## Data

File named “COUGH\_wideform” is the original data used for calculating the transition probability matrix (TPM). This data can be used to reproduce Table 1 and Table 2. Any software program like SPSS can be used to generate the cumulative 3 by 3 TPM matrices for each category. Then using the R code file named “meanpassagetime.R” the mean first passage time for each TPM and its 95% confidence bounds can be calculated to reproduce the Table 3. Description of variables: id is the unique number describing each patient, age is a continuous variable denoting the age which was categorized to (1 is <40, 2 is 40-60, 3 is >60), gender is 0=male 1 = female, day0-day7 are the state of cough on the 8 different days (1=no, 2=mild, 3=moderate/severe), lv is left ventricle ejection fraction in % which was categorized to (1 is <49% and 2 is >=49%), bh is breath holding time in seconds which was categorized to (1 is <20 and 2 is >=20), vd is ventilator duration in hours (1 is <12,2 is 12-24, 3 is >24), sur is surgery type (1=CABG,2=Valve surgery,3=Septal surgery,4=Multiple surgery).

File named “COUGH.xlsx” is the longform data prepared from the original data for running the semi-parametric regression models in SAS using the code file “AG\_PWPmodels.sas”. This will reproduce Table 4. Description of variables: id is the unique number describing each patient, regist keeps track of the number of times a specific type of change in state has occurred within each patient, typetran is the type of transition where 1 = worsening of cough and 2 = resolution of cough which is a simplification of the typetran\_name variable which shows the exact type of transition (1\*2, 1\*3 and 2\*3, which are worsening of cough and 3\*1, 3\*2 and 2\*1 which are improvement in cough), time1 is the start of the time interval and time2 is the end of the time interval in which a state change occurred or if no change occurred the last time point available, event = 1 means a state change occurred and event = 0 means no change, age is a continuous variable denoting the age which was categorized to (1 is <40, 2 is 40-60, 3 is >60), gender 0=male 1 = female, lv is left ventricle ejection fraction in % which was categorized to 1 is <49% and 2 is >=49%, bh is breath holding time in seconds which was categorized to 1 is <20 and 2 is >=20, vd1 vd2 vd3 are dummy variables created for vd\_original which is ventilator duration in hours 1 is <12,2 is 12-24, 3 is >24), sur1 sur2 sur3 sur4 are dummy variables created for sur\_original which is surgery type (CABG,Valve surgery,Septal surgery,Multiple surgery), order keeps track of the number of times each patient ID is repeated in the rows, noofchanges indicates the number of changes undergone by a patient in total, new serial indicates the new serial number after the wide to long form transformation in completed.

Data will be stripped for personal identifying information before being made available for readers. Data will be provided only upon a written and valid request submitted by the reader to the corresponding author. Upon receiving the request, the study authors’ will request study institution’s approval for sharing the data.

## Code

The codefile “meanpassagetime.R” is a R code file. It contains 3 parts. Part I supplies the transition probability matrix information to the subsequent code. Part II and III run the Monte-Carlo simulation (n=1000) to obtain the mean first passage time and its 95% confidence bounds. More information on the codes is available here: <https://www.tandfonline.com/doi/abs/10.1080/02664763.2016.1143454>. This code will reproduce the information provided in Table 3. This code file is one of the many because for each TPM we have to change the matrix given in the part I.

The codefile “AG\_PWPmodels.sas” is a SAS code file. It calls the PTRANSIT SAS macro package to run the semiparametic regression models, namely AG and PWP. More information on the arguments of the PTRANSIT command can be found here. <https://www.sciencedirect.com/science/article/pii/S0169260703001329>. This code will reproduce the information provided in Table 4.