README – Multi-state COVID-19 length of stay and forecasting model

Objectives The attached code describes a model for predicting the number of beds occupied by COVID-19 patients.

For forecasting the hospital occupancy, we consider three aspects. Firstly, we fit length of stay distributions to each hospital state, which describes how long individuals stay in a given state before moving elsewhere. Secondly, we fit the proportion of patients going down each hospital pathway. We fit these two factors simultaneously by using a multi-state competing risks model, where the shape of the risk function describes the proportion of cases following a certain pathway and the time until the transition associated with that pathway occurs. Based on these, we can take hospital admissions data and simulate the number of people in each bed over the coming weeks/months.

The third aspect we consider is the potential admissions trajectories. For the admissions up to today we can take the actual data and project where these individuals may end up. However, this does not help plan occupancy more than a few days ahead since individuals are regularly being admitted to the hospital. Therefore, we can construct potential future admissions streams based on different assumptions, which will allow us to project bed occupancy scenarios into the future. In the multi-state model, we consider two possible scenarios, but other cases can be added as required. The first scenario assumes exponential growth in the number of admissions and the second assumes a constant number of admissions per day. For each of these scenarios, we can define the current number of admissions and/or the exponential growth rate in admissions, and the model will propagate these forwards for the number of days of interest. A limitation of the exponential growth scenario is this does not consider depletion of susceptibles, which may cause the exponential growth to stop and turn over. Therefore, this will be most accurate in the next month or two, and lose accuracy after that when susceptible depletion may start playing a role.

Important note

Length of stay and outcome probabilities have changed at the majority of trusts over the course of the pandemic, even once changes in age distributions have been accounted for. Generally, outcomes appear to have improved, as clinical treatment and policy changes have improved. This means that fitting a single hazard function for each event over the entire epidemic may not be an accurate reflection of the current patient dynamics. To address this, one could add time as a covariate to the model. However, we have opted to not add this extra complexity to the modelling process. Instead, recent trends can be captured by restricting the subset of data that we use to fit the model.

By only using more recent admissions to hospital, we capture the length of stay and outcome trends of the current patients. When deciding which data to include, a balance needs to be struck between using only the most recent data and retaining a sufficiently large sample size to get reliable estimates. We recommend including at least 500 patients in the sample, if possible, ideally using between 1000-2000 patients, though this depends on the size of the hospital.

Code

Process_Data_Table – this code takes as input patient level data and processes it into the format required for the multi-state model and generates a hospital census for comparison. This breaks individuals into four age categories (0-25, 26-50, 51-75, 76+) as this allows the dynamics to be more accurately captured.

Fit_And_Model – using the processed data from *Process_Data_Table* as input, this fits the multi-state model and simulates future scenarios. This can fit either Weibull or Burr distributions to the length of stay distributions, or a combination of the two which we observed to work best. The bed occupancy can then be simulated. Results are output in figures showing occupancy over time as well as a table recording weekly occupancy.

Fit_Admissions_Growth – this is an example of a method that can be used to fit an exponential growth rate to the admissions data. We have been taking the last four weeks of daily admissions in each age group and fitting an exponential growth rate using a generalised additive model. This estimated rate can then be used as input for the exponential growth scenario in the multi-state model.

Contact

We are happy to discuss and help users with implementation of the model. Please contact:

Christopher Overton

Christopher.overton@manchester.ac.uk

Data Scientist, Clinical Data Science Unit, Manchester University NHS Foundation Trust

Research Associate, Department of Mathematics, University of Manchester