FLoSC
Forecasting Length of Stay and Cost

Introduction
Methods
User Guide
Tutorial <<
Downloads
Installation
Links
Contact

This tutorial is designed to give users a step-by-step guide to using FLoSC and the interpretation of the reports.

Content

---

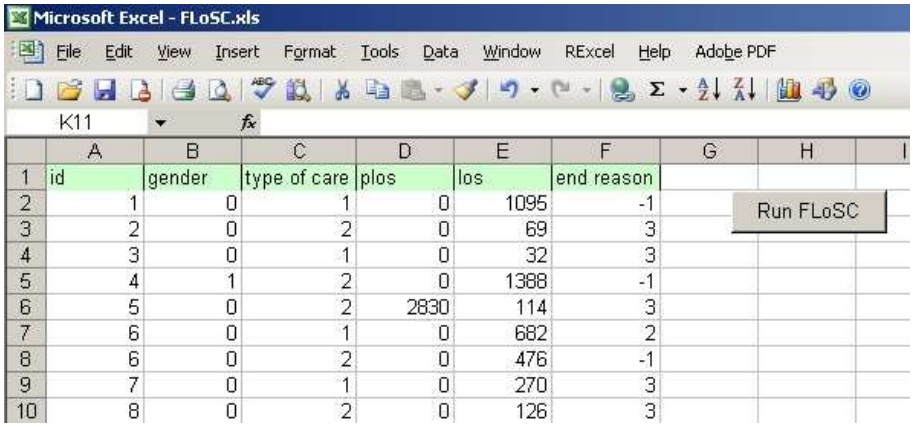
1. The data
The data used in this tutorial is generated by computer simulation. You can download the dataset here. Please have a look at the format of the data in reference to the Data requirement section of the User Guide. Also notice that, for illustration purpose, eight additional records containing missing values are appended to the simulated data.

---

2. The context
Suppose today is 7/9/2007. A local council in the Greater London area is interested in using FLoSC to forecast the cost due to their current known commitment today until the end of next financial year (i.e. 1/4/2009). Relevant data from 20/4/2003 to 7/9/2007 has been extracted from a local information system and necessary process has been carried out so that the final data is compatable with the data requirements of FLoSC.

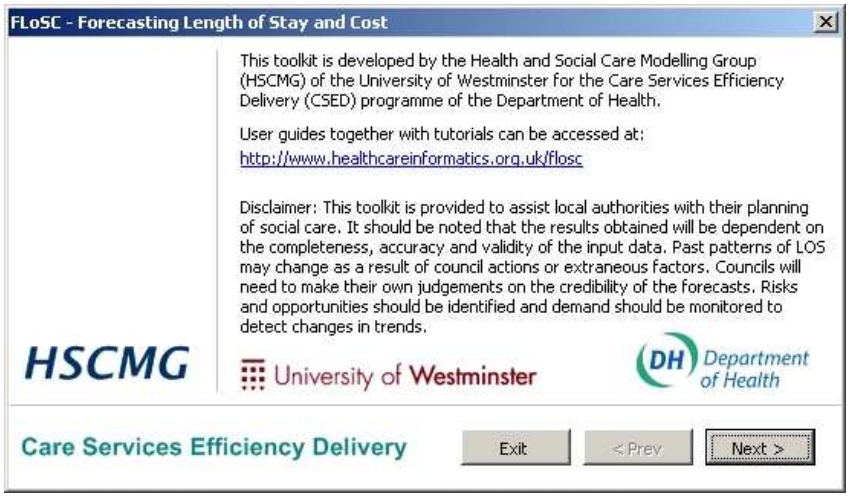
---

3. Using FLoSC
Start FLoSC from the Start Menu. Copy-and-paste the tutorial data into FLoSC, delete any previous data if necessary. Verify that the heading and order of the columns are consistent with the conditions set out in the Data requirement section. Your screen should look like the following.

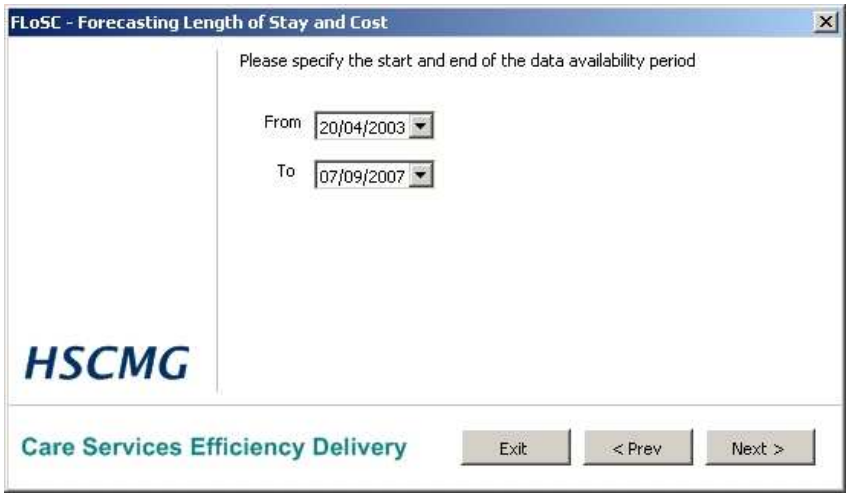


	A	B	C	D	E	F	G	H
1	id	gender	type of care	plos	los	end reason		
2	1	0	1	0	1095	-1		
3	2	0	2	0	69	3		
4	3	0	1	0	32	3		
5	4	1	2	0	1388	-1		
6	5	0	2	2830	114	3		
7	6	0	1	0	682	2		
8	6	0	2	0	476	-1		
9	7	0	1	0	270	3		
10	8	0	2	0	126	3		

