This document provides the statistical modelling and the analysis done in order to assess if the probability of type of hospital admission (elective, emergency or urgent) can be described by the length of stay, race (white or not), age group and if the patient subsequently deceased or not . The goal of this document is to provide insights into the modelling strategy and the inference methods used.

For the purpose of this analysis, the data available represents records from 1495 patients for which the length of stay, race, age category death and hospital identification number was recorded. In order to answer the question of interest the investigation used a technique called “regression analysis”, a method which incorporates several properties that could potentially explain the parameter of interest (in this case hospital admission type) and selects only the properties that actually influence the hospital admission type. The model that was used in this analysis is a multinomial model, which was considered appropriate given the nature of the data and the fact that the answer expected is the probability of each hospital admission. A multinomial model is a model that for a set of parameters returns the probability of each outcome, in this case, the probability of each of the three admission types. Out of the properties available, the model concludes that the probability of elective, emergency and urgent hospital admission is best described by a combination of length of stay, race and death while age was not considered useful for the purpose of the modelling (p-value 0.62). In order to assess which properties are statistically significant we used p-values and removed from the model any properties that were not considered relevant for the analysis.

For a patient that is not deceased, his race is not white, and spent 0 days in the hospital, the odds of emergency against elective admission is 0.03, where odds denote the ratio between the probability of being admitted in an emergency and the probability of being admitted in an elective manner. This means that the probability of the patient to be admitted in an emergency is 3% of the probability to be admitted electively, when they have the caracteristics described above. Additionally, the odds of urgent against elective is 0.26, meaninging that the probability of being admitted urgently is equal to the probability of being admitted electively times 0.26. It is worth mentioning that these results are both theoretical as a patient that does not spend any days in the hospital would not be included in the data provided, and not spending any days in the hospital is outside of the range of values that length of stay takes in the dataset.

Regardless of the race or survival status, for a patient admitted to hospital, the probability of being admitted electively decreases with the increase of number of days spent in the hospital, while the probability of being admitted in an emergency increases. Similarly, the probability to be admitted urgently increases up to about 50 days in the hospital after which starts to decrease. For a white person the probability of being admitted electively when they left the hospital after a day is about 86% (with 95% confidence that probability is between 83% and 88%) decreasing to 17% when they leave the hospital after 60 days. There is a 2% probability for a patient to be admitted in an emergency and 13% urgently if they left the hospital after 1 day and increase to 56% (between 35% - 74% with 95% confidence) and 27% (with 95% confidence that the percentage is between 14% and 45%) respectively, after 60 days .

Based on the results mentioned, we can conclude that the three properties - race, subsequent death and length of stay are useful to predict hospital admission type, however, predictions are only reliable in the interval for which data is provided. In other words, the analysis can only describe an outcome for a patient that has a length of stay in hospital between 1 and 116 days. On unseen data, randomly selected from the initial dataset, the model had 76% accuracy, calculated by how many of the predictions are correct, where a prediction represents the outcome with the highest probability out of the three probabilities returned by the model. Furthermore, the model selected requires a few assumptions to be respected in order to provide accurate results. While some of the assumptions are correct, the checks were mostly indicative. In reality, the model is in line with a rational thought - patients that leave the hospital after just a few days, have a higher probability to be admitted electively, as we would expect, these are not necessarily severe illnesses compared to patients that spent a higher number of days in the hospital which could be admitted urgently or in an emergency.