Chapter 2

Methodology

This chapter describes the methodology used in the study. They comprise data sources, path diagram, variables, and statistical methods. Graphical and statistical analyses were performed using R program.

2.1 Data Sources

Diabetes mellitus mortality data in Southern Thailand were obtained from the Bureau of Health Policy and Strategy, Ministry of Public Health from 1996 to 2006. These data were collected from death certificates. The causes of deaths are diagnosed by physician or nurse when death occurs in hospital or assessed by head of village or health personnel when death occurs outside hospital. Death certificates are issued by district registrar office. The deaths have to be reported within 24 hours when death occurred and data were entered into civil registration database. This database contains not only the deaths but also birth. The database was maintained by Ministry of Interior. Ministry of Public Health used this database for coding causes of death follows by ICD-10 code. The coded database was analyzed for public health annual report. The ICD-10 codes used in this study are ICD-10, E10-E14.9 for diabetes mellitus.

The population denominator used to calculate mortality incidence rates were obtained from the civil registration, Ministry of Interior.

2.2 Path diagram and variables

The outcome was diabetes mellitus mortality rates in Southern Thailand from 1996 to 2006. The mortality rates per 1,000 were computed from numbers of deaths divided by corresponding populations at risk.

The determinants were age, gender, year and province. The schematic diagram for this study is shown in Figure 2.1.

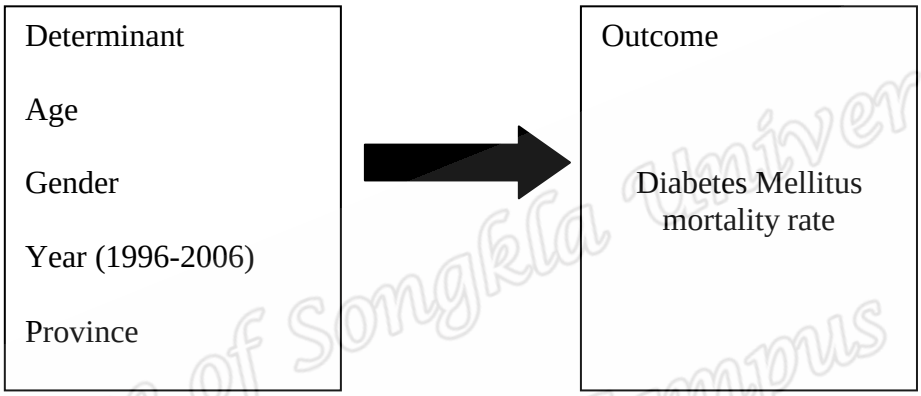

Figure 2.1: Schematic diagram of variables of interest

2.3 Statistical methods

Mortality rate

The mortality rate from diabetes mellitus was computed using equation below**.**

$$y_{ijst} = \left[K \frac{D_{ijst}}{P_{ijst}} \right] \quad (1)$$

Where D_{ijst} is the numbers of deaths for province *i*, age group *j*, gender *s*, and year *t*. P_{ijst} is population at risk and *K* is a constant.

Multiple linear regression model

Since the diabetes mellitus mortality rate was considered as a continuous outcome and the determinants comprised age group, gender, year and province, multiple regression analysis was the appropriate method for statistical modeling.

The estimated multiple regression model takes the form

$$y = \beta_o + \sum_{k=1}^{p} \beta_k x_k \qquad (2)$$

Where *y* is the outcome variable, β_0 is a constant, { β_k } is a set of parameters, and x_k is a set of determinants, *k* =1 to *p*. The model is fitted to data using least squares, which minimizes the sum of squares of the residuals. Linear regression analysis resets on three assumptions including the association is linear, the variability of the error (in the outcome variable) is uniform and these errors are normally distributed. If these assumptions are not met, a transformation of the data may be appropriate. Linear regression analysis may also be used when one or more of the determinant is categorical. In this case the categorical determinant is broken down into $c-1$ separated binary determinant, where c is the number of categories. The omitted category is taken as the baseline or referent category (McNeil, 1996). The estimated additive model for our diabetes mortality incidence rates is taken the form

$$\ln(y_{ijst}) = \mu + \alpha_i + \beta_j + \eta_s + \kappa_t \qquad (3)$$

Where y_{ijst} is the diabetes incidence mortality rates, μ is overall effect, α is the effect of province, β is the effect of age, η is the effect of gender and κ is effect of years.

Poisson model

Diabetes mellitus death is the count data that is the number of people who died from diabetes. Poisson regression model is appropriate for fitting model with count data. The probability of Poisson distribution with observed counts of *y* is given by:

$$\text{Prob (Y = } y) \ \frac{e^{-\lambda}\lambda^{y}}{y!} \tag{4}$$

Where λ is the Poisson parameter for mean and it is equal to the variance of the Poisson distribution. Poisson regression model can be fitted by using the generalized linear models (GLMs) equation with the log link function. Suppose that y_{ijst} is a number of diabetes deaths in province *i*, age group *j*, gender *s* and years *t*. Thus the Poisson regression model is taken form:

$$\ln(\lambda_{ijst}) = \ln(p_{ijst}) + \mu + \alpha_i + \beta_j + \eta_s + \kappa_t \tag{5}$$

where λ_{ijst} is the mean of y_{ijst}, p_{ijst} is number of population in province *i* age group *j* and gender *s*, α is the effect of province, β is the effect of age, η is the effect of gender and κ is effect of years. We assume $\alpha_1 = 0$, $\beta_1 = 0$, $\eta_1 = 0$ and $\kappa_1 = 0$.