Introduction, Methods, Results and Discussion/Conclusions. You can use unlimited number of figures and tables in the main text, but irrelevant and unnecessary figures and tables will be penalized. You have no restrictions in the material you can use as Supplementary. All tables and figures (in both main text and supplementary material) need to have proper formatting (including tables as images is not allowed), numbering and captions. Do not copy and paste R output directly to your report. The grading breakdown is given in the syllabus.

The R code needs to be submitted as a separate R script file, and it should use proper documentation.