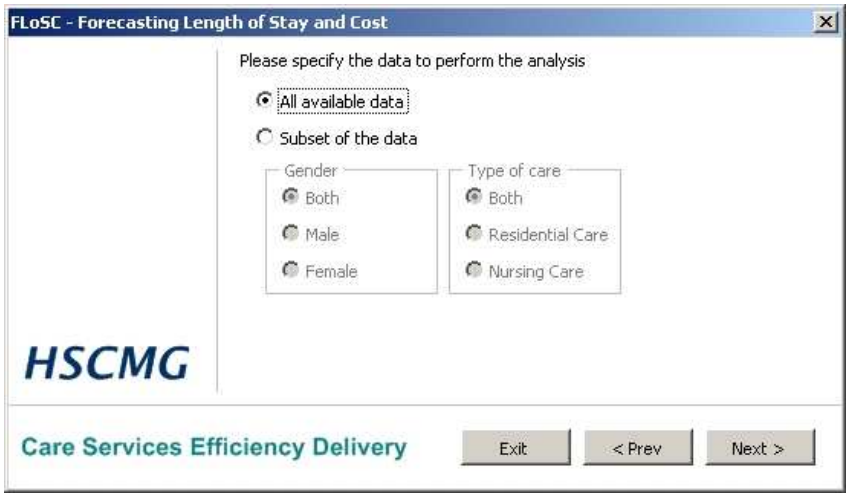
Start the FLoSC interface by clicking on the **Run FLoSC** button. The following screen will appear.



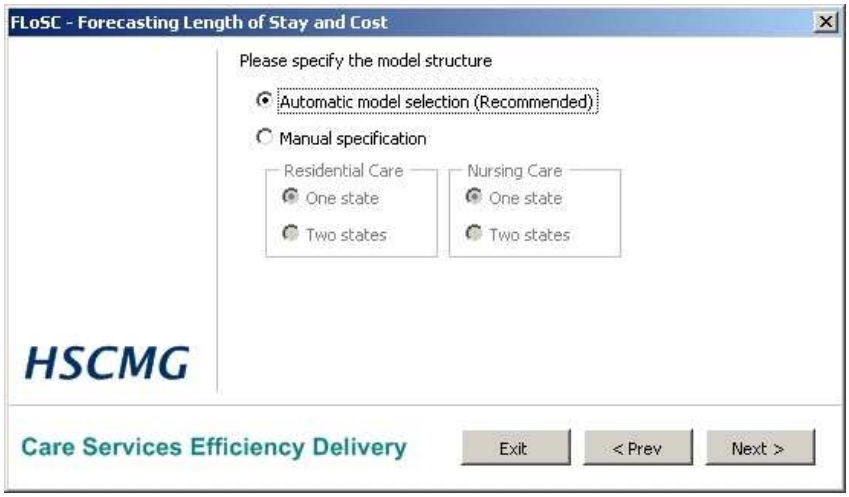
Read the disclaimer. Click **Next** to continue. No unexpected values are encountered in the data, and the following screen appears.



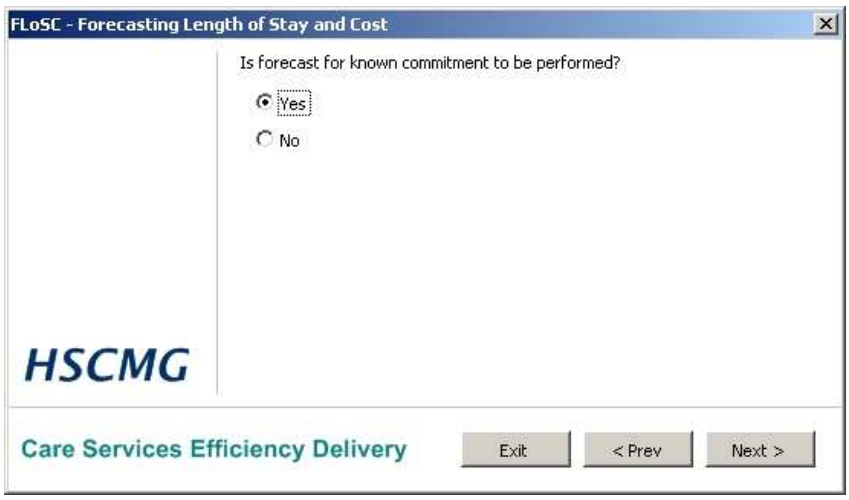
In this screen, we need to inform FLoSC the start and end date of the data availability period. In this case, it is from 20/4/2003 to 7/9/2007. Click on the drop-down arrow and a calendar will appear. Once the dates are set, click **Next** to continue.



