Chapter 3

Study Results

In this chapter we present our findings from the two studies. The first study is concerned with the statistical modeling of Length of stay of patients dying in central region hospitals in Thailand and was published in the *J Pub. Health Dev. 2011; 9(2): 168-77.*

The second study is concerned with modeling the Muslim Victims of Terrorism Violence in Southern Thailand that was submitted to the international journal.

3.1 The first study

Multiple linear regression models were used to describe the variation in length of stay (LOS) of patients who died in hospital. To transform LOS to ln(LOS+1) as the outcome and found that has a highly statistically significant with five factors of interest. This finding was presented as the proceeding to the international conference on crisis in health and turning point in August 17-19, 2009, Phuket, Thailand. From that forum gave useful suggestions to modify this paper by using multiple logistic regression models constructed to describe the variation in LOS of patients who died in hospital. Copy of the published papers is also included.

The first paper

Length of stay of patients dying in central region hospitals in Thailand

Watcharin Komolmalai, M.Sc.
  Sirindhorn College of Public Health,
  Suphanburi Province
Metta Kuning, Ph.D.,
  Faculty of Science and Technology,
  Prince of Songlka University,
  Pattani Campus

ABSTRACT

Komolmalai W, Kuning M. Length of stay of patients dying in central region hospitals in Thailand. J Pub. Health Dev. 2011; 9(2): 168-77.

The statistical model was constructed to describe the variation in length of stay (LOS) of patients who died in hospital. The total of 257,076 records of mortality cases were included to determine the effect of factors on LOS for the period of October 1999 to September 2007, reported from provincial hospitals in Central Thailand to the Ministry of Public Health.

We considered principal diagnosis (ICD-10), age, gender, hospital size, and the geographic region as predictors for LOS. Chi-squared test was to investigate associations between the factors and LOS. All predictors were found to have a statistically significant association with LOS (p-value < 0.05). Multiple logistic regressions with weighted sum contrasts were used to compare proportions for each factor after adjusting for categorical covariates. The results showed that the percentage of longer LOS (7 days or more) was 35.2 among all patients. LOS increased with age, and hospital size. Injury was the principal diagnosis most likely to have short LOS, whereas cancer patients had the highest LOS. In term of geographic region, Bangkok had the highest percentage of patients with longer LOS, whereas the Southeast of Bangkok had the lowest percentage.

Corresponding Author
  Metta Kuning, Ph.D.
  Faculty of Science and
  Technology, Prince of Songlka
  University, Pattani Campus
E-mail: kmetta@bunga.pn.psu.ac.th

ระยะเวลาที่พักรักษาของผู้ป่วยที่เสียชีวิต
ในโรงพยาบาลเขตภาคกลางในประเทศไทย

บทคัดย่อ

วัชรินทร์ โกมลมาลัย เมตตา ภูนิง. ระยะเวลาที่พักรักษาของผู้ป่วยที่เสียชีวิต ในโรงพยาบาลเขตภาคกลางใน
ประเทศไทย. ว.สาธารณสุขและการพัฒนา, 2554; 9(2): 168-77.

การศึกษาระยะเวลาที่พักรักษาของผู้ป่วยที่เสียชีวิตในโรงพยาบาล โดยใช้สมการทางสถิติ
มาอธิบายถึงความแปรปรวนของระยะเวลาที่พักรักษา ของผู้ป่วยที่เสียชีวิตในโรงพยาบาล
จำนวน 257,076 ราย ที่ได้จัดส่งรายงานมายังกระทรวงสาธารณสุข จากโรงพยาบาลต่าง ๆ
ในเขตภาคกลางของประเทศไทย ตั้งแต่เดือนตุลาคม 2542 ถึงเดือนกันยายน 2550 ซึ่งการ
ศึกษาในครั้งนี้ ได้พิจารณาโรคที่เป็นสาเหตุหลักตามระบบ ICD-10 ในการเข้ารับการรักษา
อายุ เพศ การจำแนกขนาดโรงพยาบาลตามจำนวนเตียงและเขตพื้นที่ที่เป็นที่ตั้งของโรง-
พยาบาล เป็นตัวแปรที่ใช้ในการอธิบายระยะเวลาที่พักรักษาในโรงพยาบาล โดยใช้การทดสอบ
ไคสแควร์ (Chi-square test) และการวิเคราะห์ถดถอยโลจิสติกพหุคูณ (Multiple logistic
regression) ทดสอบหาความสัมพันธ์ระหว่างตัวแปรอธิบายกับระยะเวลาที่พักรักษาใน
โรงพยาบาล

การศึกษาพบว่าตัวแปรที่ศึกษาสามารถอธิบายระยะเวลาที่พักรักษาในโรงพยาบาลได้อย่าง
มีระดับนัยสำคัญทางสถิติ (p-value < 0.05) มีผู้ป่วยที่เสียชีวิตที่ใช้ระยะเวลาพักรักษานานกว่า
หนึ่งสัปดาห์โดยเฉลี่ยร้อยละ 35.2 ของผู้ป่วยที่ศึกษาทั้งหมด สัดส่วนของการพักรักษานาน
กว่าหนึ่งสัปดาห์จะเพิ่มขึ้นตามอายุและขนาดของโรงพยาบาล เมื่อเปรียบเทียบด้วยสาเหตุ
การป่วยจากโรคต่างๆ พบว่ากลุ่มของอุบัติเหตุใช้ระยะเวลาที่พักรักษานานกว่าหนึ่งสัปดาห์
ในสัดส่วนที่ต่ำสุด ในขณะที่ผู้เสียชีวิตจากกลุ่มโรคมะเร็งใช้เวลาพักรักษานานกว่าหนึ่งสัปดาห์
ในสัดส่วนที่สูงสุด เมื่อเปรียบเทียบระยะเวลาของการพักรักษานานกว่าหนึ่งสัปดาห์ในเขต
พื้นที่ต่างๆ พบว่ากรุงเทพมหานครมีสัดส่วนสูงที่สุด ในขณะที่กลุ่มของจังหวัดทางตะวันออก
เฉียงใต้จากกรุงเทพมหานคร พบว่ามีสัดส่วนของการพักรักษามากกว่าหนึ่งสัปดาห์ต่ำที่สุด

คำสำคัญ   ระยะเวลาที่พักรักษาในโรงพยาบาล ผู้ป่วยที่เสียชีวิต การวิเคราะห์ถดถอย
โลจิสติกพหุคูณ

INTRODUCTION

Health expenditure increases substantially with length of stay (LOS). The time spent in hospital by patient differs according to disease, treatment facility, cost of treatment and discharge status. Hospital stay terminated by death is an important outcome event, with death usually reflecting the severity and burden of disease.

