Chapter 4

Statistical Modeling

In this chapter statistical modeling was performed in order to identify the strengths of association between the outcome and determinants. Poisson regression was employed to fit the model to the liver cancer data in Songkhla province.

4.1 Modeling fitted

Since the outcome is the incidence of liver cancer patients, poisson regression model is an appropriate method for assessing the effects of the determinants (gender, age groups, religion, topography, area and year) as described in Chapter 2, equation (2.4). Three different models are assessed. The first model includes all six determinants. Since the two types of liver cancer (HCC and CCA) have different causes and originate at different organs in the body, the second and third models are fit stratified by cancer type. Figure 4.1 shows the scatter plots of the observed number of liver cancer patients plotted against the fitted values of the three models (top panel) and the normal quartile plots (bottom panel). The scatter plots of the observed count and fitted values for the first and third model are scattered along the diagonal line. The normal quartile plots of the three models are similar. Therefore the first model is considered to be the best model for the data. Table 4.1 shows the values of the estimated parameters.

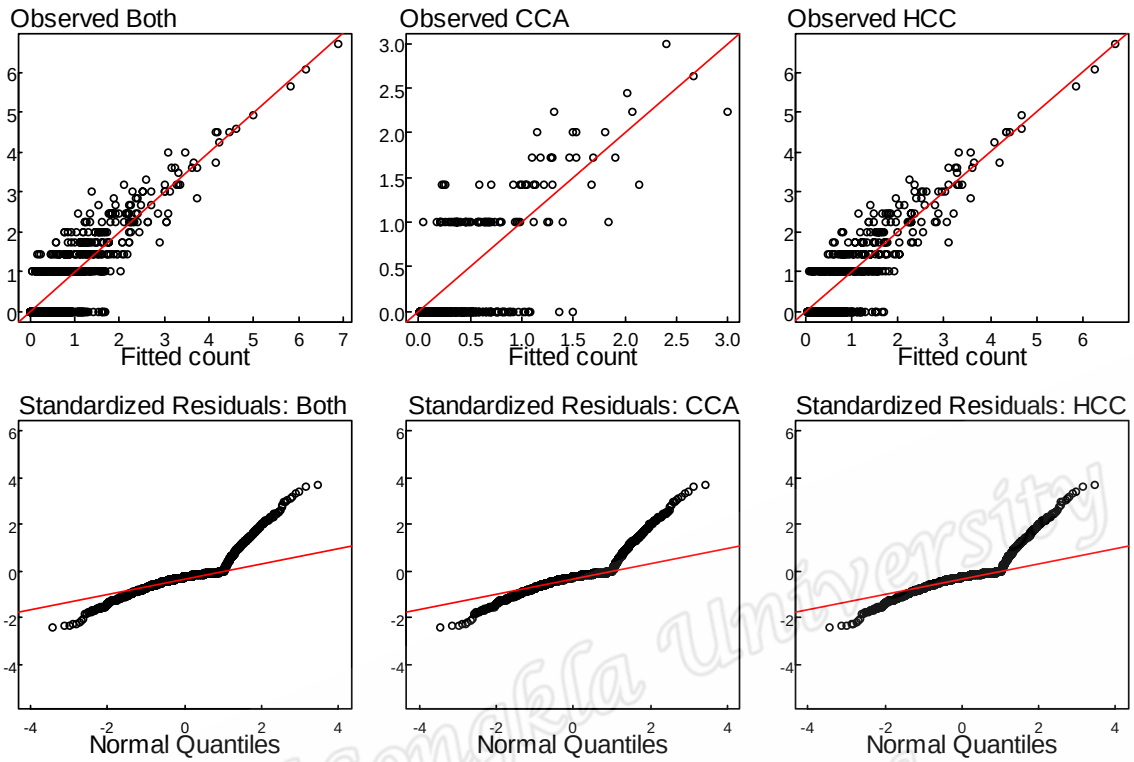

Figure 4.1 Diagnostic plots of residuals for Poisson model, top panel shows observed counts versus fitted values while bottom panel shows normal-quantile plots

The results in Table 4.1 clearly show that males are more likely to develop liver cancer than females. The liver cancer tends to occur in the older age group. The HCC liver cancer type is more prevalent than CCA. The incidence of liver cancer was highest in the period 2004-2007.