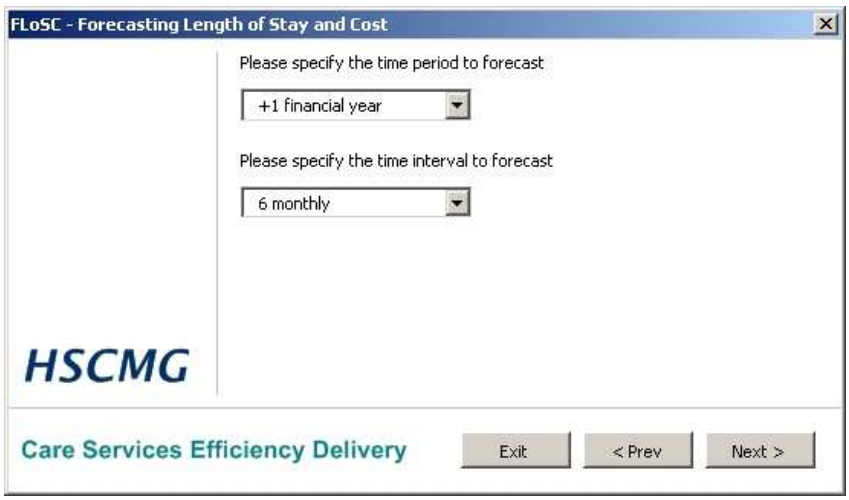
In this tutorial, we will use all the data for analysis. Click **Next** to continue.



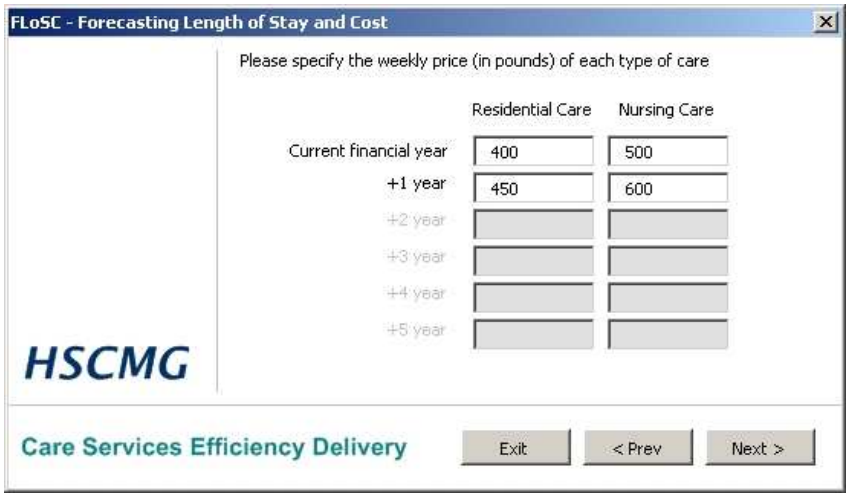
We will let FLoSC to decide the appropriate model structure, e.g. should there be only one state or a combination of a short-stay state and a long-stay state. Click **Next** to continue.



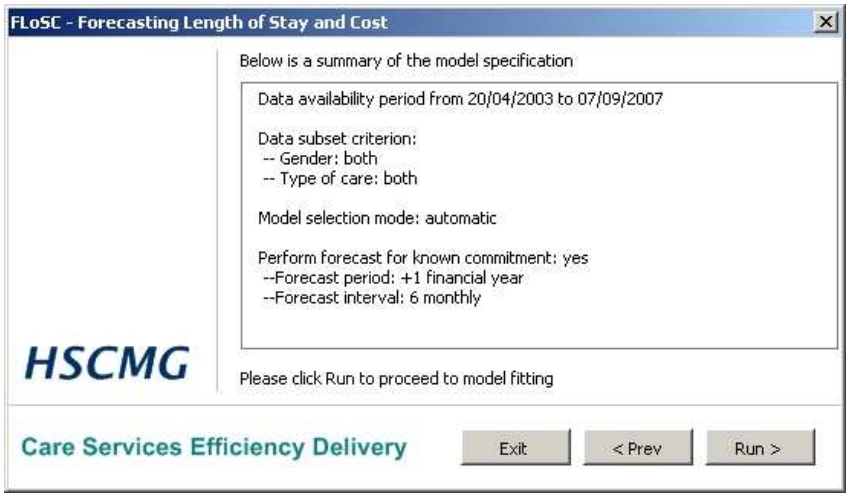
Select Yes to produce cost forecast for known commitment on 7/9/2007. Click **Next** to continue.



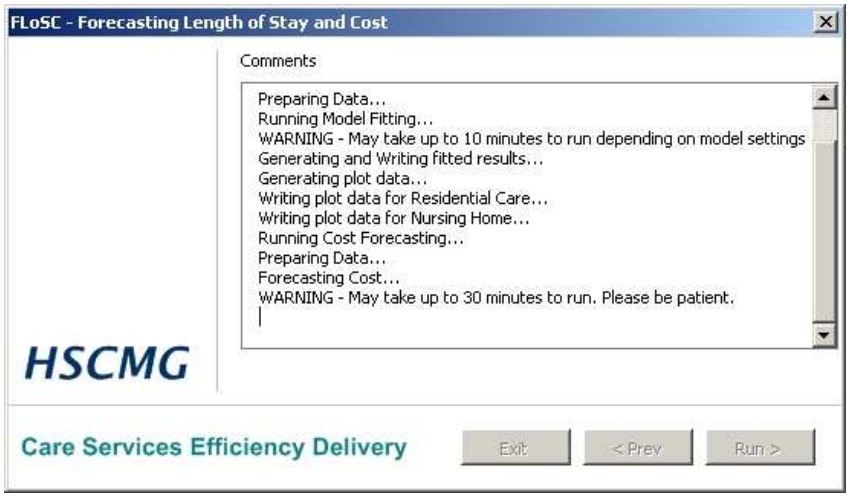
The forecast period is to one financial year after the current one. In other words, we want FLoSC to produce cost forecast for their known commitment until 1/4/2009, and the forecast in on every 6 monthly. Click **Next** to continue.