Hospital LOS has been the subject of investigation over the past few decades. Previous studies showed that a large number of factors influence LOS, mainly patient demographics such as age and socioeconomic status, and hospital characteristics [1-4].Some studies have reported that low socioeconomic status can be used to predict longer LOS (on average, longer by 5.9 days in a study by Stelianides et al.)[5]. However, LOS is influenced to varying degrees by various other factors, such as clinicians' style of practice, size of the hospital, hospital ownership, teaching status and other hospital characteristics[6-7]. A patient's LOS is also highly correlated with their injury or illness severity[8]. The LOS is high for many diseases including cancer, mental diseases, renal diseases and others[9-10].

Investigating LOS and factors influencing hospital stay is important as it is a measure of the hospital's efficiency and also health care utilization by ill and injured people. Lim & Tongkumchum[3] investigated LOS for patients who died in Southern Thailand during 2000 to 2003. Patients with cancer and elderly patients, particularly females, had higher LOS for all diagnoses. Huang et al.[11] determined factors associated with LOS for patients with suspected community-acquired pneumonia (CAP) who required hospitalization for treatment. Weight loss, functional impairment and heart, renal or neoplasm diseases were predictive of LOS greater than seven days. In many countries, reducing LOS is one of the most important health policies[12]. Reducing LOS will reduce the cost per patient episode, reduce the risk of patients being exposed to hospital acquired infections, reduce waiting times for treatment, and improve clinical outcomes.

The empirical distribution of LOS is positively skewed so LOS models have used logistic regression to avoid symmetry assumptions[3, 13]. Through a better understanding of the factors affecting longer LOS, it will be useful for health care services, treatment during hospitalization, and the allocation of resources on the basis of health care service at hospitals. In this study, we examined the variation in LOS of patients dying in hospital in the Thai central region. The patient's principal diagnosis and demographic status, and hospital size and geographic location were taken into account in the model.

METHODOLOGY

Study design and variables

The data include hospital discharge database information routinely reported to the National Health Security Office (NHSO) in the Ministry of Public Health during the 8 fiscal years from October 1999 to September 2007. We considered principal diagnosis according to International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10), age, gender, the hospital size and geographic region as predictors for LOS. The diseases were classified into 9 groups: injuries, cerebrovascular disease (CVD), digestive system disease, infectious disease, respiratory disease, genito-urinary disease (GUD), respiratory infection, malignant neoplasms (cancer), and other diseases. Age was divided into three groups: < 60 years, 60-74 years and 75 years or more. Hos-

pital size was classified into three groups (small: 60 or fewer beds, medium: 61-499 beds, and large: 500 beds or more). The 26 provinces of Central Thailand were grouped into 6 geographic regions compressing provinces as follows:

  i. North: Nontaburi, Pathumtani, Ayuthaya, Aungthong, Lopburi and Singburi
  ii. Northwest: Chainat, Kanjanaburi, Suphanburi and Nachornpatom
  iii. Centre: Bangkok
  iv. East: Saraburi, Chachengtrao, Prachinburi, Nachornnayok and Srakaeo
  v. Southeast: Samutprakarn, Chonburi, Rayong, Chuntaburi and Trad

  vi. Southwest: Rajburi, Samutsakhorn, Samut-songkhram, Pechaburi and Prajubkerekun.

Except for cancer, which had longer LOS, the median LOS in each disease group varied from 2-6 days. We classified hospital LOS into binary outcomes: "less than 7 days" and "7 days or more (adverse outcome)".

Statistical Methods

Preliminary statistical analysis involved examining the frequency distributions of the determinants and their univariate associations with the outcome using chi-squared tests. Logistic regression analyses[14-15] were performed to determine variables associated with LOS using the additive model[16]:

$$\ln\left[\frac{P}{1-P}\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 \qquad (1)$$

In this model β_0 is the intercept and the terms X_1, X_2, X_3, X_4 and X_5 are factors denoting gender, age group, disease, hospital size and geographic region, respectively. To avoid over-specification of parameters, each set of coefficients was constrained to have a mean equal to 0.

To calculate the proportion of LOS for each factor after adjusting for the effects of the other factors, equation (1) was used with the terms associated with the other factors replaced by constants, chosen to make the sum of the expected numbers of LOS as based on the model equal to the number observed[17]. The statistical model fitting and graphical displays group 95% confidence intervals were performed using R, version 2.11.1.[18] and using weighted sum contrasts[19].

RESULTS

Table 1 shows the median LOS of patients who died in each combination of disease group and age group. Principal diagnosis was classified into nine disease groups: injuries, cerebrovascular disease (CVD), digestive system disease, infectious disease, respiratory disease, genito-urinary disease (GUD), respiratory infection, cancer and other diseases which were found to have median LOS 2, 3, 4, 4, 5, 6, 6, 8 and 4 days, respectively. The largest number of patients (59,671) was in the infectious disease group. There were 54,514 and 30,824 cases in CVD and cancer groups, respectively. The shortest median LOS was found in the injuries group and the highest median LOS was found in the cancer group.

Table 1   LOS summary of inpatients dying in hospital in Central Thailand, by disease group, gender and age group (October 1999 to September 2007)

| Disease group | Age group | | | | | | | Median LOS | | | |
| | < 60 | | 60- | | 75+ | | | | | | Age group |
Gender	Male	Female	Male	Female	Male	Female	Total	< 60	60–74	75+	all
1. Injuries	14,441	4,105	1,850	939	831	862	23,028	2	3	6	2
2. CVD	11,677	7,681	10,217	9,840	5,998	9,101	54,514	2	3	4	3
3. Digestive	6,631	2,924	2,777	2,081	1,394	1,616	17,423	3	4	6	4
4. Infectious	24,280	13,842	5,876	6,396	3,655	5,622	59,671	4	4	5	4
5. Resp.Disease	2,452	1,484	3,524	1,355	3,385	1.588	13,788	3	5	6	5
6. GUD	2,010	1,960	1,494	2,062	901	1,430	9,857	4	6	7	6
7. Resp.Infection	3,651	2,131	3,198	2,067	3,438	3,169	17,654	4	7	8	6
8. Cancer	7,944	6,646	6,994	4,457	2,715	2,068	30,824	8	8	9	8
9. Other	10,541	8,612	3,023	3,617	1,738	2,786	30,317	4	6	7	4
Total	83,627	49,385	38,953	32,814	24,055	28,242	257,076				

Table 2 shows the association between determinants and LOS. All determinants were associated with LOS (p-value < 0.05). Patients in the elderly age group (75+ years of age) (42.38%), having cancer (52.73%), in a large hospital (with 500+ beds) (37.62%) and in a hospital in the Bangkok region (46.58%) had longer LOS.

