Chapter 4

Discussion and Conclusions

The main results from our study are summarized in the first section of this chapter. The implications, advantages, and limitations are discussed in the later sections. Finally, the conclusion is made in the last section.

4.1 Findings

The Poisson model incorporating additive factors for age-group and district generally provides a good fit to mortality data in Thailand.

Higher mortality rates were found in the north of Thailand for both sexes, especially for super-districts in Chiang-Rai, Chiang-Mai and Phayao provinces. Some super-districts in the northeastern region had significant high mortality than the average. In contrast, lower mortality rates were found in the southern region with the exception of Phuket and Narathiwat.

In order to examine the sub-national cause of death quality, logistic regression models for comparing ill-defined outside hospital mortality across superdistricts in Thailand were developed in the second study.

The results from this study show that the quality of cause-of-death data varies markedly across Thailand. The proportion of death in older ages in female population was higher than those in men and this contributes to higher percentage of ill-defined cause of death among females. "Senility" (ICD10: R54) was reported as the major cause of ill-defined mortality, accounting for 58% (50% for males and 66% for females) of the total ill-defined deaths. As deaths at older age were more than half of total deaths, and "senility" was a major cause of mortality in older age in the vital statistics, high proportions of ill-defined mortality in older ages were found in our study.

People in rural area tend to have less access to health institutions and this affects their cause of death quality. The superdistricts with highest percentage of ill-defined outside hospital deaths were in Ubon-Ratchathani, Pattani, Nakhon-Phanom and Loei and these all have a high proportion of their population living in rural areas (about 80%-100%) (National Statistical Office, 2002). However, Bangkok, which should have better cause diagnosis due to concentration of hospitals and physicians showed a high proportion of ill-defined deaths. This may relate to complicated illness conditions (co-morbidity) and conditions requiring advanced investigation.

Superdistricts are more effective regional divisions than districts because they provide cells with approximately equal populations, and thus equalize the standard errors of estimated effects. Using 235 superdistricts instead of 926 districts decreases the standard errors by a factor of 2, approximately.

4.2 Implications

The geographic differences in mortality rates reflect variation in health status, morbidity, and risk behaviours, that may be related to lifestyle, access to health care and social factors.

The significant differences in the proportion of ill-defined outside hospital death rates across superdistricts in Thailand give a better picture of where the cause of death quality needs improvement, in order to improve overall (national) data quality.

4.3 Advantages

The model of mortality rates developed in the present study involves fitting a Poisson regression model in which age group and superdistrict are incorporated as factors. This model gives age-adjusted mortality rates for each geographical area, and corresponds to the method of computing standardized mortality rates (SMRs) commonly used in demographic research. Its advantage over the SMR-based method is that it routinely provides standard errors for the adjusted death rates, and thus facilitates comparison of both superdistricts and age groups.

As suggested by Tangcharoensathien et al (2006), the quality of cause of death data outside hospital could be improved by training more than 70,000 village heads, but this could be very costly. Our modeling and mapping approach on analysis of ill-defined morality as indicators of poor data quality on cause-of-death is a useful preliminary tool for enabling public health researchers to plan investigations in specific areas, possibly less than in one-third of the 926 districts.

This study shows that the geographical variation of mortality can help define targets for effective interventions and resource allocation.

4.4 Limitations

In our study on geographical variation of mortality, we assumed the percentages of registration coverage were equal for all districts. There were differences in the coverage percentages of death registration across district, this information was not available. Although the coverage of the death registration system has been reported by the National Statistical Office based on the Survey of Population Change (National Statistical Office, 1998; National Statistical Office, 2007), only country and regional levels are shown. At the national level, these surveys provide useful results for under-reporting adjustment of vital registration data (Hill et al, 2006). The variations of coverage in district level can be taken into account in a study if there is available information.

In our analysis of vital registration database, we used the place of residence rather than the place of dying. The geographical differences in mortality may not reflect the real figure of district population health because of the migration issue.

4.5 Suggestions for further study

It became apparent that mortality forecasting of Thai data will be difficult because of the absence of good-quality data for more than a few recent years. The model developed in this study could extend to time series analysis of overall and cause-specific mortality when extensive historical data are available.

The regional variations of mortality within the country may relate to socio-economic deprivation. Many studies have found a significant association between socioeconomic status and mortality (see, for example, Petcharoen et al, 2006; Vapattanawong et al, 2007; Vapattanawong et al, 2008). The demo-geographic model for mortality developed in this study could be extended by adding socio-economic factors.

Our finding on the high age-adjusted morality rates in the northern and northeastern regions may relate to socio-economic factors and excessive mortality from specific causes, such as HIV/AIDS. However, this calls for further investigation on the regional variation of cause-specific mortality.

The research on the determinants of cause of death quality could provide information to explain the findings on the differences in data quality.

The differences of mortality rates across superdistricts in different age groups reflect underlying factors that justify further studies. Potential factors worthy of investigation are risk behavioural patterns, level of health resources, and social determinants of health.