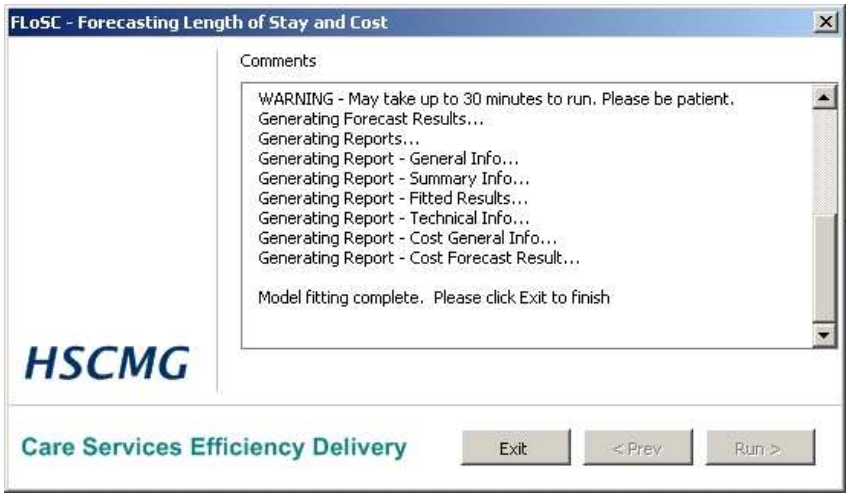
Since our forecast period covers two financial years -- the current financial year and one after, we need to provide weekly price for RC and NC for two years. Suppose, in the current financial year, the average weekly prices in the local area are £400 for RC and £500 for NC, and in the following financial year they are expected to go up to £450 and £600 for RC and NC respectively. Type in the relevant figures and click **Next** to continue.



A final review of the settings, then click **Run** to start the model fitting and forecasting.



During running, FLoSC prints out messages indicating the task in process. Some tasks will take time to run.



This message indicates that FLoSC has finished running. Click **Exit** to finish.

---

4. Results
FLoSC produces the results in the form of additional worksheets within the current workbook.

They are presented in a set of reports, namely:

- "*Report - LOS - General Info*" -- contains general information about the data and the data cleaning process.
- "*Report - LOS - Summary Info*" -- contains summary information on patterns of LOS and movements of residents within the LTC system.
- "*Report - LOS - Fitted Results*" -- contains information on the fitted results.
- "*Report - Cost - General Info*" -- contains general information about cost forecasting, such the weekly price, forecasting period and forecasting intervals.
- "*Report - Cost - Forecast Result*" -- contains description of the forecasted results on cost.
- "*Report - LOS - Technical Info*" -- contains technical information on the fitted Markov model that captures the survial and movement of residents in LTC system.

In the following sections, we will demonstrate how to interpretate these reports.

4.1 Report - LOS - General Info

This is the first part of the report, which contains general information about the data and the analysis, such as the time and date the analysis was carried out, the data availability period specified by the user, the number of records in the data.

Report: Length-of-stay Analysis

This analysis was carried out on 2007-10-09 at 16:21:58

General information

number of records	2144
period covered	between 2003-04-20 and 2007-09-07

Data cleaning and data processing

number of record(s) deleted due to missing (or unknown) value	5
---- due to missing (type of care)	1
---- due to missing (end reason)	1
---- due to missing (los)	2
---- due to missing (plos)	1
data subset criterion	gender = both ; type of care = both
number of record(s) deleted due to movement from NC to RC	2
number of record(s) in the working dataset	2137

This report also contains information on data cleaning, which is conducted before any analysis is carried out. In the above table, we can see that 5 records were deleted due to missing value -- 1 record due to missing "type of care"; 1 record due to missing "end reason"; 2 recordw due to missing "los"; and 1 missing "plos".

There are 2 records deleted due to movement from NC to RC, which is not modelled by FLoSC.

Although not the case in this tutorial, FLoSC will also delete records associated with residents who have more than 2

records in the data. As outlined in the Data requirement section of the User Guide, FLoSC expects at most one records each for stays in RC and/or NC. Therefore, for any residents having more than 2 records would represent an invalid entry.

The data subset criterion used during the analysis is reported.

After data cleaning and data selection based on subset criterion, the final working dataset contains 2137 records.

4.2 Report - LOS - Summary Info

This part of the report contains summary about the working data. The following table shows the frequency distribution of residents in RC and NC by gender.

