

# FORECASTING TUBERCULOSIS (TB) MORTALITY IN THAILAND USING MULTIVARIATE LINEAR REGRESSION

Sampurna Kakchapati<sup>1</sup>, Chamnein Choonpradub<sup>2,\*</sup>

<sup>1</sup>Program in Research Methodology, Prince of Songkla University, Pattani Campus 94000, Thailand

<sup>2</sup>Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus 94000, Thailand

---

**ABSTRACT:** Tuberculosis (TB) is a major cause of morbidity, mortality, and disability worldwide. It is still a public health problem in Thailand. The objective of the study was to model and forecast TB mortality in Thailand using death certificate reports. A retrospective analysis of the TB mortality rate was conducted. Data were obtained from the national vital registration database for the 10-year period from 2000 to 2009, provided by the Ministry of the Interior and coded as cause-of-death using ICD-10 by the Ministry of Public Health. Multivariate linear regression was used for modeling and forecasting age-specific TB mortality rates in Thailand. Gender differences existed in TB mortality in Thailand with higher deaths occurring in males. TB mortality increased with increasing age for each sex and was also higher in the Central and Northern provinces. The trends of TB mortality decreased in most age groups and remained stable in others during ten-year period (2000 to 2009). The model forecast that the TB mortality will not increase over the 6-year period, and will actually decrease in most of age group and region for both sexes. The multivariate linear regression model can be used as a simpler method for forecasting TB mortality. These findings provide information on forecasting for health authorities to help establish effective prevention programs in specific areas and groups where the TB mortality is relatively high.

**Keywords:** mortality, Tuberculosis, modeling, forecasting, multivariate linear regression, Thailand

---

## INTRODUCTION

Tuberculosis (TB) remains a major cause of infectious disease mortality, with an estimated 8.8 million new cases and 1.6 million deaths annually [1, 2]. It affects in particular low-income countries, with the highest case rates in Africa and the highest case numbers in Asia [1]. The enormous disparity and imbalance of TB globally is due to the underfunding of public health services, the spread of human immunodeficiency virus (HIV), and the emergence of multi-drug resistant TB (MDR-TB) [3]. It is a leading cause of death among people who are HIV-positive. It is also a disease of poverty; it has the greatest impact on youth and adults and has become cause of death among adults [1, 2].

Thailand is one of the 22 WHO-designated high burden TB countries, with an estimated 90,000 TB cases occurring annually. The prevalence of TB was estimated to be 192 per 100,000 population for all forms in 2007, with an incidence rate of 62 new smear-positive cases per 100,000 population and a mortality rate of 19 per 100,000 [1].

The Thailand Health Profile Report indicates that the number of TB cases declined between 1985 and 1989 but increased slightly from 1990 to 2005 due to an explosive HIV epidemic in the 1990s that

resulted in a sudden increase in TB cases [4]. Recent research based on National Tuberculosis Program (NTP) surveillance data from 2001 to 2005 shows that the rate of death from TB was more than two times higher in men than in women. Mortality was highest in persons 65 years and older and in persons aged 35-44 years [5].

Studies on TB mortality and forecasting were reviewed. Recent study compared mortality rates in TB/HIV co-infected individuals globally and by country using multivariate linear regression. The mortality rates of African countries were higher than non-African countries [6]. A study in United States developed multivariate Markov chain model to project TB incidence from 1980 to 2010. The projections of the model demonstrate an intermediate increase in TB incidence followed by continuing decline [7].

Long-term forecasts of mortality and disease burden are essential for setting current and future health system priorities. Better disease forecasting models would help public health officials to enhance the understanding of epidemic patterns in order to prepare for intervention measures in advance.

The objective of our study was to model and forecast TB (all forms) mortality in Thailand.

## MATERIALS AND METHODS

Thailand is an independent country in Southeast Asia. It has four regions (Central, North-east,

---

\* Correspondence to: Chamnein Choonpradub  
E-mail: cchamnein@bunga.pn.psu.ac.th  
Tel: +66 (0)8 9466 0803, Fax: +66 (0) 7331 2729

Northern and Southern) and 76 provinces [8].

Data for registered deaths due to TB were provided by the Bureau of Health Policy and Strategy, Ministry of Public Health. The data were collected from death certificates across the whole country. Deaths certificate are issued by a physician or nurse when death occurs in hospital and by head of village or health personnel when death occurs outside the hospital. This data is entered into the vital registration database that is maintained by Ministry of Interior. It is used by the Ministry of Public Health and coded cause of deaths using International Classification of Disease 10th edition (ICD-10). The resident population denominators used to compute mortality rates were obtained from the Population and Housing Census of 2010 undertaken by the National Statistics Office of Thailand.

Age, gender, residential area by region in Thailand and year were selected as the explanatory variables in studying the mortality rates of TB. Age was divided into seven groups (0-15, 15-24, 25-34, 35-44, 45-54, 55-64 and above 65+ yrs).

