บทความวิจัย

วิธีการทางสถิติเพื่อเปรียบเทียบอัตราการตายจากโรคอัมพฤกษ์ อัมพาต ของจังหวัดสุพรรณบุรีกับกรุงเทพมหานคร ปี พ.ศ. 2548

อารินดา มะอาลี[*]
ภัทราวรรณ ทองคำชุม[**]
นัฒกิจ พิพัฒน์จาตุรนต์[***]

บทคัดย่อ

การศึกษานี้มีวัตถุประสงค์เพื่อเปรียบเทียบอัตราการตายระหว่างประชากรสองกลุ่ม โดยใช้การถดถอยแบบปัวซง ใช้ข้อมูลการตายจากโรคอัมพฤกษ์ อัมพาต ในจังหวัดสุพรรณบุรีและกรุงเทพมหานคร ปี พ. ศ. 2548 ที่ถูกประมาณเพื่อให้มีจำนวนการตายที่ถูกต้องมากขึ้น เนื่องจากประชากรมีโครงสร้างที่ต่างกัน จึงปรับโดยใช้วิธีการปรับฐานอายุเพื่อให้สามารถเปรียบเทียบกันได้ ใช้อัตราการตายมาตรฐานเพื่อเปรียบเทียบการตายเบื้องต้น และใช้การถดถอยแบบปัวซงเพื่อสร้างตัวแบบเปรียบเทียบอัตราการตายหลังจากปรับอิทธิพลของปัจจัยด้านประชากร ผลการศึกษาพบว่า อัตราการตายมาตรฐานจากโรคอัมพฤกษ์ อัมพาต ทั้งเพศชายและหญิงในจังหวัดสุพรรณบุรีสูงกว่ากรุงเทพมหานคร แบบจำลองการถดถอยแบบปัวซงมีความเหมาะสมกับข้อมูล ซึ่งจังหวัดสุพรรณบุรีมีอัตราการเสียชีวิตสูงกว่ากรุงเทพมหานครหลังจากปรับอิทธิพลของเพศและกลุ่มอายุแล้วและยังพบอีกว่าเพศชายมีอัตราการตายสูงกว่าเพศหญิงในทุกกลุ่มอายุที่ต่ำกว่า 80 ปี วิธีการทางสถิติที่ใช้ในการศึกษานี้ สามารถนำไปประยุกต์ใช้กับการศึกษาอื่นที่ต้องการเปรียบเทียบอัตราการตายจากสาเหตุเฉพาะโรคระหว่างประชากร

คำสำคัญ: กรุงเทพมหานคร, การถดถอยแบบปัวซง, การปรับมาตรฐานอายุ, โรคอัมพฤกษ์ อัมพาต, สุพรรณบุรี

---

[*] นักศึกษาปริญญาเอก วิธีวิทยาการวิจัย ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี, และศูนย์ความเป็นเลิศทางคณิตศาสตร์ สำนักพัฒนาบัณฑิตศึกษาและวิจัยด้านวิทยาศาสตร์และเทคโนโลยี คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล

[**] ดร. (สถิติ) ผู้ช่วยศาสตราจารย์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี

[***] ดร. (วิธีวิทยาการวิจัย) สำนักงานป้องกันควบคุมโรคที่ 2 พิษณุโลก จังหวัดพิษณุโลก

RESEARCH

Statistical methods for comparing stroke death rates for Suphan Buri with Bangkok in 2005

Arinda Ma-a-lee[*]
Phattrawan Tongkumchum[**]
Nattakit Pipatjaturon[***]

Abstract

This study aimed to compare stroke death rates between two populations using Poisson regression model. The stroke deaths for Suphan Buri and Bangkok in 2005 were analyzed after adjusting for misclassification of cause of death. Age-standardization was used to adjust for population differences in gender and age group. Standardized mortality ratio (SMR) was used to explore deaths. A Poisson model was used to compare stroke death rates after adjusting for demographic factors. The results showed that SMRs were greater than 1 for both males and females. These SMRs indicated that there were excess stroke deaths for both males and females in Suphan Buri compared to Bangkok. The Poisson model was fitted well with the stroke data. The stroke death rate was statistically significant higher in Suphan Buri than those in Bangkok after adjusting for gender-age-group. The rates were higher for males than for females in all age groups under 80 years. Therefore, increasing efforts in health care and stroke prevention should target at males in suburban areas. The statistical methods used in this study can be applied to other regions in Thailand and to other countries for comparing cause-specific mortality between populations.