Summary information on the working dataset

Frequency table (by gender)

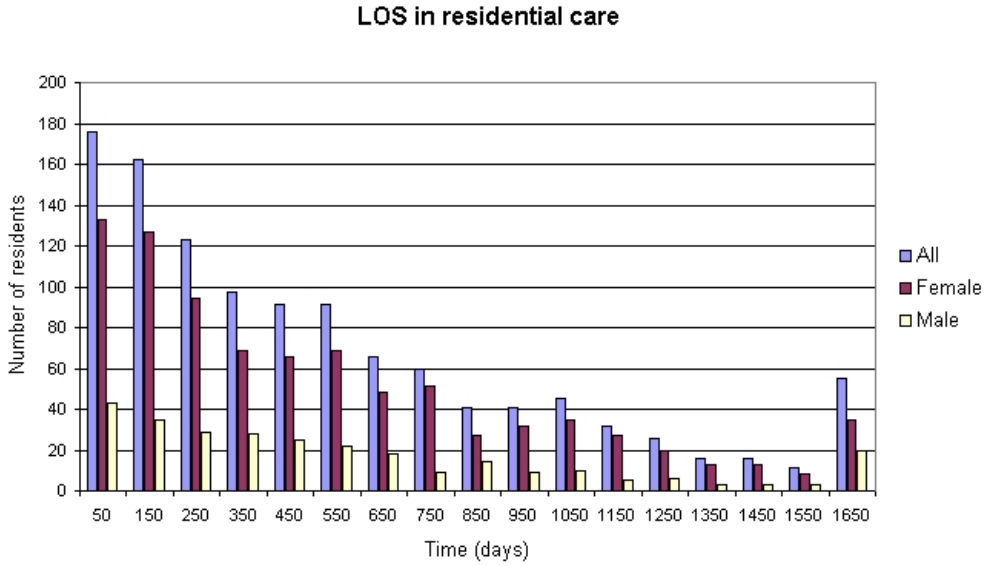
	RC	NC	Total
female	867	753	1620
male	282	235	517
total	1149	988	2137

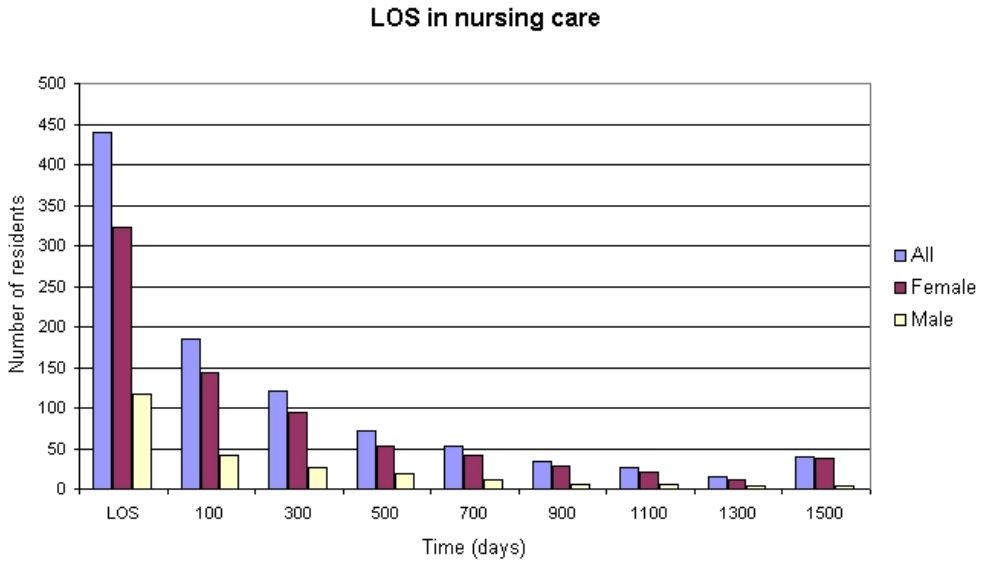
The following table (screenshot is truncated here) shows the typical summary statistics on length-of-stay, stratified by type of care and by gender. Basic measures, such as average (i.e. mean), median and standard deviation (stdev) are reported.

Summary statistics on length of stay

	count	mean	stdev	Q1
All RC residents	1149	800.6	788.7	
All NC residents	988	616.7	775.8	
All female residents in RC	867	793.8	791	
All male residents in RC	282	821.5	782.4	
All female residents in NC	753	642.2	796.7	
All male residents in NC	235	534.7	699.7	
All residents with missing gender in RC	0	NA	NA	
All residents with missing gender in NC	0	NA	NA	

Although the figures reported in the above table are useful in general, the distribution of LOS is best presented as a graph. The following two graphs are the histograms of LOS for each type of care, also stratified by gender.





The most obvious observation is that the overall pattern for the distribution of LOS is an exponential decline for both types of care and both genders. There are residents who stay in care substantially longer than the average shown in the table of summary statistics. As often with LOS data, due to the strong skewness, median LOS is often a better summary than average LOS.

It is probably worth keeping in mind that this pattern of exponential decline for the distribution of LOS data in LTC is very common. Therefore, if your data does not show such a pattern, it is often a hint that there might be problems with the data, probably due to data extraction error during the data preparation stage. If this is the case, review the percedure of data extraction before interpreting the results.

Also notice that the last set of bars seem odd as they appear to differ from the decreasing trend. This is actually quite normal as they often include all the residents who were present during the entire data availability period.

FLoSC also generates a summary on the movement of residents during the data availability period. This is useful in giving an overview of how residents move within the system.