Factor	Coefficient	SE	P-value
Constant terms	-2.142	0.252	< **0.000**
Gender			
Male	0	0	
Female	-1.474	0.070	< **0.000**
Age Group			
<30	0	0	
30-44	2.927	0.226	< **0.000**
45-59	4.346	0.219	< **0.000**
60+	5.058	0.217	< **0.000**
Religion			
Islam	0	0	
Buddhist or Other	0.423	0.090	< **0.000**
Topography			
HCC	0	0	
CCA	-1.915	0.084	< **0.000**
Areas			
Muang	0	0	
Chana + Namom	0.005	0.127	0.966
Nathawi	0.014	0.154	0.926
Thepa	-0.137	0.165	0.408
Saba Yoi	-0.797	0.244	< **0.001**
Ranod+ Krasea Sinthu+Sathing Phra+ Khuan Niang	-0.172	0.104	0.098
Rattaphum	0.202	0.139	0.147
Sadao+Khlong Hoikhong	0.089	0.119	0.453
Hat Yai	0.226	0.092	< **0.010**
Bang Klam	-0.341	0.255	0.181
Singha Nakorn	-0.270	0.149	0.070
Period			
< 1992	0	0	
1992-1995	-0.182	0.119	0.127
1996-1999	-0.123	0.117	0.295
2000-2003	0.446	0.105	< **0.001**
2004-2007	1.095	0.097	< **0.001**

Table 4.1 Coefficients, standard error and p-value from Poisson model

4.2 Confidence interval

A 95% confidence interval graph of the liver cancer incidence rate/100,000 for each factor based on poisson regression model using sum contrasts are shown in Figures 4.2-4.4. The dotted horizontal lines represent the overall mean incidence rate of 2.62.

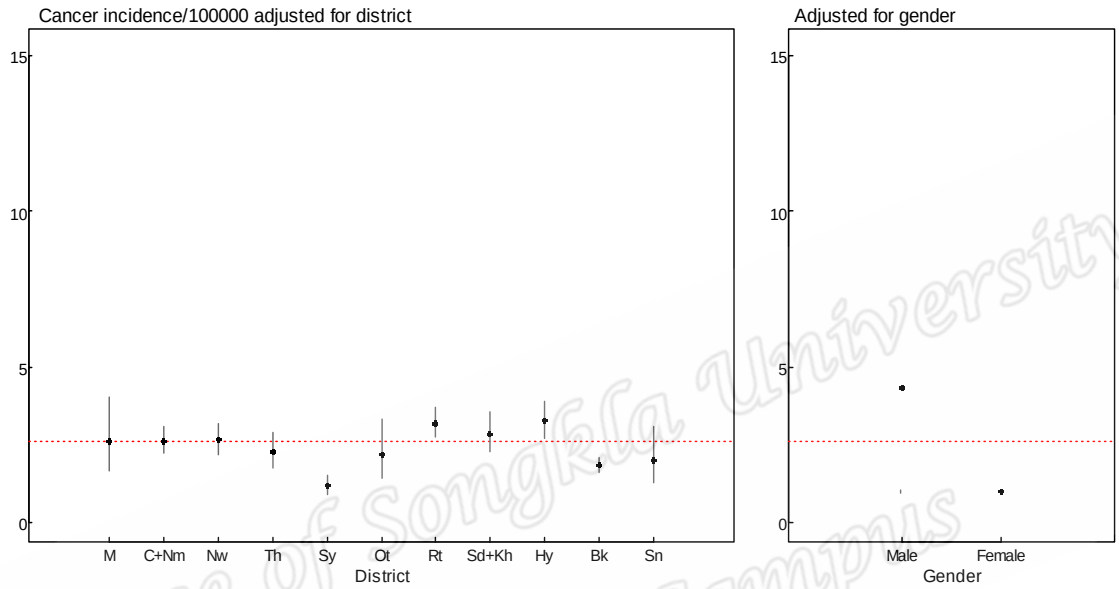

Figure 4.2 Confidence intervals for district and gender of incidence rates/100,000 adjusted for the other factors

Figure 4.2 shows the association between the incidence rate of liver cancer by district (left panel) and gender (right panel), respectively. The districts with incidence rates higher than the overall mean are Rattaphum and Hat Yai. Males have an incidence rate higher than the overall mean.