Keywords: Bangkok, Poisson regression, standardization, stroke, Suphan Buri

[*]Ph.D. student, Research Methodology, Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, and Centre of Excellence in Mathematics, CHE, Si Ayutthaya Rd., Bangkok 10400, Thailand.

[**]Dr. (Statistics) Asst. Prof., Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus.

[***] Dr. (Research Methodology), the office of Diseases Prevention and control 2th Phitsanulok, Phitsanulok.

Introduction

Stroke is the second leading cause of death and the leading cause of adult disability worldwide (Mathers and Loncar, 2006). The number of stroke deaths, particularly in developing countries, is expected to increase in future due to ageing populations, the effects of tobacco use, and the rising burden of hypertension and diabetes (Hoy et al, 2013).

In Thailand, more than 300,000 die every year. These deaths comprise more than 50,000 reported annually for stroke from the Thai Ministry of Public Health, which is a major health burden in Thailand. It is also the leading cause of death and long term disability in both males and females (Rao et al., 2010). Another study (Suwanwela, 2014) found the prevalence of stroke to be 1.88% among adults 45 years or more. It is more prevalent in men than women and the mean age of onset is 65 years. There is also a variation of stroke deaths between the five regions of Thailand with the highest prevalence in the central region (2.41%), followed by the southern (2.29%), northern (1.46%), and north-eastern regions (1.09%).

Place of residence is an important determinant of population health (Tideman et al., 2014). In many settings worldwide, there is underinvestment in health-promoting infrastructure and geographical health disparities exist. The health of persons living in suburban area is worse in many dimensions compared with urban populations. In order to assess geographical health disparities, focusing on a specific area can clarify understanding of disparities (Fukuda, Nakamura and Takano, 2004; Rodriguez-Fonseca et al., 2013).

The central region of Thailand comprises 11 provinces. Of these, Suphan Buri and Bangkok were sampled in the 2005 Verbal Autopsy (VA) survey (Rao et al., 2010). VA studies are widely used to verify causes of death where mortality data are of poor quality (Rao et al., 2010; Klinjun, Lim and Bundhamcharoen, 2014; Herbst, Mafojane and Newell, 2011).

Suphan Buri is a suburban area which is located northwest of Bangkok, and has population density in the range 100-250 persons per square kilometer, whereas Bangkok is the capital city of Thailand, and has population density in the range 400-4,060 persons per square kilometer (National Statistical Office, 2000). In addition, health care systems are located to a greater extent in Bangkok than in Suphan Buri and other places in Thailand. For these reasons our study assessed this geographical health disparity by comparing stroke mortality between Suphan Buri and Bangkok. Moreover a

comparison based on the magnitude of health disparities is of key importance in identifying opportunities to reduce health disparity.

Comparison of mortality between populations requires a method known as age- standardization. This method provides a summary measure that controls for the population age distribution (Curtin and Klein, 1996). Several studies represent the mortality difference based on the Standardized Mortality Ratio (SMR) (Anderson et al., 1998; Sather et al., 2005; Torrey and Haub, 2004). However, few studies evaluate mortality differences using statistical models, and modeling provides a means of gaining insight into details of health disparities.

This study describes methods for comparing stroke death rates for persons aged 30 years or more between two populations. The methods involve first estimating the number of stroke deaths after adjusting for misclassification of cause of death based on more accurate ICD-10 diagnosis from the 2005 VA survey. Second, age-standardization was used to compare populations with different age distributions. Third, the Poisson model was then used to compare stroke death rates after adjusting for demographic factors.

Methods

Data sources and management

The 2005 verbal autopsy (VA) and Thai death registration (DR) data were used to estimate numbers of stroke death. Both databases were obtained from the Bureau of Health Policy and Strategy, Ministry of Public Health. The VA data contain information of deceased persons for province of residence, gender, age group, location (death occurrence in or outside hospital), DR reported ICD-10 code and VA-assessed ICD-10 code.