Summary on the movement of residents

total number of residents present on 2003-04-20	636
---- in RC	382
---- in NC	254
number of admissions to RC during the period	767
number of residents died in RC during the period	628
number of residents transferred to NC during the period	137
number of residents still living in RC at the end of the period	384
number of direct admissions to NC during the period	597
number of residents still living in NC at the end of the period (including those transferred from RC)	259
number of residents died in NC at the end of the period (including those transferred from RC)	729
number of residents died in NC at the end of the period (among those transferred from RC)	93
number of residents still living in NC at the end of the period (among those transferred from RC)	44

The above table shows that, in this particular dataset, there were 636 residents present in RC (382) and NC (254) on 20/4/2003, which is the starting date of the data availability period. During the data availability period, there were 767 admissions to RC. Among all the RC residents, 628 died during the period, 137 were transferred to NC and 384 were still present in RC on 7/9/2007, the end of the data availability period. There were 597 direct admissions to NC during the period, and including those transferred from RC, 259 residents were still present in NC, whereas 729 residents died. Among the 137 RC residents who were transferred to NC, 93 of them died by the end of the data availability period and 44 were still present in NC.

4.3 Report - LOS - Fitted Results

This part of the report contains description on the fitted results of the model capturing the pattern of survival and movements of publicly funded residents in RC and NC.

First, FLoSC reports the structure of the model and the selection method specified by the user. In this case, we chose automatic model selection and FLoSC suggests that a model with 1 state in RC and two states in NC provides a good representation for the movement of residents in the system.

Model fitted results

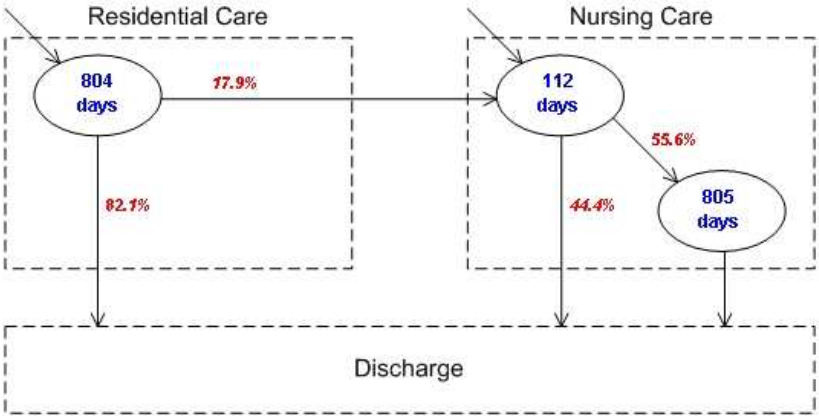
Model structure

RC	1 state
NC	2 states

model selection mode:	automatic

This is followed by a diagram showing the structure of the model together with fitted information. It gives more insights into the patterns of movement of publicly funded residents in the system.