Table 2   Association between determinants and LOS prior to death in hospitals in Central Thailand, October 1999 to September 2007

Characteristics	LOS				Chi-square p-value
	Below 7 days (n = 166,531)		7 days or more (n = 90,545)		
	n	(%)	n	(%)	
Gender					< 0.001*
Male	96,396	65.74	50,239	34.26	
Female	70,135	63.50	40,306	36.50	
Age Group					< 0.001*
Below 60 years	92,182	69.30	40,830	30.70	
60-74 years	44,218	61.61	27,549	38.39	
75 years or more	30,131	57.62	22,166	42.38	
Disease Group					< 0.001*
Injuries	18,208	79.07	4,820	20.93	
CVD	40,376	74.07	14,138	25.93	
Digestive system Disease	11,610	66.64	5,813	33.36	
Infectious Disease	39,583	66.34	20,088	33.66	
Respiratory Disease	8,293	60.15	5,495	39.85	
GUD	5,774	58.58	4,083	41.42	
Respiratory Infection	9,819	55.62	7,835	44.38	
Cancer	14,570	47.27	16,254	52.73	
Others	18,298	60.36	12,019	39.64	
Hospital size					< 0.001*
Small: Below 60 beds	11,613	66.98	5,725	33.02	
Medium: 61-500 beds	90,395	66.32	45,900	33.68	
Large: 501 beds or more	64,523	62.38	38,920	37.62	
Geographic Region					< 0.001*
Southwest	22,842	64.99	12,306	35.01	
Northwest	22,633	68.94	10,195	31.06	
North	33,436	67.42	16,158	32.58	
East	23,331	66.43	11,789	33.57	
Southeast	35,063	70.59	14,609	29.41	
Bangkok	29,226	53.42	25,488	46.58	

* Significant at p-value < 0.001

The proportion of an adverse outcome is of primary interest, so it is important to construct confidence intervals for comparing several proportions after adjusting for categorical covariates. Figure 1 shows 95% confidence intervals of the percentages of patients with LOS at least one week, using logistic regression model with weighted sum contrasts, calculated for each of the determinants. This graph shows the corresponding adjusted percentages and overall mean with respect to each of the determinants.

Gender, age group, disease group, hospital size and geographic region were associated with LOS at a statistically significant level (p-value < 0.05). Patients in age group 75 years or more had longer LOS in hospital. Cancer, respiratory infection, and GUD also had longer LOS. CVD and injury had the shortest LOS. Patient in a large hospital had longer LOS. The highest LOS occurred in hospitals in Bangkok province.
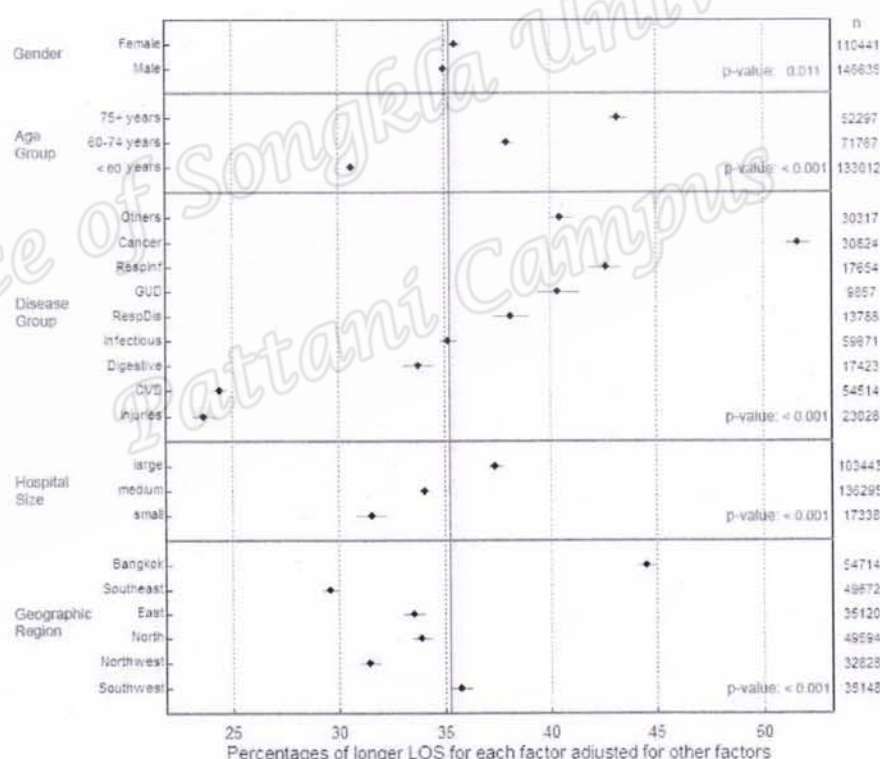

Figure 1   The percentages of longer LOS (7 days or more) prior to death in hospitals in Central Thailand, for various factors, at 95% confidence intervals of LOS

วารสารสาธารณสุขและการพัฒนา
ปีที่ 9 ฉบับที่ 2 พฤษภาคม-สิงหาคม 2554

DISCUSSION

This study investigated the variation in LOS for patients who died in Thai Central region hospitals between October 1999 and September 2007. The largest number of patients as three principal diagnoses had infectious disease, followed by cerebrovascular disease and cancer. Thirty five percent of all patients stayed in the hospital at least one week. Gender, age group, disease group, hospital size and geographic region were associated with LOS.