It is known that the DR data misclassifies cause of death (Mathers et al., 2005). The misclassification was corrected using the VA data, based on a method described in detailed by Chutinantakul et al (Chutinantakul et al., 2014). Figure 1 shows a flow diagram for estimating stroke deaths. Data management began with systematic analysis of the VA data used chapter-block classification of ICD-10 codes (WHO, 2004), where blocks are classified mainly by human organs, creating 21 major cause groups with VA counts. A logistic regression model was fitted to stroke cause with demographic (province, gender, age-group) and medical (reported cause group and location) factors as determinants. A Receiver Operating Characteristic (ROC) was used to assess the model (Sakar and

Midi, 2010). A spatial triangulation method was used to estimate coefficients for the provinces outside the VA study. The estimated probabilities from the model were applied to the DR data giving the estimated numbers of stroke deaths for the whole of Thailand. Then deaths in Suphan Buri and Bangkok were chosen for comparison.

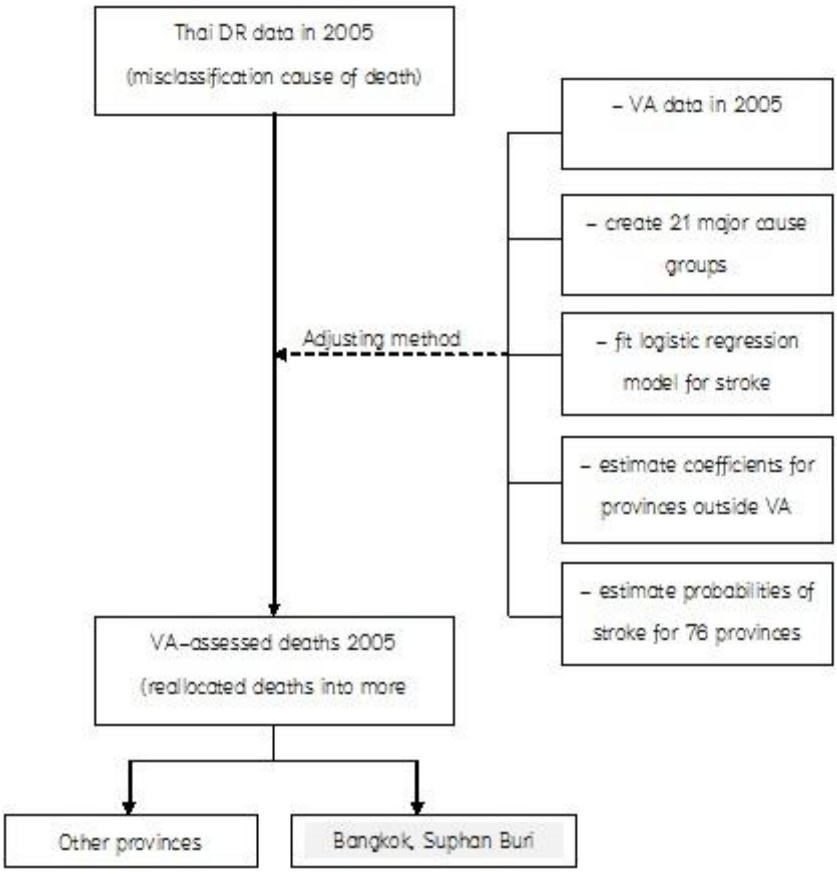

Figure 1. Diagram for estimating stroke deaths

The 2005 projected populations for Suphan Buri and Bangkok were used. The projection method is based on the assumptions about mortality and female fertility: we developed a mathematical model for projection, which is formulated in terms of the size of the population, the mortality and fertility rate per unit population. To avoid the estimation of the age-specific fertility, we assume that births decrease by 2% annually. From the model we start with gender-specific population census age distributions in 2000, and we used reported death rates by age and gender from 2000 to 2009, thereafter assumed constant.

Thus projected populations for the whole of Thailand from 2000 to 2009 were obtained. Finally Suphan Buri and Bangkok were chosen for comparison.

Preliminary analysis

Suphan Buri and Bangkok have different age distributions. After age-standardization, we applied Bangkok's gender and age-specific death rates to the Suphan Buri population. This calculation provides the expected number of deaths by gender (male and female) and age group (30-39, 40-49, 50-59, 60-69, 70-79, 80+) for Suphan Buri, based on corresponding rates for Bangkok. For each gender, the Standardized Mortality Ratio (SMR) was then calculated by dividing the total number of observed deaths by the total number of expected deaths. Since this method does not directly provide confidence interval and p-values, we fitted a statistical model based on Poisson generalized linear model.