Residents' movements and patterns of LOS

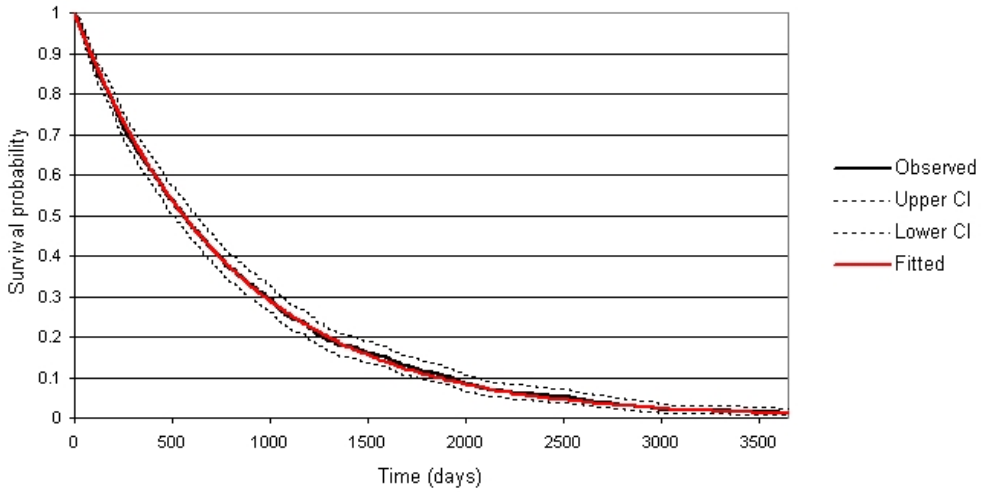


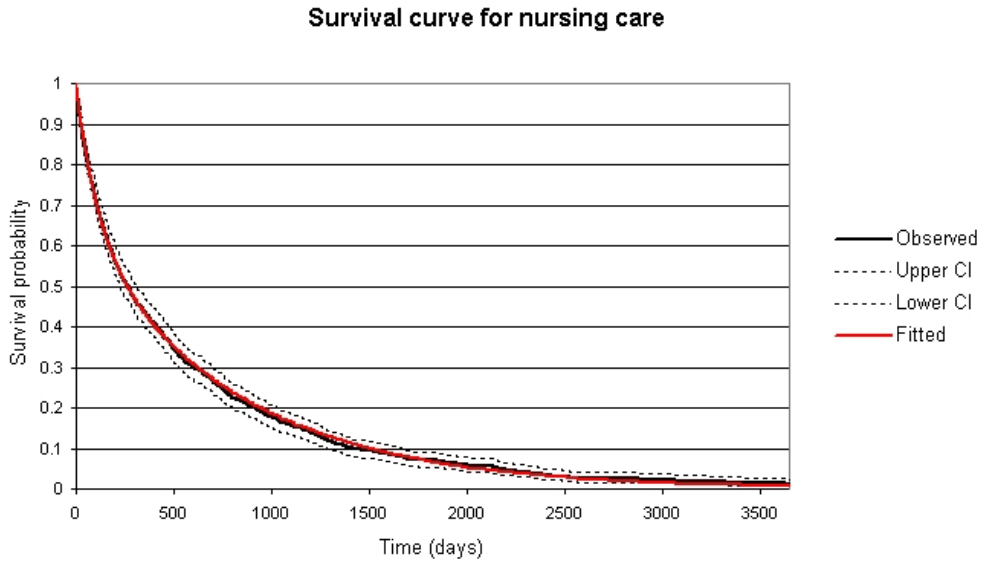
In this case, there is one state in Residential Care (RC) with an average LOS of 804 days (about 2.2 years). Remember that the LOS in a state follows exponential distribution, which has a long tail. Upon leaving RC, majority of them (82.1%) will be discharged (including discharged by death), whereas about 18% will be transferred to Nursing Care (NC). A single state in RC suggests that residents are leaving RC on a harmogenous fashion, however that does not mean residents are leaving in exactly the same way.

There are two states for NC -- a short-stay state with average LOS of 112 days and a long-stay state with average LOS of 805 days (about 2.2 years). About 56% of them will settle down and become long-stay residents, and about 44% will be discharged relative early during their stay in NC.

Next part of the report is the plot of survival curves in both types of care. Survival curve is a graphical representation on the probability of a resident staying longer than the time on the x-axis, and the curve will always be a declining line start from one when time is zero. The following plot of fitted survival curves (i.e. predicted by the model) against the observed curves estimated from the data can be used to judge how well the model fit the data, hence providing a mean of checking if the results provided by the model are supported by the observed data.

Survival curve for residential care





In both of the plots above, the black solid line represents the observed survival curve suggested by the data, the black dotted lines represent the associated 95% confident band; whereas the red solid line is that predicted by the survival model. Clearly the model is able to capture the survival pattern in both types of care as the fitted survival curves are all within the 95% confident band and follows the black solid line closely.

4.4 Report - Cost - General Info

This part contains a brief summary information on known commitment, which are the residents in care on 7/9/2007.

Report: Cost Forecast - Summary Information

This analysis was carried out on 2007-10-09 at 16:26:56

Summary information on known commitment

number of residents in system on 2007-09-07	643
data subset criterion	gender = both ; type of care = both
number of residents in working dataset	643
--- in RC	384
--- in NC	259

This table shows that on 7/9/2007 there were 643 residents in care. Among them, 384 were in RC and 259 were in NC. The data subset criterion specified by the user is also reported. Since in this tutorial, we specified to use all data, which includes both gender and both type of care, the working dataset also has 643 residents.

Frequency table (by gender)

	RC	NC	Total
female	284	205	489
male	100	54	154
total	384	259	643

The above table show the number of residents by gender and by type of care.

Summary statistics on length of stay

	count	mean	stdev	Q1
All RC residents	384	822.3	776.3	
All NC residents	259	807.7	862	
All female residents in RC	284	803.3	732.5	
All male residents in RC	100	876.4	890.9	
All female residents in NC	205	879.6	902.3	
All male residents in NC	54	534.7	622.2	
All residents with missing gender in RC	0	NA	NA	
All residents with missing gender in NC	0	NA	NA	

This table gives a brief summary statistics on LOS of the known commitments.

4.5 Report - Cost - Forecast Result

The result of the cost forecast for known commitment is given in this section.

Report: Cost Forecast - Forecasted Costs Due to Known Commitment

Forecasting period: from 2007-09-07 to 2009-04-01 year(s)
Forecasting interval: 6-monthly

Weekly cost of care

Financial year	RC	NC
2007/08	400	500
2008/09	450	600

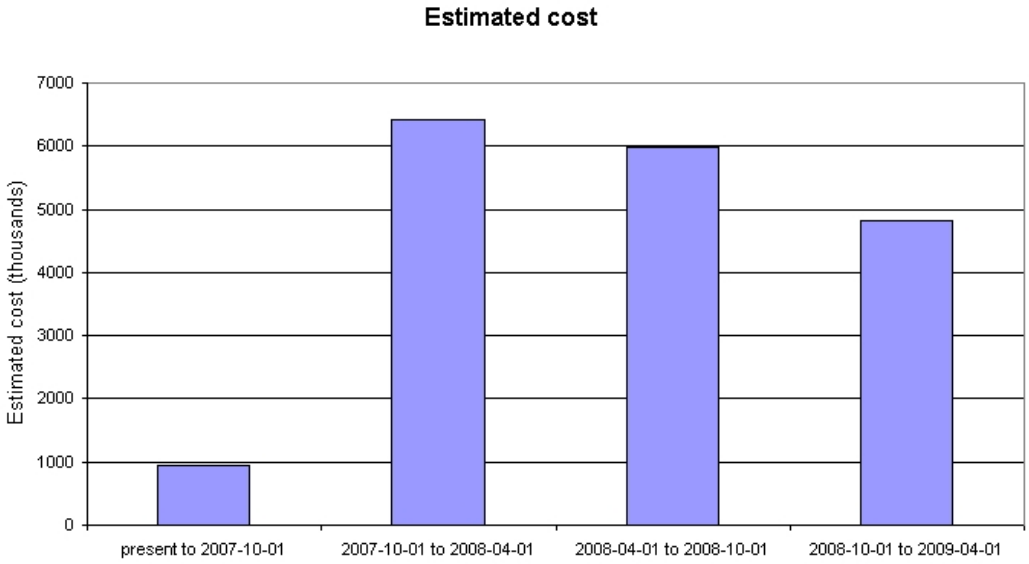
First, the input information specified by the user is report, such as the forecasting period, forecasting interval and the weekly cost of care.

