Methods for Analyzing Hospital Length of Stay with Application to Inpatients Dying in Southern Thailand

Apiradee Lim

Department of Mathematics and Computer Science

Faculty of Science and Technology, Prince of Songkla University

Pattani Campus 94000, Thailand

Tel: 66-7331-2179   E-mail: lapirade@bunga.pn.psu.ac.th

Phattrawan Tongkumchum (Corresponding author)

Department of Mathematics and Computer Science

Faculty of Science and Technology, Prince of Songkla University

Pattani Campus 94000, Thailand

Tel: 66-7331-2179   E-mail: tphattra@bunga.pn.psu.ac.th

Abstract

This study investigated length of stay (LOS) for patients who died in hospital in Southern Thailand from 2000 to 2003 with respect to principal diagnosis and demographic, geographic and hospital size factors. The computerized data of 40,498 mortality cases were obtained from the Ministry of Public Health from 167 hospitals in 14 provinces of Southern Thailand between October 2000 and September 2003 with information on age, gender, principal diagnosis, province and hospital size. Logistic and linear regression with log-transformed LOS was used to analyze the data. Patients with injuries as principal diagnosis had shortest LOS, whereas cancer patients had the longest LOS. Older patients, particularly females, had higher LOS for all diagnoses. LOS increased with hospital size except in the North and North West. Small hospitals in the South West region had the lowest LOS whereas large hospitals in the North West had the highest. The highest proportion of bed days (11.2%) occurred in males aged less than 60 diagnosed with infectious diseases. Males aged less than 60 diagnosed with injuries and digestive diseases, and aged at least 60 diagnosed with COPD, and aged less than 60 diagnosed with infectious diseases, accounted for more than double those for female patients in the same disease groups. Both logistic regressions with LOS at least 1 week as the outcome and linear regression on appropriately log transformed LOS gave consistent results. Providing suitable palliative care or allowing patients to select the place for spending their final time of life, especially for patients with chronic diseases, which can reduce hospital resource utilization.

Keywords: Spending time in hospital, Hospital mortality, Number of bed days, Analysis of length of hospital stay

1. Introduction

Length of hospital stay (LOS) is a common parameter used to indicate health resource utilization, health care cost and severity of disease (Li, 1999; Wang et al., 2002; Lee et al., 2003). Based on a literature review Martin and Smith (1996) concluded that patient demographics and hospital characteristics were the two major factors that determine patient LOS. Among demographic characteristics, studies have reported that LOS varied according to age and disease group (Goldfarb et al., 1983; McMullan et al., 2004), whereas among hospital characteristics, LOS has been reported to vary by region, hospital size, and health care service (Health Technology Case Study 24, 1983; Xiao, 1997; Clarke, 2002).

LOS can be terminated by cure, transfer or death discharge. Many studies have been concerned only with LOS for patients with cured discharge (see, for example, Cabre et al., 2004). However, hospital stay terminated by death is also an important outcome event. Patient care in hospital is the most expensive way of providing palliative care (Huang et al., 2002). Longer stays are more likely to indicate physicians' decisions or administrative inefficiencies than patients' need (Brownell & Roos, 1995). Reducing LOS is health policy in many countries (Clarke & Rosen, 2001). Setting up a

proper palliative care or providing opportunity for patients to decide where they want to stay during the last stages of their life can reduce unnecessary hospital LOS. For inpatients spending their final days in hospital in developed countries such as Canada, England and Belgium, LOS has been studied by Huang et al. (2002), Dixon et al. (2004) and van den Block et al. (2007), but similar studies for developing countries are limited.

The development of LOS models for patients dying in hospital will be useful for hospital management, particularly for prioritizing health care policies and improving health services, including the most appropriate allocation of health resources according to differences in LOS with respect to patients' health conditions and demographic and geographic factors.

It is also important to develop appropriate methods for the statistical analysis of LOS data. Generally, LOS has a highly positive skewed distribution (Liu et al., 2001; Lee et al., 2003) and some patients stay in hospital for a very long time. Using linear regression to predict LOS is likely to seriously violate the assumptions of the model (Li, 1999). Various methods have been suggested to handle outliers. For caesarean delivery LOS Lee et al. (2003) used median regression. Marazzi et al. (1998) examined the adequacy of models based on lognormal, Weibull and gamma distributions and found that the lognormal model was most appropriate but cases with LOS less than one day were omitted from their analysis because of computational problems when the LOS is zero.

In this study, two methods were used to handle skewness in the LOS distribution: (a) logistic regression with LOS 7 days or more taken as the outcome, and (b) linear regression on the natural logarithm of LOS after adding an appropriate constant to cope with LOS equal to zero. We thus examined the variation in LOS with respect to principal diagnosis, demographic, and geographic and hospital size factors for patients dying in hospital in Southern Thailand over the 4-year period from 2000 to 2003. This analytical framework provides a straightforward methodology for modeling positively skewed outcomes with a high proportion of zero occurrences.