Statistical methods

We considered Poisson regression for comparing stroke death rates after adjusting for gender-age group (Venebles and Ripley, 2002).

The Poisson regression model is appropriate for modeling count data. The formulation for the Poisson model is based on two set of predictors, if λ_{ij} denotes the mean death rate in province *i* and the gender-age group *j*, *P* is the population at risk per 100,000 for each data cell, an additive model with this distribution has mean λ_{ij} , where

$$\ln(\lambda_{ij} / P_{ij}) = \mu + \alpha_i + \beta_j . \qquad \text{(a)}$$

In this model, α_i and β_j are effects of province and gender-age group, respectively, which sum to zero, and μ is a constant encapsulating the overall the death rate. The Poisson distribution has only the single parameter λ , and the variance is also λ .

Plot of residuals against normal quantiles is used to assess how well a model fits the data. For Poisson model deviance residuals based on the likelihood are recommended. The plotted residuals should follow a straight line on the plot corresponding to the normal quantiles.

Confidence intervals (CIs) were obtained from the model for comparing the death rates within each factor with the reference group. The reference for gender-age group with maximum deaths was chosen.

R program version 2.15.2 was used for data management, statistical analysis and graphical displays (R core team, 2014).

Results

Comparison of population age distribution for Suphan Buri and Bangkok

Figure 2 shows population's age distribution for Suphan Buri and Bangkok by gender. Bar strips represent the proportions of population by age-group. These two populations have different age distributions. Bangkok has a greater proportion of population aged 30-39 and 40-49 years for both males and females than Suphan Buri whereas Suphan Buri has a higher proportion of people aged 60 years and older especially for females.

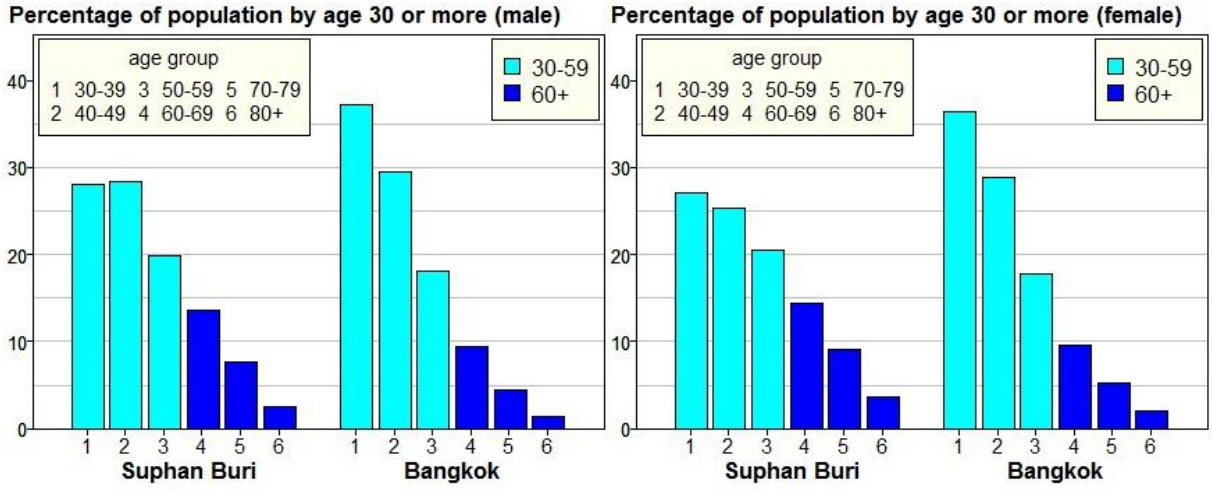

Figure 2. Percentage of population age distribution for Suphan Buri and Bangkok by gender for persons aged 30 or more.

Comparison of age-specific death rates

Figure 3 shows gender and age-specific stroke death rates. Both in Suphan Buri and Bangkok, rates for males were greater than those for females under 80 year. Both sexes, Bangkok has lower rates than those in Suphan Buri in every age group. For clarity, a cube root y-axis scale is used, however this graphs do not accurately depict the extra burden of death in Suphan Buri compared to Bangkok.