Over an eight-year period, longer LOS prior to death in hospital increased as age increased. This finding agrees with the studies by Himsworth & Goldacre[20], Brownell & Roos[21] and McMullan et al.[2] who also found that LOS increased with age. Elderly patients tend to get sick and take longer to recover from disease[22]. Most elderly are prone to developing a chronic disease whereas younger patents tend to have acute forms of disease with shorter duration[21].

Females had slightly longer LOS than males. For principal diagnoses, 'injury' had the shortest average LOS. One possible explanation is that patients tend to die earlier from some injuries. However, there is an opportunity for medical management and avoidance of a death outcome for those who survive from injuries. The longest LOS prior to death was for cases of cancer, respiratory infection and GUD, a result consistent with other studies[3,24]. A possible explanation is that these diseases require longer duration, whether the outcome is recovery or death.

LOS increased with hospital size. LOS by hospital size, for both recovery and death outcomes, has been reported in many studies[1,7] and found to vary greatly from place to place. In this study, Bangkok had the highest percentage of longer LOS, possibly because it has more than one large hospital, whereas the lowest percentage of longer LOS was found in the Southeast region of Central Thailand.

In general, LOS data contains outliers, extreme data and skewness, so the lognormal distribution, being a continuous curve, cannot accurately accommodate data where LOS takes discrete values 0 or 1, although if these shorter LOS were coded in hours rather days they might well be accommodated in the model curve[3]. However, we found that using a multiple logistic regression model to describe the factors associated with LOS of at least one week provided results consistent with a log-linear regression model. Logistic regression has an advantage of no assumptions related to the distribution, a linear relationship, or equal variances.

In conclusion, our findings highlight LOS prior to death in hospitals in the Central region of Thailand from 1999 to 2007 and factors associated with it. More research is needed to identify risk factors that are associated with LOS such as clinicians' style of practice, hospital ownership, teaching status, occupancy rate and others.

RECOMMENDATIONS

These findings could be useful for hospital management, particularly for prioritizing health care policies and improving health services, including when deciding the most appropriate allocation of health resources. It could be helpful to have more information about LOS with respect to patients' health conditions, and demographic and geographic factors. Further study is also needed to determine the relationship between regional LOS differences and health outcomes.

One limitation of the data needs to be acknowledged. Not all hospital deaths are reported by hospitals to the NHSO, and individual reporting practices differ from hospital to hospital and also over time. The data used in this study might be criticised for being incomplete but it does represent a very high proportion of all deaths in hospitals in the region studied.

REFERENCES

1. Clarke A. Length of in-hospital stay and its relationship to quality of care. Quality Safety Health Care, 2002; 11: 209-10.

2. McMullan R, Silke B, Bennett K, Callachand S. Resource utilization length of hospital stay and pattern of investigation during acute medical hospital admission. Postgraduate Medical Journal, 2004; 80(939): 23-6.

3. Lim A, Tongkumchum P. Methods for analyzing hospital length of stay with application to inpatients dying in Southern Thailand. Global Journal of Health Science, 2009; 1(1): 27-38.

4. McGregor MJ, Reid RJ, Schulzer M, Fitzgerald JM, Levy AR, Cox MB. Socioeconomic status and hospital utilization among younger adult pneumonia admissions at a Canadian hospital. BMC Health Services Research, 2006; 6: 152.

5. Stelianides S, Golmard JL, Carbon C, Fantin B. Influence of socioeconomic status on features and outcome of communityacquired pneumonia. European Journal of Clinical Microbiology and Infectious Diseases, 1999; 18: 704-8.

6. Menendez R, Cremades MJ, Moragon EM, Soler JJ, Reyes S, Perpina M. Duration of length of stay in pneumonia: influence of clinical factors and hospital type. European Respiratory Journal, 2003; 22: 643-8.

7. Xiao J, Douglas D, Lee AH, Vemuri SR. A delphi evaluation of the factors influencing length of stay in Australian hospitals. International Journal Health Planning Management, 1997; 12: 207-18.

8. Walczak S, Scorpio RJ, Pofahl WE. Predicting Hospital Length of Stay with Neural Networks. Proceedings of the Eleventh International FLAIRS Conference, 1998: 333-7.

9. Candrilli S, Bell T, Irish W, Morris E, Goldman S, Cairo MS. A comparison of inpatient length of stay and costs among patients with hematologic malignancies (excluding Hodgkin disease) associated with and without acute renal failure. Clinical Lymphoma and Myeloma, 2008; 8(1): 44–51.

10. Hampshire PA, Welch CA, McCrossan LA, Francis K, Harrison DA. Admission factors associated with hospital mortality in patients with haematological malignancy admitted to UK adult, general critical care units: A secondary analysis of the ICNARC Case Mix Programme Database. Critical Care, 2009; 13(R137): 1–28.

11. Huang JQ, Hooper PM, Marrie TJ. Factors associated with length of stay in hospital for suspected community-acquired pneumonia. Canadian Respiratory Journal, 2006; 13(6): 317-24.

12. Clarke A, Rosen R. Length of stay: How short should hospital care be. The European Journal of Public Health, 2001; 11(2): 166-70.

13. Ruttimann UE, Pollack MM. Variability in duration of stay in pediatric intensive care units: A multiinstitutional study. The Journal of Pediatrics, 1996; 128(1): 35-44.

14. Hosmer DW, Lemeshow S. Applied logistic regression. 2nd ed. New York : John Wiley and Sons; 2000.

15. Kleinbaum DG, Klein M. Logistic Regression: a Self-Learning Text. 2nd ed. New York : Springer-Verlag; 2002.

16. Chongsuvivatwong V. Analysis of Epidemiological Data Using R and Epicalc. Hat Yai, Thailand: Chammuang Press; 2008.

17. Kongchouy N, Sampantarak U. Confidence Intervals for Adjusted Proportions Using Logistic Regression. Modern Applied Science, 2010; 4(6):1-7.

18. The R Project for Statistical Computing. R version 2.11.1, 2010. Available at: http://www.r-project.org/.

19. Tongkumchum P, McNeil D. Confidence intervals using contrasts for regression model. Songklanakarin Journal of Science and Technology, 2009; 31(2): 151-6.

20. Himsworth RL, Goldacre MJ. Does time spent in hospital in the final 15 years Of life increase with age at death? A population based study. British Medical Journal, 1999; 319(7221): 1338-9.