Using Google Earth as novel tool for public health research

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. Several GIS tools for public health have been developed, such as the HealthMapper developed by the World Health Organization (WHO, 2010), CONTRAST project (Stensgaard et al, 2009), Dengue Decision Support System project--DDSS (RAMS-AID, 2010), and Malaria Atlas Project--MAP (Hay and Snow, 2006). Additionally, an on-line or web-based GIS has also been developed (Gao et al, 2008).

Current examples of GIS used for public health research in Thailand are the studies by Chansang and Kittayapong (2007), Kittayapong et al (2008), and Chaikaew (2009).

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 under three different licenses: Google Earth, a free version with limited functionality; Google Earth Plus (discontinued), which included additional features; and Google Earth Pro ($400 per year), which is intended for commercial use.

The characteristics of Google Earth versus four commonly used GIS software were outlined by Lozano-Fuentes et al (2008) as shown in Table 4.1.

Table 4.1: Basic characteristics of Google Earth and commonly used GIS package

Software	Provider	Freeware	Free access to satellite images	Basic feature making tools	Spatial data processing capability	Spatial analysis and modeling capability[a]
Google Earth	Google	Yes	Yes	Yes	Low[b]	None
Epi info/Epi Map	CDC	Yes	No	Yes	Low	Low
HealthMapper	WHO	Yes	No	Yes	Low	Low
SIGEpi	PAHO	Yes	No	Yes	Moderate	Moderate
ArcGIS	ESRI	No	No	Yes	High	High

CDC, [United States] Centers for Disease Control and Prevention; ESRI, Environmental Systems Research Institute, Inc.; GIS, geographical information system; PAHO, Pan American Health Organization.

[a] Spatial analysis and modeling going beyond visual interpretation of displayed data.

[b] Restricted to basic feature-making tool

The advantages and limitations of GE have been summarized by Lozano-Fuentes et al (2008). Its advantages are (i) free software, simple and intuitive to use, (ii) a variety of colour options for points and polygons allowing development of maps simultaneously showing information related to multiple diseases and/or public health initiatives, (iii) stand-alone desktop software, Internet access needed to first "capture" the image; thereafter the image stored and processed in Google Earth on the hard drive, (iv) inclusion of simple tools to make and label polygons, lines, and points, (v) provision of free access to satellite imagery (with image quality constantly improving), (vi) Google Earth-generated KML files readily sharable among Google Earth users; and also importable as shapefiles into a GIS software for spatial analysis of the included data, (vii) a large user community with online support groups: http://groups.google.com/group/earth-help.

However, Google Earth has several limitations. First, access to the Internet is needed to initially capture satellite images of areas of interest. Second, the quality of the satellite images is highly variable, especially for rural areas. Third, the age of the images is variable. Fourth, there is only a limited set of editing and data management tool and a lack of analysis and modeling capability. Finally, Google Earth uses the world geodetic system (WGS) 84, which is well suited for country-wide or continent-wide scales, but less well suited for small locations where UTM coordinates are preferable.

Because of these GE limitations including limited data processing and analysis functions, systems with a combination of traditional GIS software and Google Earth were then developed (see, for example, Chang et al, 2008; Lozano-Fuentes et al, 2008; Kamadjeu, 2009).

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. (OGC). KML uses a tag-based structure with nested elements and attributes and is based on the XML standard.

A recent example of using Google Earth as the visualization tool is being developed for analyzing the violence incidence rates obtained from the Deep South Coordination Centre (DSCC) using an appropriate statistical model. All statistical graphs and model were produced using R commands similar to those used in this thesis, and R is also used to dynamically generate the KML source code. The estimated incidence rates from the model for Muslim and non-Muslim residents during 2004-2009 in 23 regions (aggregated districts in Songkla, Pattani, Yala, and Narathiwat province) were then used as the objects in KML tags. Finally, KML files were created for displaying the results in Google Earth as shown in the Figure 4.1. This visualization system allows the user to compare the risk patterns across regions. In addition, the statistical graphs can also be added as "screen overlays" to provide more information from data analysis such as the time series plots of incidence rates, residuals plots, and the two dimensional maps (Figure 4.2). More information about the region can put in specific point on the Google Earth map for providing the related details.