For each region and gender combination; multivariate linear regression model was used to investigate and forecast TB mortality by age group and year. This model is expressed as

$$\log(m_{xt}) = a_x + b_x t \quad (1)$$

where  $m_{xt}$  is the central death rate (per 100,000) in age group  $x$  and year  $t$  for the specified gender and geographical region. The mortality rates generally

have positively skewed distributions so it is conventional to transform them by taking logarithms. The factors  $a_x$  and  $b_x$  describe the level and annual increase, respectively of the age-specific mortality rate.

Since some cells or mortality in some age group and year had no reported cases, to allow log-transformation, we replaced zero counts by a suitably chosen small constant, without changing any values of  $m_{xt}$  greater than 0.

The multivariate linear regression estimates the same coefficients as one would obtain using separate univariate regression models [9]. This model has the additional advantage as it takes account of correlations between data in different age groups.

The R program, version 2.10 was used for all statistical analysis and creating graphs [10, 11].

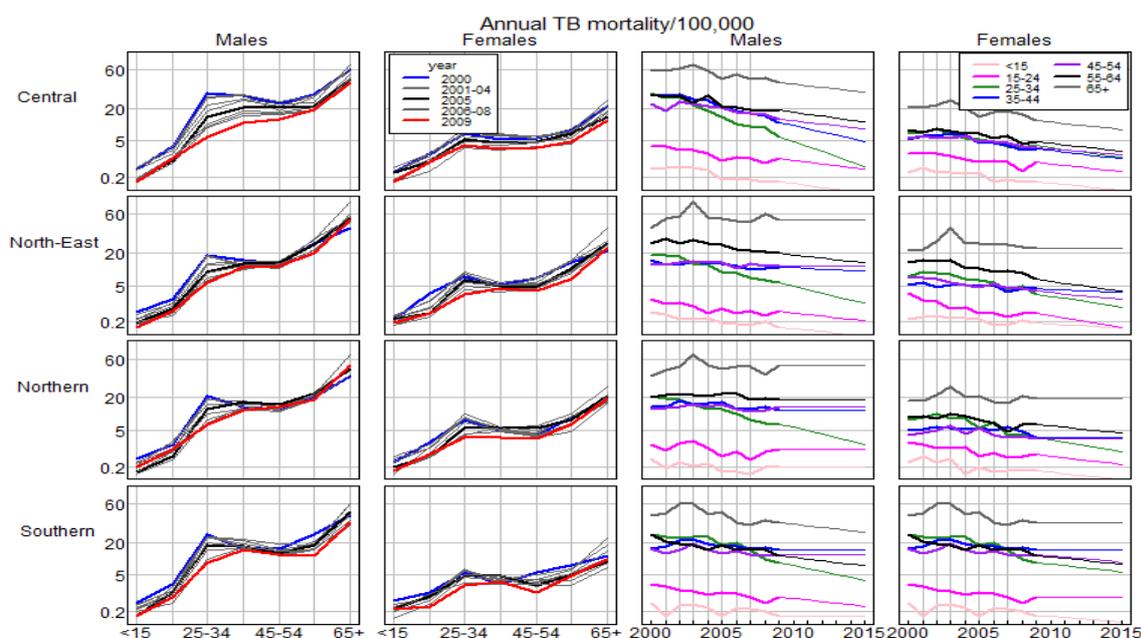
## RESULTS

### Preliminary analysis

A total of 57259 death due to TB or during TB treatment for 10 year period (2000-2009) in Thailand. Among the deaths, 40,565 (70%) were males and 16,694 (30%) were females.

### Statistical analysis

To log-transform the counts, we replaced zeros by 0.5 before fitting the model. The left panels of figure 1 show the TB mortality rates plotted by age group for each year; the right panels show the trends for 2000-2009 for each age group, together with the forecasts based on the model.



**Figure 1** Plot of TB mortality rates by age group for each year (left panels) and trend with forecasts for each age group (right panels) for the four regions of Thailand

Figure 1 shows that mortality increased up to age group 25-34 years and decreased slightly before increasing in the higher age groups for both sexes. In all regions, age groups and years, males had higher mortality than females. In the central region, there was marked bulge in mortality among males between 25 and 55 years of age. However, there were fewer bulges in mortality in other provinces.

The time trends shown in the right panels indicate that the TB mortality decreases in most age groups and remained stable in others over the 10 year period. The model forecast that TB mortality did not increase over the 6-year period, but decreased for 39 and remains steady for 17 of the 56 combinations of age group, gender and region.

The correlation varied substantially between age groups. Highest correlation was observed between 35-44 years and 45-54 years in females of the central region (0.88), followed by 55-64 years and above 65 years in female of north east region (0.69) and 35-44 years and 45-54 years in male in central region (0.66) (data not shown).

## DISCUSSION AND CONCLUSION

In this paper, we applied multivariate linear regression to model and forecast the TB mortality in Thailand. The findings showed that gender differences existed in the mortality of TB; males had higher mortality rates than females in most age groups and regions, as consistent with TB mortality and gender patterns found in studies in Thailand [4, 5]. However, in few age groups, mortality was higher in females than males.

Age has a major impact on TB mortality [12, 13]. In our study, mortality was highest in age groups above 55 years for both sexes. Co-morbidity and decreased immune function are important factors in the increasing TB mortality among the elderly [12, 13].