21. Brownell MD, Roos NP. Variation in length of stay as a measure of efficiency in Manitoba hospitals. Canadian Medical Association Journal, 1995; 152: 675-82.

22. Lesourd B. Immune response during disease and recovery in the elderly. Proceedings of the Nutrition Society Cambridge Journals, 1999; 58: 85-98.

23. Cavanaugh JC, Blanchard-Fields F. Adult development and aging. 5th ed. Belmont (CA): Wadsworth Publishing/Thomson Learning. 2006.

24. Huang J, Boyd C, Tyldesley S, Zhang-Salomons J, Groome PA, Mackillop WJ. Time Spent in Hospital in the Last Six Months of Life in Patients Who Died of Cancer in Ontario. Journal of Clinical Oncology, 2002; 20(6): 1584-92.

3.2 The second study

We consider methods for modeling and displaying an adverse event incidence rates that vary over space and time with respect to demographic factors (gender and age), region and years). The GLMs were used to fit to the injury incidence rates from the terrorism violence to Muslim in the southern Thai region for 6 years (2004-2009). Since the outcomes are counts per person-year, two regression models are used to compare risks of violence to Muslim residents in target area:

- negative binomial (an extension of Poison model with a parameter that allows for over-dispersion), and

- log-transformed linear model modified to handle zero counts (0 to 0.5)

We create two complex categorical factors: (1) age group & sex (6 levels); (2) region & year (138 levels). Then we fit a model with two factors for Muslim victims of terrorism violence. Also we construct confidence intervals for comparing risks for each level of factor with overall mean, after adjusting for other factors.

With the model comparisons, each model gave different results, but each showed that while specific regions were at higher risk at different times and these patterns could not be easily predicted, risks in different demographic groups remained relatively constant.

The second paper

TITLE OF THE PAPER

Title: Muslim Victims of Terrorism Violence in Southern Thailand

Authors   Watcharin Komolmalai [1], M.Sc. (Research Methodology)

Metta Kuning [2], Ph.D. (Epidemiology)

Affiliation: [1] Sirindhorn College of Public Health, Suphanburi Province,

Suphanburi, 72000, Thailand

[2] Department of Mathematics and Computer Science, Faculty

of Science and Technology, Prince of Songkla University,

Pattani Campus, 94000,Thailand

Correspondence   Metta Kuning

Department of Mathematics and Computer Science,

Faculty of Science and Technology,

Prince of Songkla University, Pattani Campus, 94000, Thailand

Tel: +66 (8) 6747 8210, Fax: +66 (0) 7331 2179

E-mail address: kmetta@bunga.pn.psu.ac.th

Type of contribution: Original article

Number of tables: 1

Number of figures: 3

MUSLIM VICTIMS OF TERRORISM VIOLENCE

IN SOUTHERN THAILAND

Abstract. Terrorism violence is an important public health issue and in many aspects. It is similar to a deadly disease that can reach epidemic proportions. We investigated statistical models for describing the incidence rate of injuries to civilian resident victims of violence from terrorism in the target area defined as Pattani, Yala and Narathiwat provinces and four eastern districts of Songkhla province. From January 1, 2004 for six years, until December 31, 2009 there were 4,143 Muslim residents and 3,544 other (mainly Buddhist) residents of the target area have been recorded as victims by the Deep South Coordination Centre (DSCC). The overall incidence rates per 100,000 residents are 48 for Muslims and 121 for non-Muslims. We focused on the Muslim population and fitted negative binomial and log-normal models to incidence rates classified by gender, age group, region and year, with the objective of comparing relative risk by these factors, after adjusting for other factors to remove confounding. The models gave different results, but each showed that while specific regions were at higher risk at different times and these patterns could not be easily predicted, risks in different demographic groups remained relatively constant.

1. INTRODUCTION

A dramatic raid on an army base in Narathiwat on January 4, 2004 clearly signaled a renewed outbreak of terrorism violence in southern Thailand. Subsequently, the martial law has been declared in the three southern provinces, Pattani, Yala and Songkhla, on January 5, 2004 (Cline, 2007). Two further flashpoints were followed on April 28, 2004 when insurgents attacked 15 security posts with the army retaliating which culminated in a bloody siege of the historic Kru Se mosque in Pattani and killing over 100 people and a mass demonstration at Tak Bai, Narathiwat on October 25, 2004, where 78 unarmed protestors died, mainly from suffocation, after being locked up and spending more than five hours lying in the back of army trucks. These victims are mainly Muslim men signified a sharp deterioration in the security situation among Muslim residents in Pattani, Yala and Narathiwat provinces and four eastern districts of Songkhla province (Ward & Hackett, 2004; McCargo, 2009).

Several studies have been written about the facts surrounding and alleged causes of the situation that has developed in the seven years since the violence escalated. A recent report by the International Crisis Group (2010) gives a detailed history of the relevant background and events. Nakata (2010) recently devoted a complete issue to this subject. These papers reveal a wide range of scholarly views, but little if any serious analysis of data. To our knowledge, there are only three substantial scholarly analyses of the data that are publicly available. The first is a publication by Marohabout et al (2009) that fitted a statistical model to events classified by location and month (in 2004 and 2005), using data files provided in police reports in the terrorism target area (defined as the three provinces Pattani, Yala and Narathiwat and

the four easternmost districts of Songkhla province). The second is a conference paper presented by Khongmark et al (2011) that modeled injury incidence rates for non-Muslims in the target area for years 2004-2009 inclusive, using data recorded in the database of the Deep South Coordination Center (DSCC). The third is an unpublished report by Jitpiromsri (2010) that provides statistical graphs and summaries of the 9,446 terrorism incidents resulting in approximately 4,100 deaths and 6,500 non-fatal injuries, again using the Deep South Watch database for the 73 months from January 2004 to January 2010 inclusive.

It should be noted that these articles all defined events as occurrences on both sides of the conflict, that is, the victims included both civilians and non-civilians (defined as army and police personnel). Since such non-civilians could be regarded not just as victims but also as protagonists, we focus in this paper on Muslim civilian victims. Our objective is to provide a detailed analysis of these victim incidence rates, using appropriate statistical models that take into account the gender, age-group, location and year of the event.