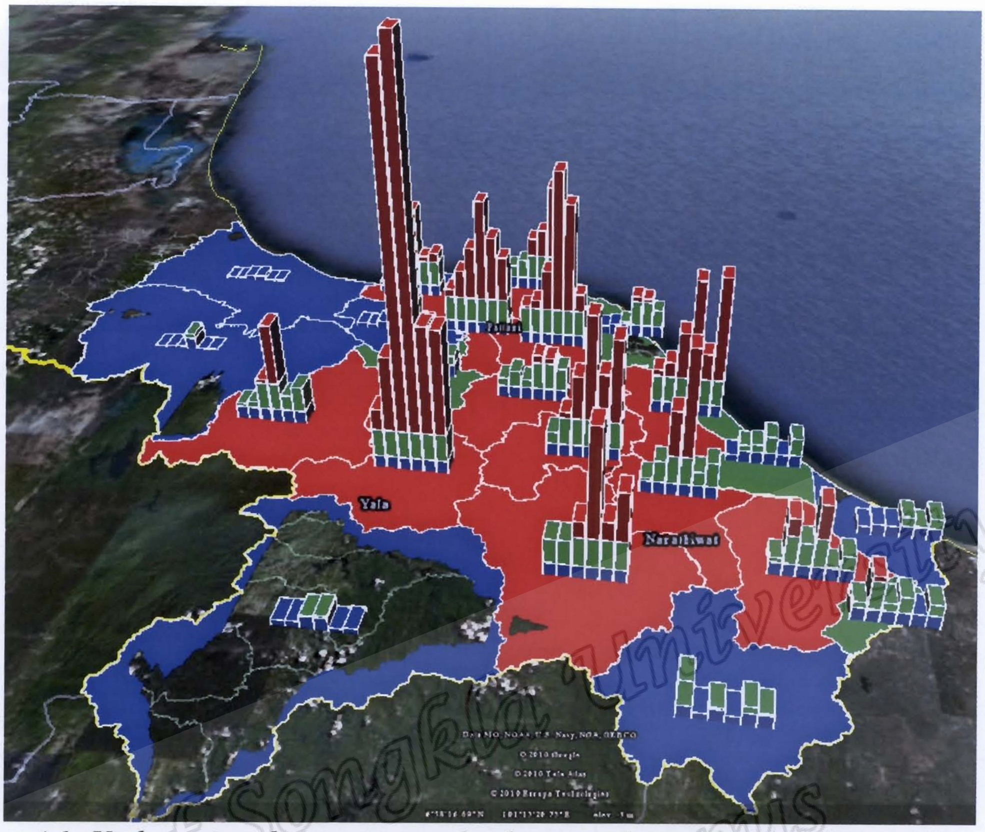

Figure 4.1: Violence incidence rates in the deep-southern Thailand during 2004-2009

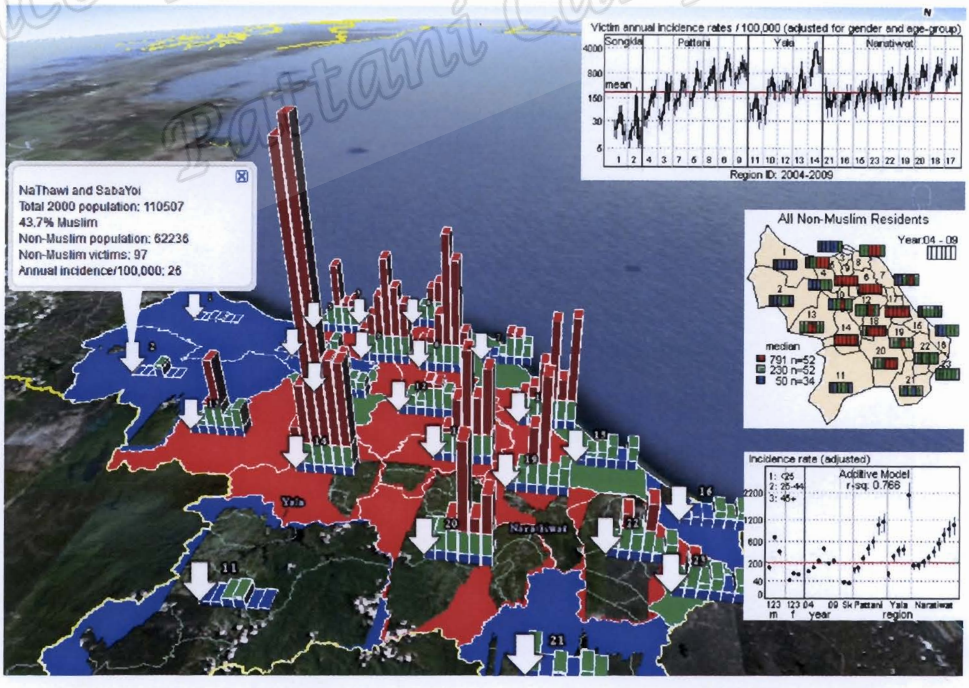

Figure 4.2: Statistical and visualization results and of violence incidence rates for non-Muslim population in the deep-southern Thailand during 2004-2009

KML files are being developed for displaying results similar to those from our studies, making regional comparisons of various mortality rates and patterns of interest.

4.6 Conclusion

The two statistical models were used to analyze mortality in Thailand during 1999-2001 and were presented through two articles.

In the first study, we developed a statistical method for comparing mortality rates across geographical regions. The method involves first fitting a statistical model based on the Poisson distribution in which age group and region are incorporated as factors. The estimated mortality rates and their 95% confidence intervals were then used to produce thematic maps, so that valid statistical conclusions could be made.

Variation of ill-defined mortality across superdistricts in Thailand is reasonably fitted to the logistic regression model, as shown in the second article. A number of contributing factors were considered in terms of age, gender, and residence.

Appropriate graphs and maps developed from the statistical results provide useful tools for assessing the geographical variations of mortality over the country.