2. Material and methods

2.1 Data source and variables

The data for this study comprised 41,134 cases of mortality records routinely reported to the National Health Security Office (NHSO) of the Ministry of Public Health from hospitals in 14 provinces of Southern Thailand according to date of discharge during the 4 fiscal years from October 2000 to September 2003. We excluded 636 cases with no information on patient demographics or principal diagnosis. As gender and age by themselves are known to be insufficient to adequately model LOS (National Center for Health Statistics, 1973), we also used principal diagnosis group and hospital location and size. Two nominal explanatory factors were defined: (1) the combination of diagnosis, gender and age group (54 categories), and (2) the combination of region and hospital size (19 categories). Age and gender were grouped into six categories by dividing age into three groups: 0-59 years, 60-74 years and 75 and over. Principal diagnosis using ICD 10 was regrouped into nine broad classes: injuries, digestive diseases (DD), unspecified septicemia (ICD-10 code A41.9), other infectious diseases (ID), chronic obstructive pulmonary diseases (COPD), respiratory infection (RI), cardiovascular diseases (CVD), cancer, and other diseases. The 14 provinces were reduced to seven regions as follows: Chumphon and Ranong (N), Surat Thani (ST), Phangnga, Phuket and Krabi (NW), Nakhon Si Thammarat (NST), Satun and Trang (SW), Songkhla and Phattalung (CS) and Pattani, Yala and Narathiwat (SE). These seven regions were than combined with hospital size (60 or fewer beds, 61-499 beds, 500 beds or more) to give the second factor as shown in Table 1.

2.2 Statistical methods

Two statistical models were used to examine the influences of the two factors on length of stay for patients who died in hospital.

In the first model LOS was treated as a binary variable with patients staying in hospital at least 7 days as the outcome of interest. Logistic regression (Hosmer & Lemeshow, 2000; Kleinbaum & Klein, 2002) was then used to estimate the proportion p_{ij} of these outcomes in diagnosis-demographic group i and region-hospital size group j using the model

$$\ln \frac{p_{ij}}{1 - p_{ij}} = \mu + \alpha_i + \beta_j . \tag{1}$$

To avoid over-specification of the parameters, for each factor the category having the largest group size was taken as the referent with corresponding parameter 0. To calculate the adjusted prevalence $p_{i\bullet}$ for category i of the first factor, the term β_j in equation (1) was replaced by a constant β_0, chosen to make the sum of the expected number of outcomes equal to the sum of the observed number, that is,

$$\sum_{i=1}^{m} p_{i\bullet} n_i = \sum_{i=1}^{m} p_i n_i , \tag{2}$$

n_i being the sample size in category i. The constant β_0 was computed using an iterative procedure. Similarly, to calculate the adjusted prevalence $p_{\bullet j}$ for category j of the second factor, the term α_i in equation (1) was replaced by a constant α_0, again chosen to ensure that the sum of the expected number of outcomes equaled the total observed.

In the second model LOS was treated as a continuous outcome, by taking its natural logarithm after adding a constant d to handle zero days stay, giving the transformed outcome y. The linear regression model is thus similar to equation (1), namely

$$y_{ij} = \mu + \alpha_i + \beta_j. \tag{3}$$

Estimates of LOS for different levels of the first factor after adjusting for the second factor were calculated by replacing β_j in equation (1) by a constant β_0 to give fitted values $y_{i\bullet}$ and then reversing the transformation to give $\exp(y_{i\bullet})-d$, with β_0 chosen to match the overall fitted mean LOS with the overall observed mean.

Confidence intervals for these parameters were obtained by using the standard errors obtained through fitting each model.

All statistical analysis was carried out using the R program (R Development Core Team, 2007).

3. Results

Median LOS and proportions staying seven days or more are given in Tables 2 and 3 for each determinant. The largest number of patients (8,461) occurred in the cardiovascular disease group. There were 6,261 and 5,698 patients in the other infectious disease and injuries groups, respectively. The number of male patients in the digestive disease, COPD and respiratory infectious disease groups were close to double the number of female patients in the same groups. The shortest median LOS were found in the injuries group among males aged less than 75 and females aged less than 60 where 56 percent had LOS less than 7 days. The highest median LOS was found in the cancer group among females aged 75 and over. More than half of the patients with cancer had LOS at least 7 days. The North West region had the highest proportion of patients staying 7 days or more, followed by Nakhon Si Thammarat and the Central South. The highest median LOS occurred in the North West region with more than one third of LOS at least one week.

Figure 1 shows a histogram of the overall distribution of LOS after transforming the data by adding $d = 0.4$ and taking natural logarithms. The fit to this log-normal distribution is good for LOS greater than 1 day (apart from three high outliers at 1,179, 1,830 and 2,741 days). However, the lognormal distribution, being a continuous curve, cannot accurately accommodate data where LOS takes discrete values 0 or 1, although if these shorter lengths of stay were coded in hours rather than days they might well be accommodated in the model curve. Note that if the model is to be used to provide estimates of hospital costs, it is more important to accurately model longer stays.