2. MATERIALS AND METHODS

2.1 Data and variables

We considered incidence rates per 100,000 population for Muslim resident civilian victims of terrorism events classified by gender, age group (<25, 25-44 and 45 or more), district of residence and year (six years from 2004 to 2009 inclusive). The data provided by the Deep South Coordination Centre (DSCC) database, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, Thailand. The population denominators were obtained from the 2000 population and housing census of Thailand, National

Statistical Office. We restricted the study to Muslim victims because previous studies (in particular the 2010 report by the International Crisis Group and the 2010 conference paper by Khongmark) suggest that their patterns of violence by location and period are different from those of non-Muslim victims. The 37 districts were aggregated districts with Muslim population less than 30,000 into 23 larger regions as listed in Table 1.

Province	RegionID: Districts	Population	
		Muslim	Total
Songkla	1: Chana/Thepha	94,178	156,799
	2: SabaYoi/Na Thawi	48,271	110,507
Pattani	3: Mueang Pattani	67,149	104,145
	4: Kok Pho/Mae Lan	40,816	75,628
	5: Nong Chik/Mayo/Kapho	61,305	70,118
	6: Yaring	79,051	81,495
	7: Panare/Sai Buri/Mai Kaen	88,471	108,188
	8: ThungYang Dang	69,745	73,545
	9: Yarang	73,919	78,740
Yala	10: Mueang Yala	79,343	154,634
	11: Betong/Than To	31,487	68,193
	12: Raman	54,451	62,756
	13: Yaha/Kabang/Krong Pinang	50,522	56,546
	14: Bannang Sata	69,892	73,408
Narathiwat	15: Mueang Narathiwat	72,665	104,615
	16: Tak Bai	45,781	61,157
	17: Bacho/Yi-ngo	82,424	85,225
	18: Rueso	53,333	59,108
	19: Rangae	69,530	80,550
	20: SiSakon/Chanae	50,075	54,039
	21: Sukirin/Waeng	52,141	63,765
	22: Su-ngaiPadi/Cho-airong	75,688	89,251
	23: Su-ngaiKolok	41,317	64,640
Total Muslim Population		1,451,554	1,937,052

Table 1:   Regions used in analysis of Muslim victims of terrorism in southern Thailand

2.2 Statistical methods

Linear regression (see, for example, Cook & Weisberg, 1999) is a statistical method widely used to model the association between a continuous outcome and a set of fixed determinants. The model expresses the outcome variable as an additive function of the determinants. For example, if there are two categorical determinants with levels indexed by subscripts i and j, the model takes the form

$$Y_{ij} = \mu + \alpha_i + \beta_j. \tag{1}$$

In this case the number of parameters is $r + c - 1$ where r and c are the number of levels of the factors α and β, respectively, thus requiring two constraints, taken as $\Sigma\alpha_i = 0$ and $\Sigma\beta_j = 0$ so that μ encapsulates the average of Y. We also assume that the errors are independent and normally distributed with mean 0 and constant standard deviation. The model may be fitted to the observations y_{ij} by least squares, giving estimates and confidence intervals for the parameters. Equation (1) generalizes straightforwardly to any specified number of categorical determinants.

This method also applies to data that need to be transformed to satisfy the normality assumption, by first applying the method to the transformed data and then rescaling the result to ensure that the overall means of the untransformed data are the same before and after adjustment. It also extends straightforwardly to any number of covariate factors.

The Poisson generalized linear model is widely used for modeling event counts in incidence rates (see, for example, Crawley, 2005). For two additive factors as in the linear model given by equation (1), if P_{ij} is the population denominator, the expected value of the cell count N_{ij} is expressed as

$$E[N_{ij}] = P_{ij} \exp(\mu + \alpha_i + \beta_j) \tag{2}$$

However, the Poisson model often does not fit incidence data in practice because it assumes that the variance is equal to the mean, and in many situations the variance is substantially greater than the mean (see, for example, Jansakul & Hinde, 2004; Kaewsompak et al, 2005; Paul & Saha, 2007; Kongchouy & Sampantarak, 2010). The standard negative binomial GLM is a generalization of the Poisson model with the same mean λ, but the variance is $\lambda(1 + \lambda/\theta)$ where $\theta > 0$ (see, for example, Chapter 7 of Venables & Ripley, 2002). This over-dispersion is often the result of clustering (see, for example, Demidenko, 2007).

By analogy with the method used for means based on the linear regression model, we define the adjusted incidence rate for level j of factor β as $\exp(\hat{\beta}_j + c)$, where the constant c is chosen to ensure that the total number of adverse events based on the fitted model matches the number observed, that is,

$$\sum n_{ij} = \sum P_{ij} \exp(\hat{\beta}_j + c). \tag{3}$$

We used R software (R Development Core Team, 2010) to produce all statistical results and graphs.

2.3 Analysis strategy

To remove skewness in the linear model we transformed the incidence rates by taking their logarithms, after replacing zero counts by 0.5 to ensure finiteness.

We fitted models with two additive factors as determinants, one comprising the combination of gender and age-group, and the other comprising the combination of

region and calendar year. This model differs from that used by Khongmark (2010), who fitted an additive model comprising age-group, year, and the gender-region combination as three factors, for statistical reasons aimed at reducing the standard errors of the estimated parameters. However, we restricted to combining gender and age-group as a single demographic factor and define a single further factor combining region and year. This model was chosen because it is arguably more appropriate for studying patterns of conflict where highly mobile attackers can choose the time and place of their attack, and thus these times and place are largely unpredictable, rather than following the predictable patterns inherent in additive models.

Even though they are valid and often preferred models, the Poisson and negative binomial models also have problems with zero counts, failing to converge in situations where no events occur for a level of a factor (in contrast to linear models). This problem was circumvented by making a minimal change in the data by shifting a single non-zero count from one level to a neighboring level, thus keeping the total number of events constant.