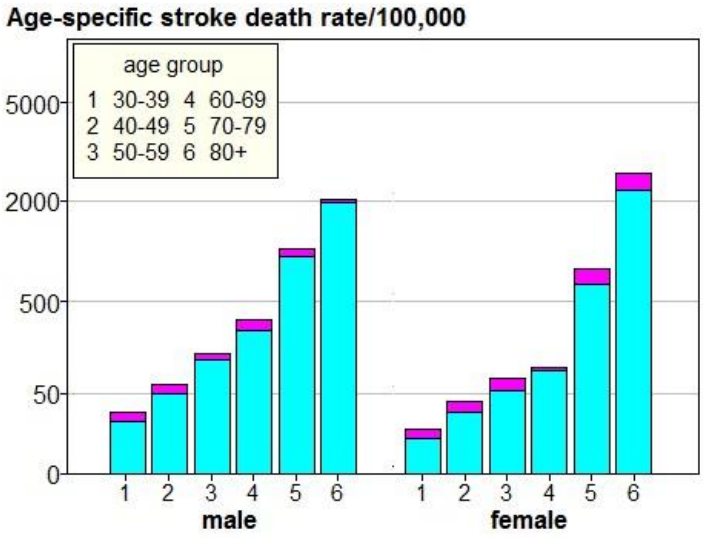

Figure 3. Age-specific death rates for stroke between Suphan Buri and Bangkok by gender-age group for persons aged 30 or more.

Comparison of age-standardized deaths

Applying Bangkok's gender and age-specific stroke death rates to the Suphan Buri population, the total number of stroke deaths for Bangkok (929) would be 83% of the number of stroke deaths in Suphan Buri (1,119) as shown in Figure 4. The excess stroke deaths in Suphan Buri are only 16% (559/483) for males although female deaths are 26% (560/446) higher. Suphan Buri excess stroke death rates are higher than Bangkok in every age group. To detect statistical significant excess an appropriate statistical model is needed.

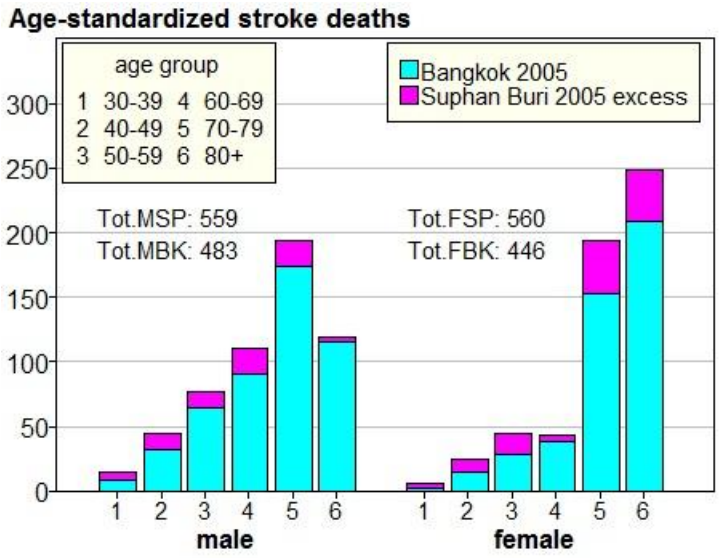

Figure 4. Age-standardized stroke deaths for Suphan Buri and Bangkok by gender-age group for persons aged 30 or more.

Statistical model results

The results of the statistical analysis are presented in Figure 5. The plot in the upper panel shows residual plot for the Poisson model. The model was fitted well with the data, and the chi-squared test indicates that the fit is statistically acceptable. The lower panel shows the adjusted stroke death rates with 95% confidence intervals. These confidence intervals use treatment contrasts to compare death rate in Suphan Buri with Bangkok. Bangkok is a reference group for province and male aged 70-79 is a reference group for gender-age group. The + symbols denoted the unadjusted ("crude") death rates. There is a little difference between the crude and adjusted death rates, so there is little evidence of confounding.

The stroke death rates (per 100,000 population) was statistically significant higher in Suphan Buri (232; 95%CI. 159-271) than those in Bangkok (193) after adjusting for gender-age group. Rates were higher in males than in females in every age group under 80 years.