The data were grouped by principal diagnosis, age, gender, region and hospital size from 40,498 individual cases into 1,026 records. Figure 2 shows a scatter plot of these observed counts and fitted values of LOS at least one week in the left panel and a residual plot in the right panel. The model predicts the proportions in the 1026 cells quite well, as shown in the residuals plot. However, the model could not be expected to accurately predict individual LOS.

Figure 3 shows the fitted prevalence of LOS of at least 1 week based on the logistic regression model. The dotted horizontal line on each graph gives the overall prevalence of patients having LOS at least 7 days, (30.9 percent). The top panel shows the prevalence of LOS at least 1 week by diagnosis-demographic groups after adjusting for the geographic-hospital size factor. Males aged less than 60 years admitted to hospital from injury comprised the referent group. The increasing trend of LOS at least one week in older age groups occurred in all disease groups except females diagnosed with digestive disease. The lowest LOS occurred among males aged less than 60 diagnosed with injuries while the highest LOS occurred among females diagnosed with cancer aged 75 and over.

The graph in the lower panel of Figure 3 shows the prevalence of LOS at least 1 week by the geographic-hospital size factor after adjusting for the diagnosis-demographic factor. The large hospital in Nakhon Si Thammarat was taken as the referent group. Patients from all hospital sizes in the North, Surat Thani, South East, and from small and medium sized hospitals in Nakhon Si Thammarat and small hospitals in the Central South were less likely to have LOS at least 1 week. Patients from large hospitals in the Northwest were more likely to have LOS at least 1 week. The increasing trends of LOS with hospital size appeared in all regions except for the North and the North West.

The second method involved fitting a linear model to the log-transformed outcome ln(0.4+LOS). Figure 4 shows a normal quartile plot for the standardized residuals from this model, and indicates that the normality assumption is reasonable for LOS greater than zero. However, the model gave an r-squared of only 7.8%, confirming that individual LOS cannot be accurately predicted. The components of this r-squared were 1.7% for age and gender alone and 6.6% for age-gender and principal diagnosis combined.

Figure 5 shows estimates of LOS from the model. The dotted line on each graph gives the overall mean of LOS (8.9

days). The top graph shows LOS by the diagnosis-demographic factor after adjusting for the geographic-hospital size factor. It shows a similar pattern of increasing trend of LOS by disease group, gender and age group as Figure 3. Patients with injuries had the lowest LOS while patients with cancer had the highest LOS.

The graph in the lower panel of Figure 5 shows mean LOS for the geographic-hospital size factor after adjusting for the demographic-diagnosis factor. Again the patterns are similar to those given by the logistic model.

Table 4 shows numbers of patients and corresponding expected proportions of bed days given by the model for each category of the diagnosis-demographic factor, thus giving a breakdown of expected contributions to the costs associated with patients dying in hospital in southern Thailand. Among patients aged less than 60, the highest proportion (11.2%) occurred in males diagnosed with infectious diseases while patients aged 60 to 74, males diagnosed with cancer, patients diagnosed with the CVD, and males diagnosed with COPD, respectively, had the highest proportions. The highest proportions for patients aged more than 74 years were found in the CVD group. Males diagnosed with injuries accounted for over 4% of all bed days. Males aged less than 60 diagnosed with injuries and digestive diseases, and aged at least 60 diagnosed with COPD, and aged less than 60 diagnosed with infectious diseases accounted for more than double those for female patients in the same disease groups.

4. Discussion

This study analysed LOS using logistic and linear regression. The first model was used for predicting the factors associated with LOS at least 1 week. This method provided reasonable and explainable results and the patterns of LOS across diagnosis-demographic and region-hospital size were clearly illustrated from the plot of the prevalence (Figure 3). The linear model with log transforming LOS provided results consistent with the logistic model (Figure 5). We used the log-linear model to calculate average LOS and total number of bed days as they cannot be calculated from the logistic model.

Spending the remaining time to death in the hospital increases hospital cost (Sashamani & Gray, 2004). In this study, LOS increased with age. Similar results were reported by Himsworth and Goldacre (1999), Brownell and Roos (1995) and McMullan (2004). The simple explanation is that older patients tend to take longer to recover from disease, and most have chronic disease whereas younger patents tend to have acute disease with shorter duration. However, our result contrasts with the finding by Dixon et al. (2004) that average LOS did not increase with age. The longest LOS was found in females aged at least 75 diagnosed with cancer. Patients with injuries as principal diagnosis had shortest average LOS. Reducing LOS is a policy for many health care systems (Clarke & Rosen, 2001). Brownell and Roos (1995) indicated that longer stays in hospital were more likely to reflect physicians' decisions or administrative inefficiencies than patient needs, and LOS among patients dying from injuries could not be managed due to the short stay and uncertainty of the discharge status. However, hospital LOS among cancer patients can be reduced by providing palliative care or providing the opportunity for patients to decide where they want to stay during their last stage of life.

LOS variation by region and hospital size has been reported in many studies (List et al., 1983; Xiao, 1997; Clarke, 2002). In our setting, small hospitals in the South West region had the lowest average LOS whereas large hospitals in the North West had the highest average LOS.