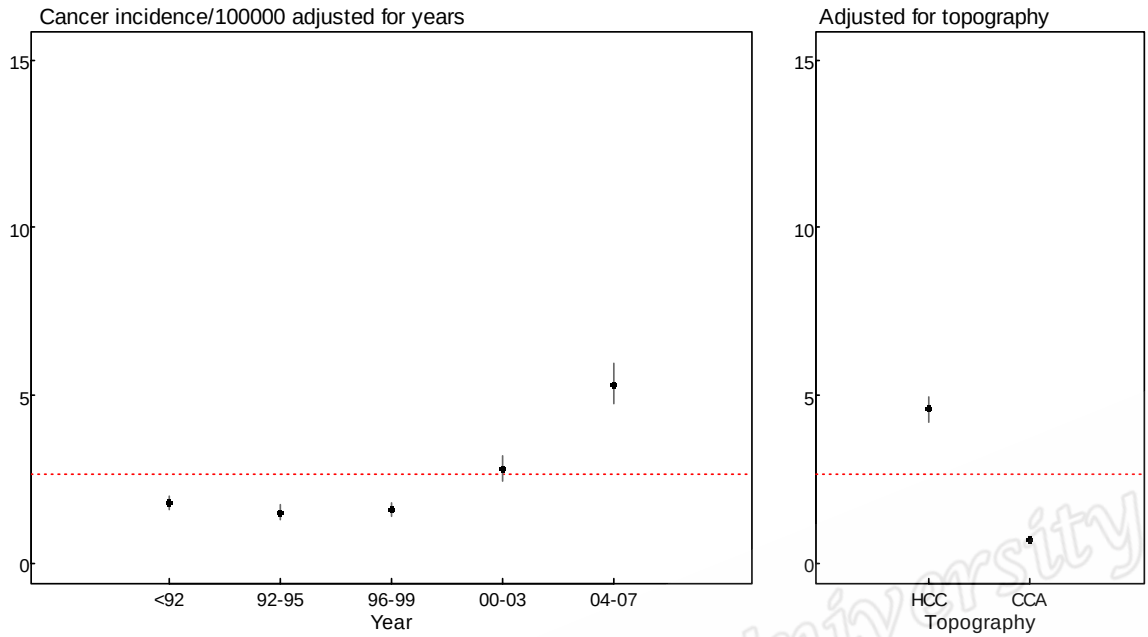

Figure 4.3 Confidence intervals for year and topography of incidence rates/100,000 adjusted for the other factors

Figure 4.3 shows the association between the incidence rate of liver cancer by year (left panel) and topography (right panel), respectively. The incidence rate by year decreased from 1989 to 1995 and increased from 1996 to 2007, year 2004-2007 the incidence rates were higher than overall mean. The incidence rate of HCC was higher than CCA.