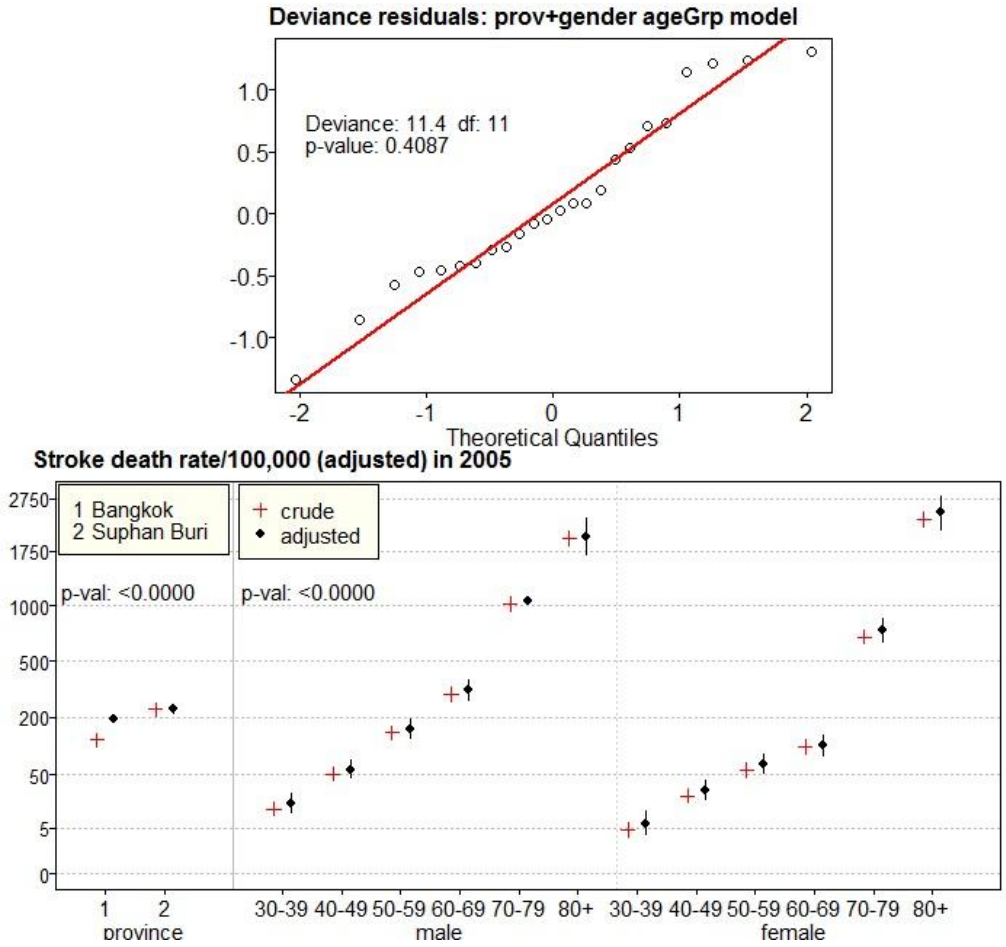

Figure 5. Poisson regression analysis for comparing stroke death rates between Suphan Buri and Bangkok after adjusting for gender-age group, with these factors unadjusted (denoted by + symbols) and after adjusting for other factor (denoted by solid circle): Bangkok as a reference group for province and male aged 70-79 as a reference group for gender-age group.

Discussion

In this study, SMR and Poisson model were the appropriate methods for comparing stroke death rates between populations in suburban (Suphan Buri) and urban city (Bangkok).

The SMRs were higher in Suphan Buri for both males and females. Several studies have used SMRs to compare all cause and cause-specific excess deaths including stroke (Fukuda, Nakamura and Takano, 2004; Torrey and Haub, 2004; Faramnuayphol, Chongsuvivatwong and Pannarunothai, 2008).

However, the main objectives of these studies was not compare death rates based on the statistical model.

Our study used a Poisson regression model to compare stroke death rates between populations. The model provided a good fit to stroke mortality data as shown in the upper panel plot on Figure 5. This model gives death rates for two factors (province and gender-age group) adjusted for other factor. A Poisson model has the advantage over the SMR method as it directly provides a 95% confidence interval and p-value for adjusted death rates, therefore it helps to facilitate comparison of death rates between populations.