Beside average LOS, total numbers of bed days provide useful information on the financial cost of patients who die in hospital. Thus the use of health care services and resources can be identified by considering this aspect, allowing health planning and budget allocations to be fairly distributed to the most suitable groups. In the United States, hospital inpatient stays accounted for nearly one-third of total health care expenses (Machlin & Carper, 2004). Our results show that males aged less than 60 diagnosed with infectious diseases accounted for 11.2% of the total, whereas females in the same group accounted for only 5.4% of the total. The higher proportion in males is possibly due to excess deaths for males from HIV/AIDS (Lim & Choonpradub, 2007). Excess male mortality from other diseases related to HIV/AIDS such as tuberculosis and pneumonia were also reported by Rumakom et al. (2002). Males who died from injuries aged less than 60 accounted for 4% of total bed days, possibly due to traffic accidents with higher mortality in males than females (Bureau of Policy and Strategy, 2002).

A strength of this study is that data were taken for 4 years and from every hospital size in every province in Southern Thailand. Gustafson (1968) mentioned that the poor performance of linear regression analysis for LOS due to the small sample size. In our study, the sample size was large enough for accurate analysis with each statistical method. Extremely positive skewed data and zero days stay were not omitted from the analysis. Outliers can give vital information for health services management in several aspects such as resources allocation or assessment of health intervention (Marazzi, 1998). While information concerning physician characteristics for data used in this study was limited, this may not have unduly biased the results of our study. Westert et al. (1993) noted that the difference in LOS between treating physicians within the same hospital was much less than the difference between hospitals.

In conclusion, providing suitable palliative care or allowing patients to select the place for spending their final time of life, especially for patients with chronic diseases, which can reduce hospital resource utilization.

References

Brownell, M.D. & Roos, N.P. (1995). Variation in length of stay as a measure of efficiency in Manitoba hospitals. *Canadian Medical Association Journal*, 152, 675-682.

Bureau of Policy and Strategy. (2002). Thailand Health Profile 1999-2000. Ministry of Public Health: Express Transportation Organization. pp. 189-247.

Cabre, M., Bolivar, I., Pera, G. & Pallares, R. (2004). Factors influencing length of hospital stay in community-acquired pneumonia: a study in 27 community hospitals. *Epidemiology and Infection*, 132, 821-829.

Clarke, A. & Rosen, R. (2001). Length of stay, how short should hospital care be ?. *European Journal of Public Health*, 11, 166-170.

Clarke, A. (2002). Length of in-hospital stay and its relationship to quality of care. *Quality Safety Health Care*, 11, 209-210.

Dixon, T., Shaw, M., Frankel, S. & Ebrahim, S. (2004). Hospital admissions, age, and death: retrospective cohort study. *British Medical Journal*, 328, 1288-1290.

Goldfarb, M.G., Hornbrook, M.C. & Higgins, C.S. (1983). Determinants of hospital use: A cross-diagnostic analysis. *Medical Care*, 211, 48-66.

Gustafson, D.H. (1968). Length of stay: prediction and explanation. *Health Services Research*, 3, 12-34.

Health Technology case study 24. (1983). Variations in Hospital Length of Stay: Their relationship to health outcomes. Washington D.C.: U.S. Government Printing Office. pp 9-22.

Himsworth, R.L. & Goldacre, M.J. (1999). Does time spent in hospital in the final 15 years of life increase with age at death? A population based study. *British Medical Journal*, 319, 1338-1339.

Hosmer, D.W. & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed). New York: John Wiley & Sons, (Chapter 1) .

Huang, J., Boyd, C., Tyldesley, S., Zhang-Salomons, J., Groome, P.A. & Mackillop, W.J. (2002). Time spent in hospital in the last six months of life in patients who died of cancer in Ontario. *Journal of Clinical Oncology*, 20, 1584-1592.

Kleinbaum, D.G. & Klein, M. (2002). *Logistic regression: a self-learning text.* (2nd ed.). New York: Springer-Verlag, (Chapter 1).

Lee, A.H., Fung, W.K. & Fu, B. (2003). Analyzing hospital length of stay: mean or median regression? *Medical Care*, 41,681-686.

Li, J. (1999). An application of lifetime models in estimation of expected length of stay of patients in hospital with complexity and age adjustment. *Statistics in Medicine*, 18, 3337-3344.

Lim, A. & Choonpradub, C. (2007). A statistical method for forecasting demographic time series counts, with application to HIV/AIDS and other infectious disease mortality in southern Thailand. *Southeast Asian Journal of Tropical Medicine Public Health*, 38, 1029-1039.

List, N.D., Fronczak, N.E., Gottlieb, S.H. & Baker, R.E. (1983). A cross-national study of differences in length of stay of patients with cardiac diagnoses. *Medical Care*, 21, 519-530.

Liu, Y., Phillips, M. & Codde, J. (2001). Factors influencing patients' length of stay. *Australian Health Review*, 24, 63-70.

Machlin, S.R. & Carper, K. (2004). Expenses for Inpatient Hospital Stays. Statistical Brief #164, March 2007. Agency for Healthcare Research and Quality, Rockville. [Online] Available: http://www.meps.ahrq.gov/mepsweb/data_files/publications/st164/stat164.pdf (July 26, 2008).

