Trend and Levels of Liver Cancer Mortality in the Upper North-eastern Region of Thailand: A Case Study of Using Google Earth in Public Health

Patarapan Odton¹, Don McNeil² and Chamnein Choonpradub²

¹ International Health Policy Program, Ministry of Public Health, Thailand
² Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Thailand

Abstract

Objectives: (i) To investigate the demographic, temporal, and spatial variation of Liver cancer mortality in the upper north-eastern region of Thailand (Public Health Area 6). (ii) To develop a visualization tool for displaying the statistical analysis of liver cancer mortality.

Methodology: All registered home deaths aged 30 years and over due to liver cancer in the upper north-eastern region of Thailand during 1996-2006 were analyzed using a Poisson regression model. The statistical model and statistical graphs were developed using R commands. In addition, KML source code for displaying the analysis results in Google Earth was created using R commands.

Results: The Poisson model with gender, age group, year and super-district (a collection of neighboring districts in the same province with population approximately 200,000) as the factors provided a good fit to liver cancer mortality in the study region. All super-districts in Sakon-Nakhon province have higher than average liver cancer mortality, while all super-districts in Khon-Kaen and Loei provinces have lower mortality. Finally, a mortality map displayed on Google Earth gives insight into district-level variation of liver cancer mortality across the study region.

Conclusions: There were differences in liver cancer mortality across 25 super-districts in the upper north-eastern region of Thailand. The combination of Google Earth and open statistical package such as R provide an inexpensive functional tool for analysis and visualization for public health research.
Introduction

Thailand has a high burden of liver cancer, which appears to have increased substantially in the past 10 years. During 1996-2006, 24% of cancer deaths were assigned as liver cancer with ICD-10 coding as C22. Liver cancer was the fourth leading cause of death for males and the fifth for females in 1999 (Bundhamcharoen et al, 2002). A recent burden of disease study reported that liver cancer is the fourth leading cause of deaths and Years of Life Lost (YLLs) for males and the fifth for females in 2004 (Thai Working group on Burden of Disease, 2007).

Liver cancer mortality varies over the regions in Thailand. An analysis of mortality in the Thai population during 2001-2005 (Bureau of Policy and Strategy, 2007) found that residents in the north-east region had higher liver cancer mortality than other residents.

There were 88,532 deaths due to liver cancer during 1996-2004 in Thailand, with 81.6% outside hospital. The sixth Public Health Area, containing Nong-Bua-Lam-Phu, Khon-Kaen, Udon-Thani, Loei, Nong-Khai, Kalasin and Sakon-Nakhon province, accounted for 26.1% of all deaths in the country during 1996-2006.

A Geographical Information System (GIS) is valuable for public health research for understanding the relationship between geographic and disease transmission patterns, access to health care and health outcomes. A recent freely accessible satellite imagery tool is Google Earth. Google Earth is a virtual globe, map and geographic information program that was originally called EarthViewer 3D, and was created by Keyhole, Inc, a company acquired by Google in 2004. It maps the Earth by the superimposition of images obtained from satellite imagery, aerial photography and GIS 3D globe. It is available both as a free version with limited functionality, and as an enhanced version for commercial use ($400).

Google Earth uses the KML file format to visualize geographic data in an Earth browser (Google Code KML, 2010). We can thus create KML files to pinpoint locations, add image overlays, and expose rich data in new ways. KML is an international standard maintained by
the Open Geospatial Consortium, Inc. KML uses a tag-based structure with nested elements and attributes and is based on the XML standard.

**Methodology**

All registered deaths at age 30 years and over due to liver cancer (ICD10: C22) occurring outside hospital in the sixth public health area (PHA6) of Thailand were obtained from the national vital registration database. This database is provided by the Ministry of the Interior and coded as cause-of-death using ICD-10 by the Bureau of Policy and Strategy, Ministry of Public Health. The 26,640 liver cancer deaths from the vital registration database were aggregated by sex, age group (30-44, 45-59, 60-74, 75+), year, and super-district. We combined demographic variables into sex-age factor.

The population denominators were obtained from the 2000 population and housing census (National Statistical Office, 2002). Since populations of districts in Thailand vary substantially, we analyzed the mortality incidence rates in aggregated districts called “super-districts”, defined as regions comprising contiguous districts in the same province with a total population of at least 200,000 (Lim and Choonpradub, 2007). We thus obtained 25 super-districts for PHA6.

We first fitted a Poisson model containing all interactions between pairs of the four factors, and this reduced thus model by eliminating interactions that were not largely statistically significant in the ANOVA, and the combined pairs of these factors to model express the mean incidence rate per 100,000 populations as

\[
\lambda_{ij} = \exp(\mu + \alpha_i + \beta_j)
\]

In this model, the factors \(\alpha_i\) and \(\beta_j\) are combinations of age-group and years (44 levels) and gender and regions (50 levels), respectively.