3. RESULTS

During 2004 to 2009, the data covers 5,169 terrorist acts with civilian were the target, and the overall incidence rates per 100,000 residents are 48 for Muslims. There were no Muslim victims in the Chana/Thepa region in 2005, and as a result the Poisson and negative binomial models failed to converge, but convergence was achieved when the year of occurrence was changed to 2005 for one of the 4 male victims aged 25-44 who were injured in 2006. Residuals plots for the Poisson, negative binomial with dispersion parameter $\theta = 10$, and log-transformed normal models are shown in

Figure 1. In each case the high outlier corresponds to male residents aged 25-44 in

Tak Bai in 2004. The value $\theta = 10$ for the negative binomial dispersion parameter was

chosen because after removing this outlier the deviance dropped to a value for which

the chi-squared test was statistically significant, indicating a plausible fit of this
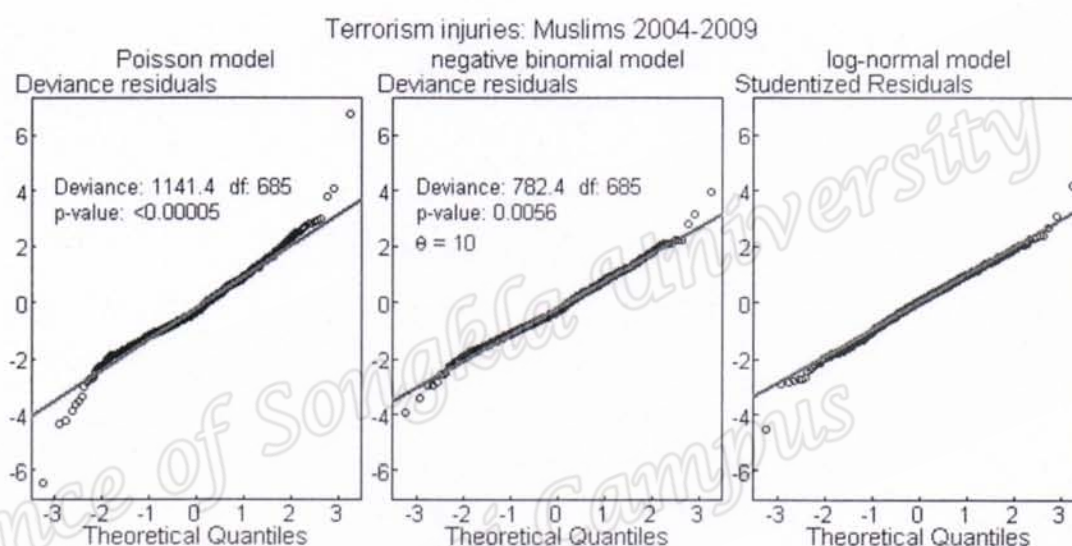
model.

Figure 1: Residuals plots against normal quantiles for three statistical models

fitted to injury rates of Muslim victims of terrorism in southern

Thailand: 2004-2009.

In Figure 2 incidence rates for each factor after adjusting for the other factor are

plotted for the negative binomial model, together with corresponding adjusted

incidence rates for the log-normal model. Estimates in gender-age groups are more

accurate (shorter confidence intervals) because each has a larger sample size (n=90).

Region-year risk estimates (n=6) are much less accurate. The regions have different

trend patterns. Some rose steadily apart from a dip in 2008 (Pattani City, Batong,

Bacho, Rueso and SungaiPadi/Cho-airong), whereas others rose and fell (Chana, Yala
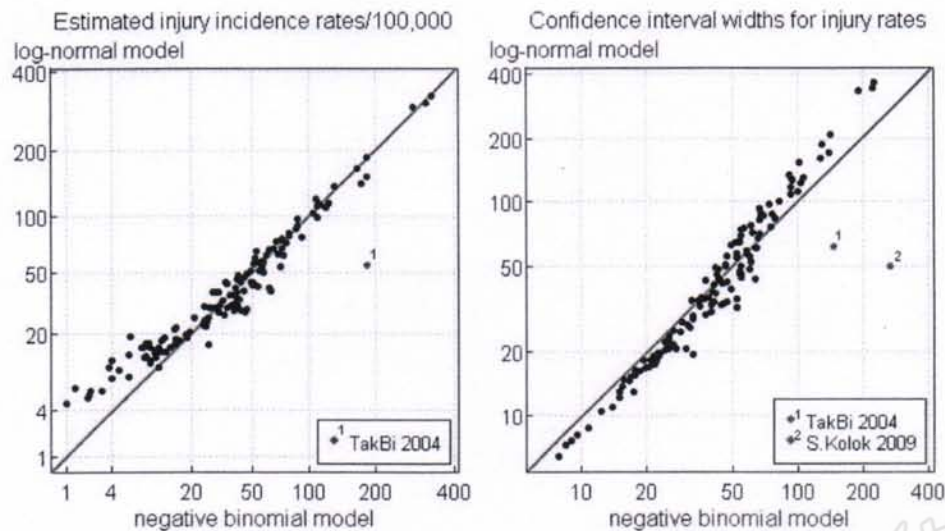
City, Raman, Yaha, BanangSta, Narathiwat City, and SungaiKolok), and rural

districts of Pattani rose up to the overall incidence rate and two (SabaYoi, Batong,

Tak Bai and Sukirin/Waeng) fell and rose and fell again.

Figure 2: Plots of incidence rates by gender-age group adjusted for region-year (left panel) and for region-year adjusted for gender-age group (right panel), with 95% confidence intervals, for the negative binomial model with dispersion parameter $\theta = 10$. The red-coloured points denote corresponding adjusted incidence rates based on the log-normal model.

In Figure 3 the results from the two models are compared with respect to estimated incidence rates and widths of confidence intervals. Estimated incidence rates (left panel) give an outlier from Tak Bai in 2004, whereas confidence interval widths (right panel) cover the outliers from Tak Bai in 2004 and SungaiKolok in 2009.

Figure 3: Comparison of incidence rates (left panel) and confidence interval widths (right panel) for log-normal and negative binomial models.

4. DISCUSSION

Terror-related injuries and deaths occurred in many countries such as Iraq, Afghanistan, Pakistan, India and Sri Lanka with emergence to southern Thailand. This paper investigates statistical models for describing the incidence rate of injuries to Muslims victims of violence from terrorism in Deep South (Pattani, Yala and Narathiwat) and four eastern districts of Songkhla province.

The generalized linear models (GLMs) were fitted to the injury incidence rates from the terrorism violence to Muslim victims in the southern Thailand region for 6 years (2004-2009). The negative binomial and transformed linear models fit equally well in overall. However, the log-normal model gives higher estimated incidence rates when incidence rates are low. Also, the log-normal gives higher standard errors when incidence rates are high. Both models fail to handle zero counts. Estimates of adjusted

incidence rates in both models need to be sealed to ensure that the overall mean incidence rate is the same before and after adjustment for covariates.