Marazzi, A., Paccaud, F., Ruffieux, C. & Beguin, C. (1998). Fitting the distributions of length of stay by parametric models. *Medical Care*, 36, 915-927.

Martin, S. & Smith, P. (1996). Explaining variations in inpatient length of stay in the National Health Service. *Journal of Health Economics*, 15, 279-304.

McMullan, R., Silke, B., Bennett, K. & Callachand, S. (2004). Resource utilisation, length of hospital stay, and pattern

of investigation during acute medical hospital admission. *Postgraduate Medical Journal*, 80, 23-26.

National Center for Health Statistics. (1973). Average length of Stay in Short-Stay Hospitals: Demographic Factors, United States-1968. Vital and Health Statistics, Series 13-No 13. Washington D.C.: U.S. Government Printing Office. pp. 1-40.

R Development Core Team. (2007). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Rumakom, P., Prasartkul, P. & Punpuing, S. (2002). Change to the epidemiological transition in Thailand due to HIV/AIDS: Implications for population and health policies. The 2002 IUSSP Regional Population Conference, Bangkok, 10-12 June 2002. Institute for Population and Social Research, Nakhon Pathom: Mahidol University. pp. 1-12.

Sashamani, M. & Gray, A. (2004). Time to death and health expenditure: an improved model for the impact of demographic change on health care costs. *Age and Ageing*, 33, 556-561.

Van den Block, L., Deschepper, R., Drieskens, K., Bauwens, S., Bilsen, J., Bossuyt, N. & Deliens, L. (2007). Hospitalisations at the end of life: using a sentinel surveillance network to study hospital use and associated patient, disease and healthcare factors. *BMC Health Services Research*, 7, 69.

Wang, K., Yau, K.K.W. & Lee, A.H. (2002). A hierarchical Poisson mixture regression model to analyse maternity length of hospital stay. *Statistics Medicine*, 21, 3639-3654.

Westert, G.P., Nieboer, A.P. & Groenewegen, P.P. (1993). Variation in duration of hospital stay between hospitals and between doctors within hospitals. *Social Science & Medicine*, 37, 833-839.

Xiao, J., Douglas, D., Lee, A.H. & Vemuri, S.R. (1997). A delphi evaluation of the factors influencing length of stay in Australian hospitals. *International Journal Health Planning Management*, 12, 207-218.

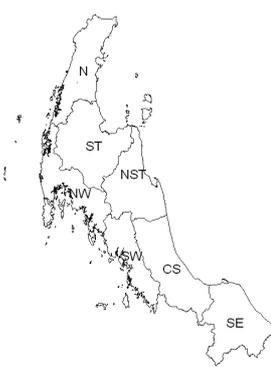Table 1. Classification of hospitals by region and size in southern Thailand

	Region	Number of beds			Total
		≤ 60	61 - 499	≥ 500	
	North (N)	14	2	1	17
	Surat Thani (ST)	17	5	2	24
	North West (NW)	16	4	1	21
	Nakhon Si Thammarat (NST)	14	5	1	20
	South West (SW)	16	2	-	18
	Central South (CS)	26	5	2	33
	South East (SE)	27	7	-	34
	Total	130	30	7	167

Table 2. LOS summary for patients dying in hospital by disease group, gender and age group

Disease group	Gender	Total	Median LOS			Number of LOS $\geq$ 7 days (%)		
			Age group			*Age group*		
			0-59	60-74	75+	0-59	60-74	75+
Injuries	Male	4,130	1	1	3	476 (13.9)	120 (25.5)	79 (32.4)
	Female	1,568	1	2	3	170 (15.5)	65 (25.6)	75 (34.1)
CVD	Male	4,581	1	2	3	283 (16.7)	434 (24.8)	316 (27.8)
	Female	3,880	2	2	3	234 (21.8)	356 (25.2)	427 (30.7)
Septicemia	Male	1,485	2	2	3	149 (19.6)	125 (30.1)	101 (32.6)
	Female	1,545	2	3	3.5	122 (19.1)	131 (27.1)	139 (32.9)
Other ID	Male	4,339	3	4	4	1,274 (33.0)	117 (37.5)	59 (35.5)
	Female	1,922	4	4	5	555 (33.8)	61 (36.1)	43 (39.4)
DD	Male	1,482	2	3	4	197 (23.8)	137 (34.7)	98 (37.5)
	Female	704	4	5	4	102 (38.8)	99 (41.1)	76 (38.0)
COPD	Male	1,886	2	5	5	125 (26.3)	332 (43.2)	275 (42.8)
	Female	835	2	4	5	82 (27.2)	111 (40.5)	111 (42.7)
RI	Male	1,408	2	4	7.5	161 (29.4)	168 (41.1)	232 (51.3)
	Female	967	3	5	6	123 (34.3)	109 (46.4)	182 (48.8)
Cancer	Male	1,917	5	7	7	358 (44.3)	397 (51.5)	174 (51.6)
	Female	1,246	7	7	9	322 (52.5)	228 (54.9)	139 (63.8)
Other	Male	3,393	2	5	5	709 (28.8)	243 (42.3)	161 (45.1)
	Female	3,210	2	5	7	663 (30.3)	247 (43.0)	236 (52.6)