The contrast matrix used in the model was “sum contrasts” so that each parameter estimate has a standard error enabling comparison with the mean (Venables and Ripley, 2002).

To compare differences in liver cancer mortality across the region of interest, visualization methods were used. Super-districts were classified into three groups, according to whether
the confidence interval for adjusted death rate was (a) totally above the mean, (b) crossing the mean, or (c) totally below the mean. Google Earth was used to display this information using corresponding colours, (a) red, (b) green, and (c) blue.

Statistical modeling and graphical displays used R commands (R Development Core Team, 2008). Furthermore, R commands were used to dynamically generate KML source codes in order to display the analysis results on the earth surface through Google Earth.

**Results**

The annual death rate due to liver cancer in the upper north-eastern of Thailand during 1996-2004 was 72 per 100,000 populations, while the average for the nation was 26 deaths per 100,000 populations.

The fitted mortality rates from applying model (1) to liver cancer mortality in the study regions are displayed in Figure 1 and Figure 2. The combined factors between age-group and year, and gender and region are statistically significant as shown on the ANOVA results in Table 1. Figure 1 shows the time series plots of liver cancer mortality in four age groups after adjusted for gender and region. The residents aged less than 45 years had lower mortality rates than the others. Liver cancer mortality rates in aged 60 and over had increased rapidly during the years 1997-2003.

**Table 1:** Analysis of variance for Poisson model of liver cancer morality rates

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Deviance residual</th>
<th>Df residual</th>
<th>Deviance residual</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age-group × Year</td>
<td>43</td>
<td>25216</td>
<td>2199</td>
<td>39091</td>
<td>&lt;0.000001</td>
</tr>
<tr>
<td>Gender × Region</td>
<td>49</td>
<td>10720</td>
<td>2107</td>
<td>3156</td>
<td>&lt;0.000001</td>
</tr>
</tbody>
</table>
Liver cancer deaths per 100,000 (adjusted for gender and region)

Figure 1: Plots of 95% confidence intervals for comparing liver cancer mortality trends

Figure 2 shows liver cancer mortality rates and their 95% confidence intervals for males (left) and females (right) in 25 super-districts after adjusted for age-group and year. The gender-specific mortality rates were 100 per 100,000 for males and 42 per 100,000 for females, as indicated by the horizontal grey lines. The overall mean was 72 per 100,000 indicated by the horizontal red line. Every super-district in Sakon-Nakhon had significantly higher mortality than the average for both sexes. Super-districts in Loei province had lower mortality than the average. All super-districts in Khon-Kaen had lower mortality than the average, except the super-district containing BanFang, PhraYun, NongRua, ManchaKhiri, KhokPhoChai, and BanHaet districts.

Figure 2: Plots of 95% confidence intervals of liver cancer mortality rates by gender and region adjusted for age-group and year.
The 95% confidence intervals of mortality rates for males and females in the 25 super-districts, as shown in the Figure 3, were compared with the average for produce bar graph on Google Earth. Fifty bars were coloured depending on their confidence interval, red-coloured for above the mean, green-coloured for across the mean, and blue-coloured for below the mean. Finally, a KML file was created for displaying the results in Google Earth as shown in Figure 4. The 3-D colour-coded bars were created using “polygon” in KML, a pair of bars on each region representing male and female morality rates. This visualization system allows the user to compare the difference in mortality rates across regions. In addition, the statistical graphs can also be added as “screen overlays” to provide more information from data analysis such as plots of confidence intervals of morality rates.

More information about the region can put in specific points on the Google Earth map for providing the related details. The example of a pop-up window, as shown in the Figure 3, contains information of a specific region, such as district names, populations, and liver cancer mortality rates for males and females.

Figure 3: Liver cancer morality map for residents in the upper north-eastern of Thailand
**Conclusions**

There was a significant difference in liver cancer mortality across 25 super-districts in the upper north-eastern region of Thailand. Poisson regression model with demographic, region, and time as factors provided a reasonable fit to liver cancer mortality in the region of interest. The differences of liver cancer mortality rates across super-districts reflect underlying factors that justify further studies. Potential factors valuable of investigation are risk behavioral patterns, level of health resources, and social determinants of health.

The combination of Google Earth and an open statistical package like R provide an inexpensive functional tool for analysis and visualization for public health research. Appropriate graphs and maps developed from the statistical results provide useful tools for assessing the geographical variations of cause-specific mortality over the study area.

**Acknowledgements**

Thank to the Thailand Bureau of Policy and Strategy, Ministry of Public Health, under the Burden of Diseases Project for providing the data. We also thank Nittaya Tongrong for providing the boundary coordinates of all 25 super-districts in the study region.

**References**