Projected total cost of current commitment

Financial period	Total cost
present to 2007-10-01	955418
2007-10-01 to 2008-04-01	6411815
2008-04-01 to 2008-10-01	5980944
2008-10-01 to 2009-04-01	4819921

Then the projected total cost of maintaining the current known commitment is given in the above table. Based on the survival model derived from past data concerning residents, also taking into account their stay so far in the system, FLoSC forecasts that the cost of maininging the 643 residents currently presents to 1/10/2007 (i.e. the starting date of the second half of the current financial year) will likely be just under £1 million (gross). Please note that this is the gross cost to local council. The actual cost for each resident will heavily depend on individual circumstances, such as if they have a pension, etc.

The cost arised from this group of residents is projected to be over £6.4 million the second half of the financial year, etc.

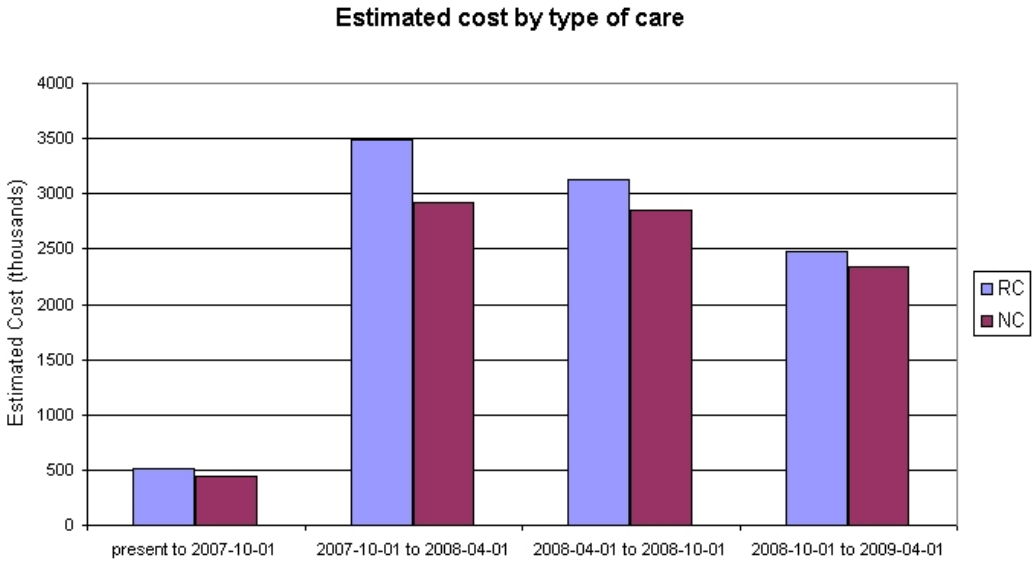


The projected figures is also presented graphically above. Notice that a year later (i.e. the second half of the coming financial year), the existing commitments today will still be costing almost £5 million to maintain.

The projected tocal cost due to known commitment is also broken down to show the cost due to each type of care. This information is also presented graphically in the following plot.

Break-down of the projected total cost by type of care

Financial period	RC	NC
present to 2007-10-01	518869	436548
2007-10-01 to 2008-04-01	3487233	2924582
2008-04-01 to 2008-10-01	3126751	2854193
2008-10-01 to 2009-04-01	2479862	2340058



4.6 Report - LOS - Technical Info

This section of the report contains more technical information about the survival model, which underpins the cost forecast. The information is included for users who are more technically minded.

Model fitted results (details)

Fitted model parameters

parameter	estimate	s.e.	cil	ciu
q_{12}	0.000223	0.00002	0.000188	0.000263
q_{23}	0.00496	0.00124	0.003037	0.008099
q_{14}	0.001021	0.00004	0.000944	0.001104
q_{24}	0.003963	0.00037	0.003302	0.004755
q_{34}	0.001243	0.00008	0.001091	0.001415

This table summarises the estimates of the fitted parameters, i.e. the transition rates of the Markov model.

Fitted survivor function

RC	exp(-0.00124*x)
NC	0.354*exp(-0.00892*x) + 0.646*exp(-0.00124*x)

The above are the fitted survivor function for RC and NC, which describes the red solid lines in the plot of survival curves.

Fitted transfer probability from RC to NC

transfer prob	0.179

The transfer probability from RC to NC is estimated to be about 18%.

Last updated: 2007-10-10