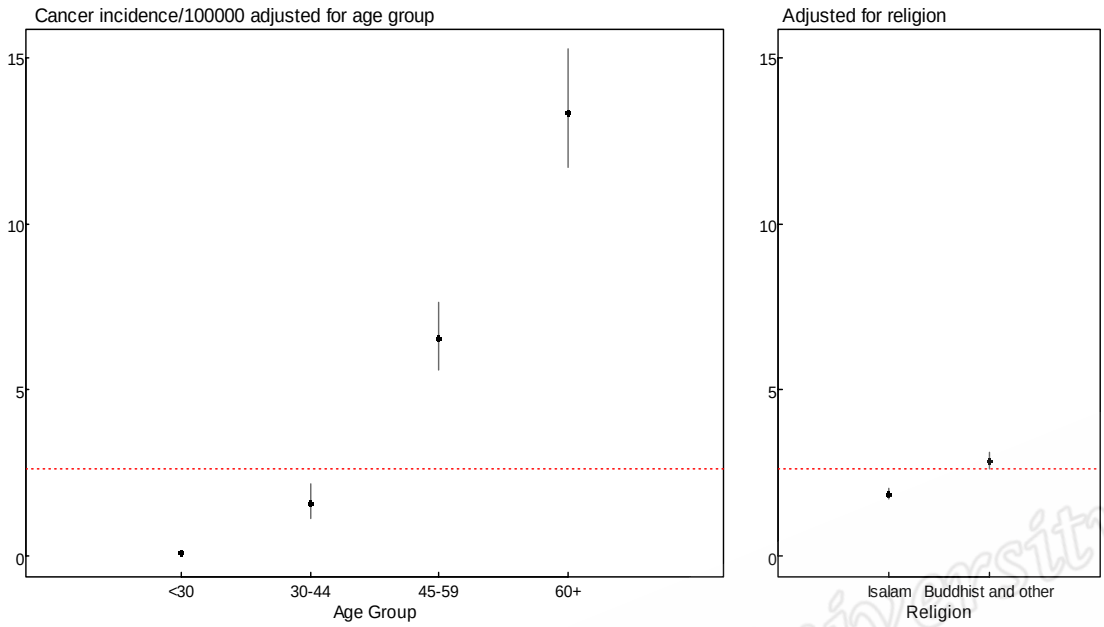

Figure 4.4 Confidence intervals for age group and religion of incidence rates/100,000 adjusted for the another factor

Figure 4.4 shows the association between the incidence rate of liver cancer by age group (left panel) and religion (right panel), respectively. The result showed that incidence rate are increasing with the older age group and peak at age more than 60 years. The incidence rate of Buddhist or other religions are higher than Islam.

4.3 Distribution of incidence rate of liver cancer in Songkhla province 1989-2007 classified by districts adjusted for the another factor

The distribution of liver cancer incidence rates/100,000 population by districts is displays by schematic maps using the information from the right panel of Figure 4.2. In Figure 4.5 the darkest shades represent the incidence rate higher than overall mean, gray shades represent the incidence rates that were similar to overall mean and light gray shades represent the incidence rates that were below the overall mean. Hat Yai and Rattaphum had incidence rates higher than the overall mean, and Saba Yoi and Bang Klam had incidence rates lower than the mean. The other districts had similar to overall mean.

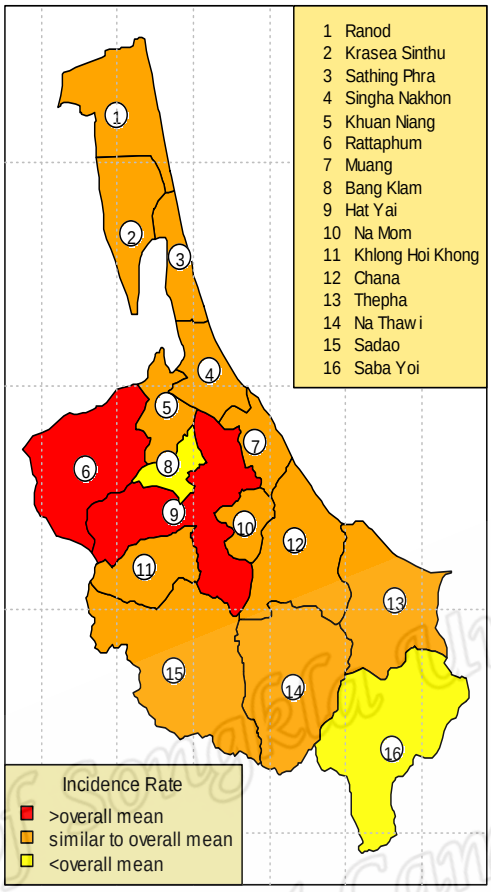

Figure 4.5 Schematic maps of liver cancer incidence rate adjusted for another factor