Table 3. LOS summary for patients dying in hospital by region and hospital size group

Region	Total	Median LOS			Number of LOS $\geq$ 7 days (%)		
		Number of beds			*Number of beds*		
		$\leq$ 60	61-499	500+	$\leq$ 60	61-499	500+
North	4,051	3	3	3	29 (23.8)	321 (31.2)	881 (30.4)
Surat Thani	5,512	2	2	3	101 (26.0)	108 (20.8)	1,323 (28.7)
North West	5,591	4	3	4	79 (41.8)	956 (36.1)	1,054 (38.2)
Nakhon Si Thammarat	7,151	3	3	3	125 (27.6)	249 (31.8)	1,933 (32.7)
South West	4,636	2	2	-	55 (18.1)	1,040 (27.6)	-
Central South	6,868	3	3	3	148 (26.4)	1,134 (32.5)	899 (31.8)
South East	7,249	2	3	-	130 (19.7)	1,943 (29.5)	-

Table 4. Number of patients and bed days by disease group, gender and age group

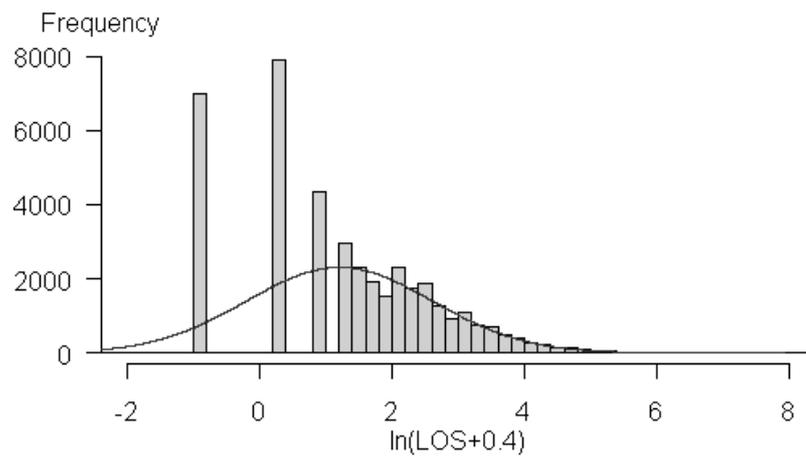| Disease group	Gender	Number of patients			Percent of bed days		
		Age group			*Age group*		
		0-59	60-74	75+	0-59	60-74	75+
Injuries	Male	3,415	471	244	4.1	0.7	0.5
	Female	1,094	254	220	1.4	0.5	0.5
CVD	Male	1,692	1,753	1,136	2.3	3.2	2.4
	Female	1,073	1,414	1,393	1.8	2.8	3.3
Septicemia	Male	760	415	310	1.2	0.9	0.7
	Female	640	483	422	1.0	1.0	1.1
Other ID	Male	3,861	312	166	10.0	0.9	0.5
	Female	1,644	169	109	4.4	0.5	0.3
DD	Male	826	395	261	1.5	1.0	0.8
	Female	263	241	200	0.7	0.8	0.7
COPD	Male	475	768	643	0.9	2.7	2.4
	Female	301	274	260	0.6	0.9	1.0
RI	Male	547	409	452	1.2	1.3	2.1
	Female	359	235	373	0.9	1.0	1.7
Cancer	Male	809	771	337	3.0	3.6	1.7
	Female	613	415	218	2.9	2.2	1.4
Other	Male	2,462	574	357	5.1	2.1	1.3
	Female	2,187	574	449	4.5	2.1	2.1

Figure 1. Distribution of ln(LOS+0.4) for in-hospital deaths in southern Thailand 2000-2003
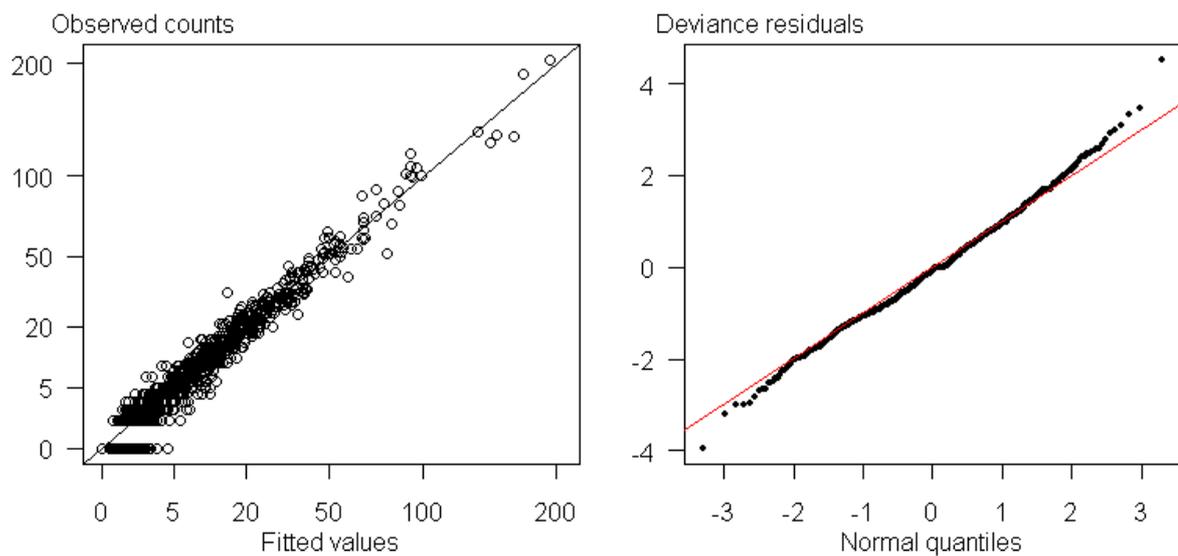