The annual risk of becoming a victim for Muslims (48 per 100,000) is less than half that for non-Muslims (121 per 100,000). The majority of victims were non-Muslim, but the percentage of Muslim deaths was higher than non-Muslim (Jitpiromsri, 2010). Muslims were targeted more by gunshot than others since those who died were more likely to be shot, while victims of bombings more likely to be non-Muslim (Chirtkiatsakul, 2011).

Male has higher risk than female, particularly with male at ages 25 or more, consistent with studies of Peleg et al. (2003; 2004), Sheffy et al. (2006). There were high risks attack occurred in the rural area than the city except for Yala City had the highest incidence rates. The risk with respect to region and year, in 2007 had the highest incidence rates in many region of each province such as Songkhla province in SabaYoi/NaTawi, Yala province in Yala City, Raman, Yaha and BannangSata, Narathiwat province in Rueso, Rangae and Sisakon. Except for Tak Bai, which had the greatest number from what happened there in October 2004, which sparked the ensuing violence, but after 2004 this region had relatively low annual incidence rates.

However, the ongoing insurgent has continued to escalate (Mavin, 2007; Wikipedia, 2011), with almost daily bombings, drive by shooting, arson and beheadings. The victims of insurgents include both Muslim (local) and non-Muslim (Buddhist), which effects on individuals, families and communities increased on emergency health care and mental health burden created by insurgency. There is urgent need of public health

approach to be expanded beyond treatment for individuals who are most severely affected to comprehensive prevention and health promotion.

5. ACKNOWLEDGEMENTS

6. REFERENCES

Chirtkiatsakul, B. (2011), Injuries from unrest situation in Deep South of Thailand, 2004-2010 (Dissertation). Pattani, Prince of Songkla University; 2011. 37p.

Cline, L. (2007), Thailand and the Insurgency in the South. Small wars and Insurgencies, 18(2), 275-287.

Cook, R.D., & Weisberg, S. (1999), Applied regression including computing and Graphics. Hoboken NJ: John Wiley, New York, USA.

Crawley, M.J. (2005), Statistics: an introduction using R. John Wiley & Sons. Chichester, England.

Demidenko, E. (2007), Poisson Regression for Clustered Data. International Statistical Review, 75(1), 96-113.

International Crisis Group (ICG) (2010), Stalemate in Southern Thailand, 3
    November 2010, Asia Briefing N 113, [Cited 2011 April 1]. Available
    from: URL: http://www.unhcr.org/refworld/ docid/4ceb84162.html

Jansakul, N., & Hinde, J.P. (2004), Linear mean-variance negative binomial models
    for analysis of orange tissue-culture data. Songklanakarin J Sci. Technol.,
    26, 683-96.

Jitpiromsri, S. (2011), Sixth Year of the Southern Fire: Dynamics of Insurgency and
    Formation of the New Imagined Violence. 10 March 2010, Deep South
    Watch. Center for the Study of Conflict and Cultural Diversity, Prince of
    Songkla University. [Cited 2011 April 1]. Available from: URL:
    http://www.deepsouthwatch.org/ node/730

Kaewsompak, S., Boonpradit, S., Choonpradub, C., & Chaisuksant, Y. (2005),
    Mapping acute febrile illness incidence in Yala province. Songklanakarin
    Medical Journal, 23(6), 455-62.

Khongmark, S., & Kuning, M. (2011), Modeling Incidence Rates of Terrorism
    Injuries in southern Thailand. Proceedings of the 2011 National Statistics
    and Applied Statistics Conference.

Kongchouy, N., & Sampantarak, U. (2007), Confidence Intervals for Adjusted
    Proportions Using Logistic Regression. Modern Applied Science, 4(6), 2-
    7.

Marohabout, P., Choonpradub, C., & Kuning, M. (2009), Terrorism Risk Modeling in
    Southern Border Provinces of Thailand during 2004 to 2005.
    Songklanakarin J. of Social Sciences & Humanities, 15(6), 883-95.

McCargo, D. (2009), Thai Buddhism, Thai Buddhists and the southern conflict.
Journal of Southeast Asian Studies, 40(1), 1–10

Melvin, N.J. (2007), Conflict in Southern Thailand Islamism, Violence and the State
in the Patani Insurgency. Stockholm International Peace Research
Institute: CM Gruppen, Bromma, Sweden.

Nakata, T. (2010), Terrorism in the Three Southern Border Provinces: Its Major
Causes. Songklanakarin J. of Social Sciences & Humanities, 16(6), 911-
32.

Paul, M., & Saha, R. (2007), The generalized linear model and extensions: A review
and some biological and environmental applications. Environmetrics, 18,
421-43.

Peleg, K., Daniel, L., Michael, M., Shapira, S. & Israeli Truama Group. (2003),
Patterns of Injury in Hospitalized Terrorist Victims. American journal of
emergency medicine, 21(4), 258-62.

Peleg, K., Daniel, L., Stein, M., Michaelson, M., Kluger, Y., Simon, D., Israeli
Trauma Group., & Noji, E. (2004), Gunshot and Explosion Injuries
Characteristics, Outcomes, and Implications for Care of Terror-Related
injuries in Israel. Annals of Surgery, 239(3), 311-8.

R Development Core Team. R. (2010), A language and environment for statistical
computing. R Foundation for Statistical Computing,   Vienna, Austria.
ISBN 3-900051-07-0. [Cited 2011 April 1]. Available from: URL:
http://www.R-project.org

Sheffy, N., Mintz. Y., Rivkind, A., & Shapira, S. (2006), Terror-Related Injuries: A
    Comparison of Gunshot Wounds Versus Secondary-Fragments-Induced
    Injuries from Explosives. J Am Coll Surg, 203(3), 297-303.

Venables, W.N., & Ripley, B.D. (2002), Modern Applied Statistics with S. fourth ed.,
    Springer, New York, USA.

Ward, A., & Hackett, J. (2004), Thailand's southern discomfort Crackdown against
    Muslim militancy. Strategic Comments, 10(4), 1-2

Wikipedia. (2011), Southern Thailand Insurgency. The Free Encyclopedia,. [Cite
    2011 August 1]. Available from: URL: http://en.wikipedia.org/wiki/
    South_Thailand_insurgency