Supplementary Table 1. Percentage of Missing Data in Selected Features

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Category** | **Parameter** | **Units** | **Percentage Missing in Training and Test Data, Subject to Imputation** | **Percentage Missing in PRO-ACT Database** |
| Demographic | Sex^ |  | 0.0% | 0.0% |
| Age^ | years | 0.0% | 28.2% |
| Site (bulbar/limb/both/other)^ |  | 0.0% | 12.4% |
| Clinical | Days since Onset of Symptoms | days | 0% | 35.9% |
| Days since diagnosis | days | 39.7% | 63.1% |
| Riluzole Use |  | 22.4% | 17.8% |
| Height | m | 10.9% | 86.6% |
| Weight | kg | 24.8% | 29.7% |
| BMI | kg/m^2 | 28.5% | 23.8% |
| FVC (as % of Predicted for Normal Individual) | proportion | 33.6% | 6.7% |
| Diastolic Blood Pressure | mmHg | 17.7% | 21.0% |
| Systolic Blood Pressure | mmHg | 17.7% | 21.0% |
| Pulse | bpm | 17.5% | 21.0% |
| Respiratory Rate | breaths/min | 27.1% | 45.4% |
| ALSFRS-R |  | 0.0% | 36.1% |
| Adverse Event | Total Number of Adverse Events |  | 0.0% | 0.0% |
| Number of Adverse Respiratory Events |  | 0.0% | 0.0% |
| Number of Adverse Nervous Events |  | 0.0% | 0.0% |
| Number of Adverse Psychological Events |  | 0.0% | 0.0% |
| Number of Adverse Metabolic Events |  | 0.0% | 0.0% |

(A) Percentage of Missing Data for Demographic, Clinical and Adverse Event Features

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Category of Laboratory Parameter** | **Parameter** | **Units** | **Percentage Missing in Training and Test Data** | **Percentage Missing in PROACT Database** |
| Muscle | ALT | U/L | 21.2% | 5.2% |
| AST | U/L | 21.4% | 9.5% |
| Creatinine | μmol/L | 21.1% | 21.0% |
| CK | U/L | 32.1% | 39.8% |
| Metabolic | Total Cholesterol\* | mmol/L | 43.1% | 52.1% |
| Triglycerides\* | mmol/L | 46.9% | 55.1% |
| HbA1c\* | % | 62.0% | 71.4% |
| Glucose | mmol/L | 21.2% | 25.0% |
| Phosphorus | mmol/L | 28.8% | 44.3% |
| Electrolytes | Sodium | mmol/L | 21.1% | 22.4% |
| Potassium | mmol/L | 21.3% | 22.6% |
| Chloride | mmol/L | 24.3% | 24.6% |
| Calcium | mmol/L | 22.0% | 32.5% |
| Bicarbonate | mmol/L | 30.7% | 30.2% |
| Kidney | BUN | mmol/L | 21.1% | 24.9% |
| Uric Acid\* | μmol/L | 72.8% | 70.6% |
| Liver | Alkaline Phosphatase | U/L | 0.1% | 28.8% |
| GGT | U/L | 43.7% | 39.2% |
| Bilirubin | μmol/L | 21.1% | 37.0% |
| Nutrition and Liver | Protein | g/L | 29.3% | 39.6% |
| Albumin | g/L | 22.0% | 34.5% |
| Blood and Immune Parameters | Hb | g/L | 1.6% | 17.6% |
| Haematocrit | % | 22.0% | 17.9% |
| RBC | 10E9/L | 28.1% | 18.4% |
| WBC | 10E9/L | 28.1% | 18.4% |
| Platelets | 10E9/L | 29.1% | 25.3% |
| Absolute Eosinophil Count | 10E9/L | 30.8% | 27.5% |
| Eosinophils Percentage | % | 41.3% | 49.0% |
| Absolute Lymphocyte Count | 10E9/L | 42.6% | 35.1% |
| Lymphocyte Percentage | % | 29.5% | 41.3% |
| Absolute Basophil Count | 10E9/L | 36.4% | 35.2% |
| Basophil Percentage | % | 29.6% | 41.3% |
| Absolute Monocyte Count | 10E9/L | 42.6% | 35.1% |
| Monocyte Percentage | % | 29.5% | 41.3% |
| Absolute Neutrophil Count | 10E9/L | 42.6% | 35.1% |

(B) Percentage of Missing Data for Laboratory Parameters

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Parameter** | **Units** | **Percentage Missing** |
| Demographic | Sex^ |  | 0.0% |
| Age^ | years | 0.0% |
| Site (bulbar/limb/both/other)^ |  | 0.0% |
| Clinical | Days since Onset of Symptoms | days | 0.0% |
| Days since diagnosis | days | 0.0% |
| Riluzole Use |  | 0.0% |
| BMI | kg/m^2 | 26.8% |
|  | |  |
| FVC (as % of Predicted for Normal Individual)^ | proportion | 26.6% |
| Diastolic Blood Pressure | mmHg | 25.2% |
| Systolic Blood Pressure | mmHg | 25.2% |
| Pulse | bpm | 25.2% |
| Respiratory Rate | breaths/min | 25.6% |
| ALSFRS-R |  | 0.0% |
| Adverse Event | Total Number of Adverse Events |  | 0.0% |
| Laboratory | ALT | U/L | 16.8% |
| AST | U/L | 16.8% |
| Creatinine | μmol/L | 16.6% |
| Glucose | mmol/L | 99.2% |
| Sodium | mmol/L | 16.5% |
| Potassium | mmol/L | 16.6% |
| Chloride | mmol/L | 16.5% |
| Bicarbonate | mmol/L | 16.6% |
| BUN | mmol/L | 16.7% |
| Bilirubin | μmol/L | 17.3% |
| Hb | g/L | 16.8% |
| Haematocrit | % | 16.8% |
| RBC | 10E9/L | 16.8% |
| WBC | 10E9/L | 16.8% |

(C) List of Features Used in Holdout Validation Dataset, the Trial of Celecoxib, and the Percentage of Missing Data.

^ patients missing this data were excluded, due to the importance of these parameters, and them being used for missing data imputation

\* factors were included despite being missing in >50% of patients, due to clinical importance

BMI: Body Mass Index

ALT: alanine transaminase

AST: aspartate aminotransferase

CK: creatine kinase

HbA1c: haemoglobin A1c test

BUN: blood urea nitrogen

GGT: gamma-glutamyl transpeptidase

Hb: haemoglobin

RBC: red blood cell

WBC: white blood cell

Supplementary Table 1: Percentage of Missing Data in Parameters