It is known that Poisson regression is a method to model count data (Venables and Ripley, 2002). Previous studies were successfully used Poisson model to compare stroke mortality among different populations (André et al., 2006; Saito et al., 2016). For example, it has been used to verify the change for stroke mortality in different regions in Brazil from 1980 to 1982, 1990 to 1992 and 2000 to 2002 (André et al., 2006). The model clearly assesses the rates of changes in stroke mortality between populations from five regions.

The adjusted stroke death rate from the model was much higher in Suphan Buri than in Bangkok. This mortality disparity may partly be due to differences in access to quality stroke care and the prevalence of stroke risk factors among these areas (Tideman et al., 2013; Yiengprugsawan et al., 2010; OECD/WHO, 2014). An important understanding of these situations will assist in efforts to reduce health disparities.

Suphan Buri was found to be high in prevalence of hypertension which is a major risk factor of stroke (Chongsuvivatwong et al., 2010). A study in China showed that the overall relative risk of stroke associated with hypertension was 5.43 (Wang et al., 2009). This evidence led to more emphasis in its prevention, and therefore more emphasis given towards rising hypertension, especially in population who live in suburban areas such as Suphan Buri (Chongsuvivatwong et al., 2010).

Our finding also found that stroke death rates for males were greater than those for females with an exception for females aged 80 years or more. Excess female stroke mortality in oldest age group needs to be investigated further.

Using estimated deaths based on more accurate ICD-10 diagnosis from the 2005 VA data has the obvious advantage that it is reporting the real situation within a population. It gives a sense of anchorage in understanding prevailing epidemiological conditions, and, in doing so, enables the development of context-specific program for improving the health of that population (Hoy et al., 2013; Klinjun, Lim and Bundhamcharoen, 2014).

Further investigation will be needed to investigate the differences in health care resources, lifestyle behavior, and other risk factors in suburban and urban areas that might explain the suburban/urban difference in stroke death rates.

Conclusion

This study reveals that age-standardization must be used to adjust for population differences in age distribution. SMR was used to explore excess deaths and Poisson model was used compare death rates after adjusting for demographic factors. Stroke is still the main threat to the health of population in Suphan Buri (suburban area). The statistical methods used in this study can be applied to other regions in Thailand and to other countries for comparing death rates between populations.

Acknowledgements

This study was funded by the Program Strategic Scholarships Fellowships Frontier Research Networks (Specific for Southern Region) for the Ph.D. Program Thai Doctoral degree from Office of the Higher Education Commission of Thailand and also by the Centre of Excellence in Mathematics, the Commission on Higher Education, Thailand. We would like to thank the Ministry of Public Health and Dr. Kanitta Bundhamcharoen for providing us data. We also would like to express our thanks to Prof Don McNeil for assistance and helpful guidance.

References

Anderson, L.B., Vestbo, J., Juel, K., et al. 1998. "A comparison of mortality rates in three prospective studies from Copenhagen with mortality rates in the central part of the city, and the entire country", **European Journal of Epidemiology**. 14, 579 – 585.

André, C., Curioni, C.C., Cunha, C.B.d., et al. 2006. "Progressive decline in stroke mortality in Brazil from 1980 to 1982, 1990 to 1992, and 2000 to 2002", **Stroke.** 37, 2784–2789.

Saito, I., Yamagishi, K., Kokubo, Y., et al. 2016. "Association between mortality and incidence rates and coronary heart disease and stroke: The Japan Public Health Center–based prospective (JPHC) study", **International Journal of Cardiology**. 222, 281–286.

Chongsuvivatwong, V., Yipintsoi, T., Suriyawongpaisal, P., et al. 2010. "Comparison of cardiovascular risk factors in five regions of Thailand: Inter ASIA data", **Journal of the Medical Association of Thailand**. 93, 17–26.

Chutinantakul, A., Tongkumchum, P., Bundhacharoen, K., et al. 2014. "Correcting and Estimating HIV mortality in Thailand based on 2005 verbal autopsy data focusing on demographic factors, 1996–2009", **Population Health Metrics**. 12, 1 – 8.

Curtin, L., & Klein, R. 1996. "Direct standardization (Age–adjusted death rates)", **Department of Health and Human Services. Public Health Service**. 6, 1 – 11.