Figure 2. Scatter plot of observed counts and fitted values and residuals plot
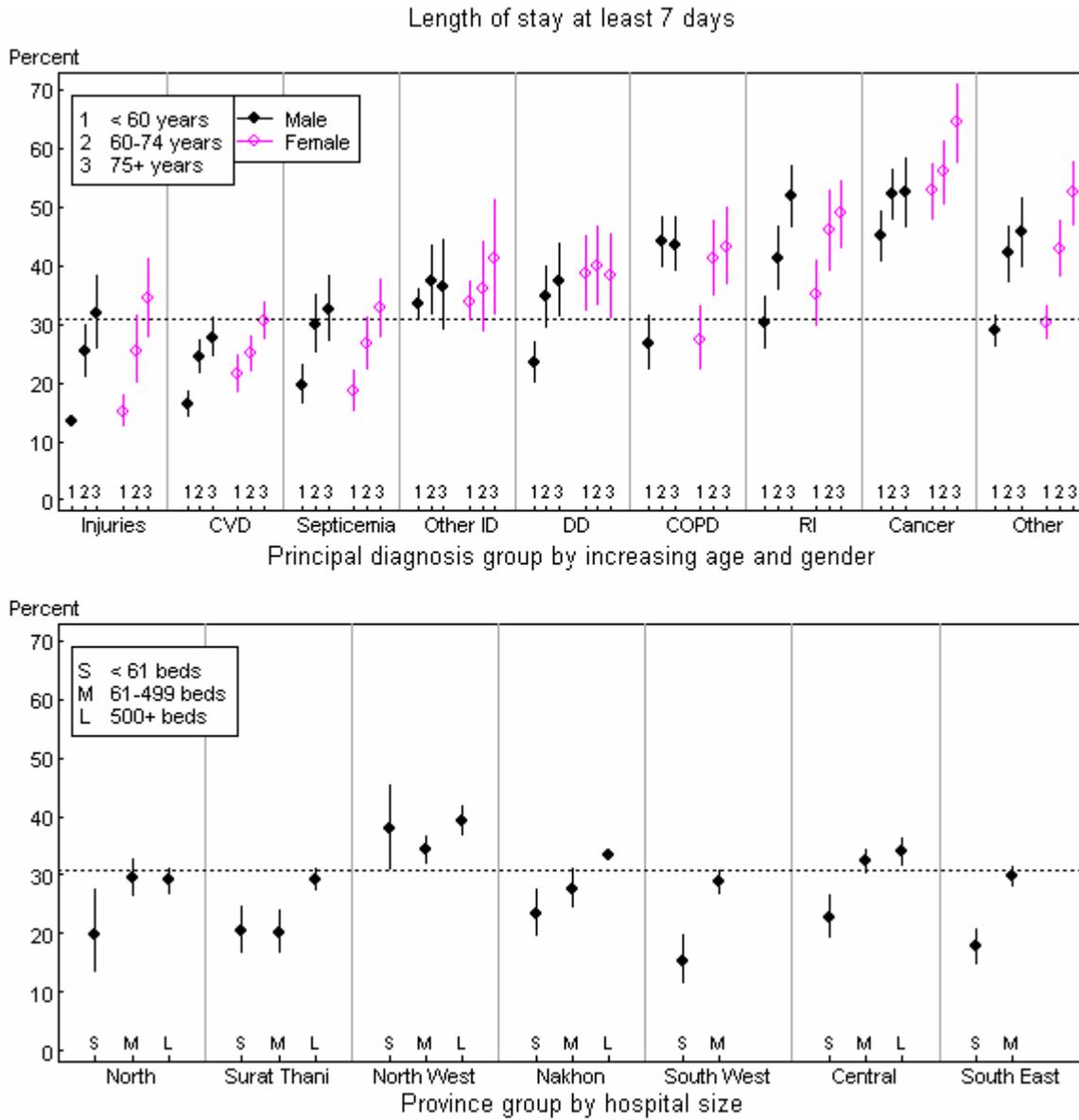
Figure 3. Prevalence of LOS at least 7 days adjusted by demographic and geographic factors.
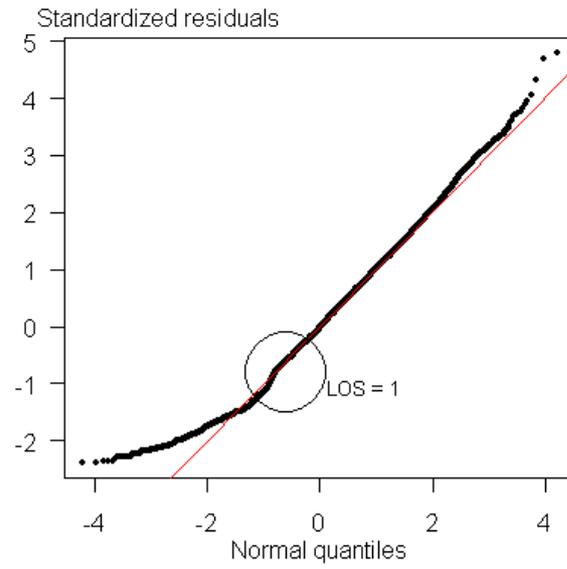