In the ten-year period (2000 to 2009), mortality decrease in most age groups and remained stable in others. The decreased in TB mortality may be attributed to successful TB control programmes in the country with the expansion of DOTS, case finding and treatment success in the recent years in Thailand [1, 5]. The model demonstrates a gradually decreased in mortality for most combinations of age group, gender and region for 6-year period. Thus, the forecasting suggests that the mortality rates of TB will now continue to decline or remain constant. Although TB mortality decrease and remain constant, the rates are high, indicating that TB is still endemic in most parts of Thailand, making TB a public health problem in Thailand.

There was pronounced bulge in mortality among males between 25 and 55 years of age in the Central region. Possible reasons for the bulge were not investigated in our study, but could be due to

increased HIV incidence during 2001-2004, which attributed to higher incidence of TB [4, 5].

Although the Lee-Carter model is often used for forecasting [14, 15], this non-linear model cannot be fitted by ordinary regression methods, and thus does not routinely provide standard errors for estimated parameters. The multivariate linear regression has the additional advantage is that it takes account of correlations between data in different age groups. We recommend this method as an additional or alternative method to forecast disease mortality as it provide evidence of better disease forecasting.

The study had some limitations. Reliable data on TB mortality are unfortunately not available. Currently available data from the national vital registration and TB register or Hospital register in Thailand are an underestimate, as deaths occurring among defaulters are not included and the diagnosis of the disease is not done by public health people. Besides this, the general mortality from TB also includes deaths due to co-morbidities and other external causes. Ideally, TB mortality should include only deaths due to TB.

Further study is needed to investigate TB mortality in smaller regions, such as provinces or districts for detail investigation of mortality by specific location. The graphical method provides an informative display of the variation in mortality by gender, age group and region. Such graphs can be used by public health authorities for applying preventive measures to control TB mortality.

In conclusion, our study highlights the trend and forecasting of TB mortality by gender, age groups and location illustrated by a graph. These findings require further investigation, to assess the forecasted results with real data, but highlight the importance of selectively monitoring geographic locations and planning future intervention strategies according to prioritized risk groups.

## ACKNOWLEDGEMENTS

We thank the Ministry of Public Health, for providing the data. This study was partially funded by the Graduate School, Prince of Songkla University. We are grateful to Prof. Don McNeil who supervised our research.

## REFERENCES

1. World Health Organization. Tuberculosis control in South East Asia region WHO Annual report 2009. [cited 2011 Jul 17]. Available from: [http://www.searo.who.int/LinkFiles/TB\\_Day\\_Kit\\_3Annual\\_Report.pdf](http://www.searo.who.int/LinkFiles/TB_Day_Kit_3Annual_Report.pdf)
2. van der Werf MJ, Borgdorff MW. Targets for tuberculosis control: how confident can we be about the data? *Bull World Health Organ.* 2007 May; 85(5): 370-6.
3. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: systematic analysis of population

- health data. *Lancet*. 2006 May 27; 367(9524): 1747-57.
4. Wibulpolprasert S. Thailand health profile 2005–2007. Bangkok: The War Veterans Organization of Thailand; 2007.
  5. Jittimane S, Vorasingha J, Mad-asin W, Nateniyom S, Rienthong S, Varma JK. Tuberculosis in Thailand: epidemiology and program performance, 2001-2005. *Int J Infect Dis*. 2009 Jul; 13(4): 436-42.
  6. Au-Yeung C, Kanters S, Ding E, Glaziou P, Anema A, Cooper CL, et al. Tuberculosis mortality in HIV-infected individuals: a cross-national systematic assessment. *Clin Epidemiol*. 2011; 3: 21-9.
  7. Debanne SM, Bielefeld RA, Cauthen GM, Daniel TM, Rowland DY. Multivariate Markovian modeling of tuberculosis: forecast for the United States. *Emerg Infect Dis*. 2000 Mar-Apr; 6(2): 148-57.
  8. World atlas. [online]. [cited 2010 Jul 17]; Available from: <http://www.worldatlas.com/webimage/countrys/asia/th.htm>
  9. Mardia KV, Kent JT, Bibby JM. Multivariate analysis. 2<sup>nd</sup> ed. United States: Academic Press Inc; 1979.
  10. Venables WN, Ripley BM. An introduction to R: notes on R: a programming environment for data analysis and graphics version 2.6.2; 2008.
  11. Murrell P. R graphics. London: Chapman & Hall/CRC; 2006.
  12. Dutt AK, Stead WW. Tuberculosis in the elderly. *Med Clin North Am*. 1993 Nov; 77(6): 1353-68.
  13. Salvadó M, Garcia-Vidal C, Vázquez P, Riera M, Rodríguez-Carballeira M, Martínez-Lacasa J, et al. Mortality of tuberculosis in very old people. *J Am Geriatr Soc*. 2010 Jan; 58(1): 18-22.
  14. Lee RD, Carter LR. Modeling and forecasting U. S. mortality. *J Am Stat Assoc*. 1992; 87(419): 659-71.
  15. Lee R, Miller T. Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*. 2001 Nov; 38(4): 537-49.