Faramnuayphol, P., Chongsuvivatwong, V., & Pannarunothai, S. 2008. "Geographical variation of mortality in Thailand", **Journal of the Medical Association of Thailand**. 91, 1455 – 1460.

Fukuda, Y., Nakamura, K., & Takano, T. 2004. "Increased excess deaths in urban areas: quantification of geographical variation in mortality in Japan, 1973–1998", **Health & Policy**. 68, 233 – 244.

Herbst, A.J., Mafojane, T., & Newell, M.L. 2011. "Verbal autopsy–based cause–specific mortality trends in rural Kwazulu–Natal, South Africa, 2000–2009", **Population Health Metrics**. 9, 1 – 13.

Hoy, D.G., Rao, C., Hoa, N.P., et al. 2013. "Stroke mortality variations in South–East Asia: empirical evidence from the field", **International Journal of Stroke**. 8, 21 – 27.

Klinjun, N., Lim, Apiradee., & Bundhamcharoen, K. 2014. "A logistic regression model for estimating transport accident deaths using verbal autopsy data", **Asia Pacific Journal of Public Health**. 1 – 7.

Matcher, C.D., Fat, D.M., Inoue, M., et al. 2005. "Counting the dead and what they die from: an assessment of the global status of cause of death data", **Bulletin of the World Health Organization**. 83, 171 – 177.

Mathers, C.D., & Loncar, D. 2006. "Projections of global mortality burden of disease from 2002 to 2030", **PLoS Medicine**. 3, 2011 – 2030.

National Statistical Office. 2000. **The 2000 Population and Housing Census**. Bangkok, Thailand.

Organization for Economic Co-operation Development/World Health Organization. 2014. "Health at a Glance: Asia/Pacific, 2014: Measuring Progress towards Universal Health Coverage", **OECD**.

R Core Team. 2014. A language and environment for statistical computing, Available from: http://www.R-project.org/ (20 August, 2014).

Rao, C., Porapakkham, Y., Pattaraarchachai, J., et al. 2010. "Verifying causes of death in Thailand: rationale and methods for empirical investigation", **Population Health Metrics**. 8, 1 – 13.

Rodriguez-Fonseca, M., Palencia, L., Mari-Dell'Olmo, M. et al. 2013. "Evolution of socio-economic inequalities in mortality in small geographical areas of the two largest cities in Spain (Barcelona and Madrid), 1976-2007", **Public Health**, 127, 716 – 721.

Sakar, S., & Midi, H. 2010. "Importance of assessing the model adequacy of binary logistic regression", ***Journal of*** **Applied Science**s. 10, 479 – 486.

Sather, D., Fryer, Jr.G.E., McCann, J., et al. 2005. "What if we were equal? A comparison of the Black-White mortality gap in 1960 and 2000". **Health Affairs**. 24, 459 – 64.

Suwanwela, N. 2014. "Stroke epidemiology in Thailand", **Journal of Stroke**. 16, 1 – 7.

Tideman, P., Taylor, A.W., Janus, E., et al. 2013. "A comparison of Australia rural and metropolitan cardiovascular risk and mortality: the Greater Green Triangle and North West Adelaide population surveys", **British Medical Journal**. 3, 1 – 9.

Torrey, B., & Haub, C. 2004. "A comparison of US and Canadian Mortality in 1998", **Population Council**. 30, 519 – 530.

Venables, W., & Ripley, B.D. 2002. **Modern applied statistics with S.** 4th ed. New York: Springer-Verlag.

Wang, X., Jiang, G., Choi, B.C.K., et al. 2009. "Surveillance of trend and distribution of stroke mortality by subtype, age, gender, and geographical areas in Tianjin, China, 1999-2006", **International Journal of Stroke**. 4, 169 – 174.

World Health Organization. 2004. **Special tabulation lists for mortality and morbidity: mortality tabulation list 1. In International Statistical Classification of Diseases and Health Related to Problems-Tenth Revision (Vol 1)**. Geneva, Switzeland: World Health Organization, 1163 – 1175.

Yiengprugsawan, V., Carmichael, G.A., Lim, L.L., et al. 2010. "Has universal health insurance reduced socioeconomic inequalities in urban and rural health service use in Thailand?". **Health & Place**. 16, 1030 – 1037.