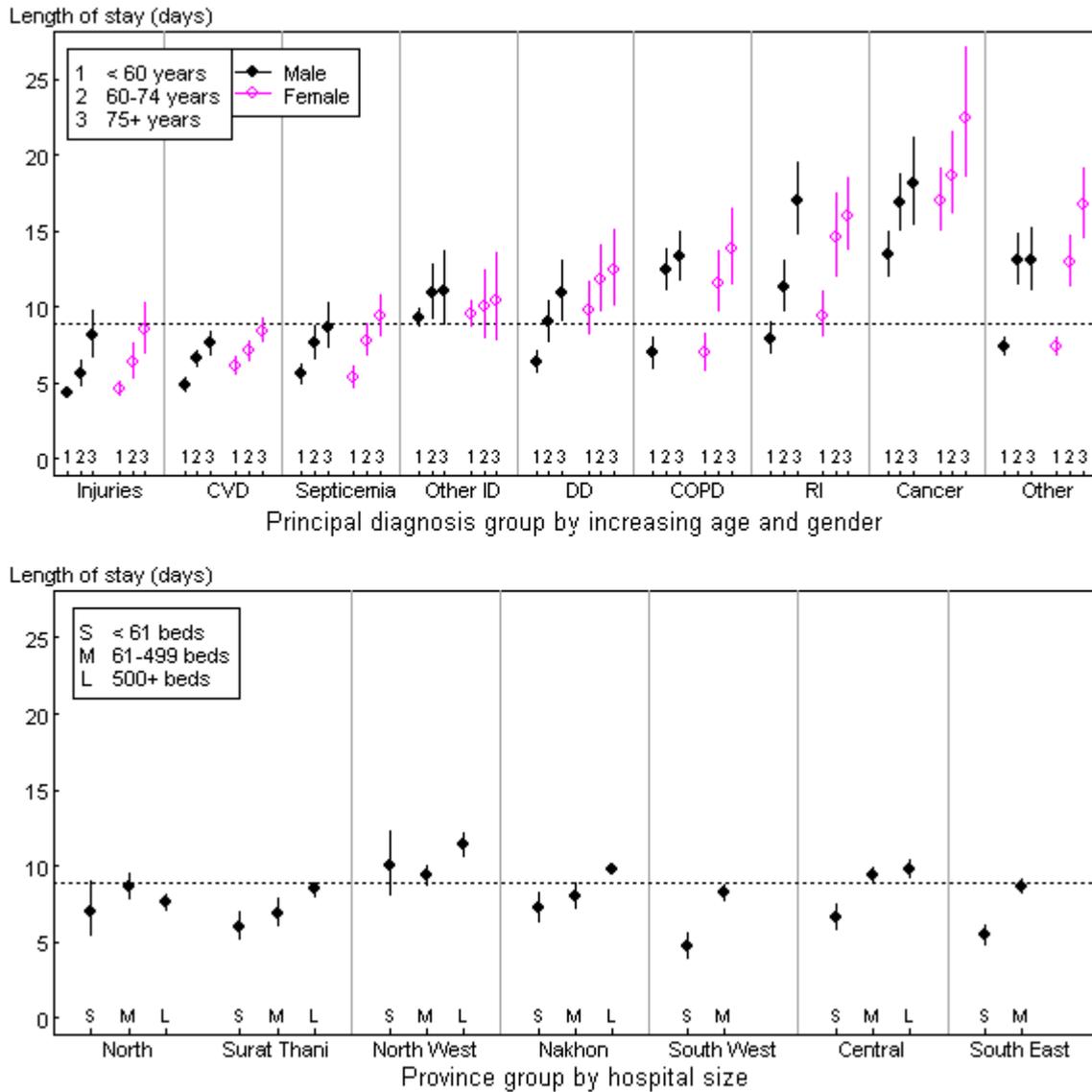Figure 4. Standardized residuals against normal quantiles plot

Figure 5. Confidence interval plots for LOS as continuous variable for patient who died in hospital adjusted for demographic and geographic factors.