SPATIAL AND TEMPORAL VARIATIONS IN TUBERCULOSIS INCIDENCE, NEPAL

Sampurna Kakchapati, Chamnein Choonpradub and Apiradee Lim

Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, Mueang, Pattani, Thailand

Abstract. Tuberculosis (TB) is an important public health problem in Nepal. The aim of this study was to investigate the spatial and temporal variations in TB incidence in Nepal. Data regarding TB cases were obtained from the Nepal National Tuberculosis Center (NTC) for 2003-2010 and analyzed. Models were developed for TB incidence by gender, year and location using linear regression of log-transformed incidence rates. Apart from a relatively small number of outliers, these models provided a good fit, as indicated by residual plots and the *r*-squared statistic (0.94). The overall incidence of TB was 1.31 cases per 1,000 population with a male to female incidence rate ratio of 1.83. There were trends of increasing incidence in TB for recent years among both sexes. There were marked variations by location with higher rates occurring in the Terai region and relatively moderate and low rates of TB in the Hill and Mountain regions, respectively. TB incidence was also higher in the capital city Kathmandu and other metropolitan cities. A log-linear regression model can be used as a simple method to model TB incidence rates that vary by location and year. These findings provide information for health authorities to help establish effective prevention programs in specific areas where the disease burden is relatively high.

Keywords: TB, incidence, spatial variation, temporal variation, log-linear model, Nepal

INTRODUCTION

Tuberculosis (TB) is a major cause of infectious disease mortality, with a worldwide estimated 8.8 million new cases and 1.6 million deaths annually (WHO, 2009a). The global burden of TB is large, and is likely to remain a major public health problem in the coming decades (Lopez *et al*, 2006). The spread of human immunodeficiency virus (HIV/AIDS) and the emergence of multidrug-resistant TB (MDR-TB) are contributing to the worsening impact of this disease. TB is a leading cause of death among people who are HIV-positive (WHO, 2009a). It is a disease of poverty; it has the greatest impact on youth and adults and is a major cause of mortality among adults (WHO, 2009a).

TB is a leading public health problem in Nepal, causing significant morbidity and mortality. Every year approximately 30,000-40,000 people in Nepal develop active TB; among them 20,000 have infectious (smear- positive) cases (National Tuberculosis Center, 2010). These 20,000 are able to spread the disease to others. Although TB

Correspondence: Chamnein Choonpradub, Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, Pattani 94000, Thailand.
Tel 66 (0) 89 4660803; Fax 66 (0) 73 335130
E-mail: cchamnein@bunga.pn.psu.ac.th

is curable, the mortality and incidence are high (Department of Health Services, 2008; National Tuberculosis Center, 2010). TB in Nepal has been worsened by the HIV/AIDs epidemic, population demographics, urbanization, changes in socio-economic standards and, more recently, emerging drug resistant TB that has fueled the high levels of disease in the country (National Tuberculosis Center, 2010).

Several studies and the National Tuberculosis Center (NTC) annual report suggest marked regional differences in the incidence of TB in Nepal (Kakchapati *et al*, 2010; National Tuberculosis Center, 2010). The incidence rate of TB is generally higher in urban areas than in rural areas and in the Terai parts of the country, bordering India (Kakchapati *et al*, 2010; National Tuberculosis Center 2010). TB is a major public health problem in Nepal, especially in the highly populated lowland Terai and in urban cities, including the capital Kathmandu (Kakchapati *et al*, 2010; National Tuberculosis Center 2010).

The application of statistical models to TB was initiated in the 1960s, and since then these statistical models have been used to develop strategies to control TB (Legrand *et al*, 2008), co-infections with HIV and TB (Williams *et al*, 2005) and MDR-TB (Dye *et al*, 1998). Several researchers have studied spatial and temporal variations in TB incidence. Kongchouy *et al* (2010) used log-transformed linear regression to examine trends, seasonal and geographic effects on TB incidence in the fourteen southern provinces of Thailand, and identified the upper western and lower southern parts of Thailand as the region at the highest risk for TB. Kakchapati *et al* (2010) studied TB incidence rates in various districts of Nepal from 2003 to 2008 using a negative binomial model and found spatial variations with higher rates occurring in the Terai region, followed by the Hill, and Mountain regions. Uthman (2008) applied Poisson regression models to determine spatial and temporal variations in TB incidence in Africa and found southern, eastern and central Africa were experiencing an increasing TB incidence and suggested 25 countries were at increased risk of TB.

Public health officials are often required to evaluate disease incidence in a country. They need to compare the standardized disease incidence rate within an area over time to determine what and when necessary actions need to be taken. Thus, understanding their spatial and temporal distribution is essential for further health intervention policies.

MATERIALS AND METHODS

Study area and data source

Nepal, officially the Federal Democratic Republic of Nepal, is a landlocked country in South Asia and is the world's youngest republic. It is bordered to the north by the People's Republic of China, and to the south, east, and west by the Republic of India. It is divided into 14 zones (75 districts), and has a current population of approximately 27.7 million and an annual population growth of 2.2% (World Atlas, 2010). It is divided into three areas: mountain, hill and Terai, in order of decreasing altitude.

Data regarding TB were obtained from the NTC during an 8-year period: 2003-2010. These data are the characteristics of the disease, gender, location and treatment status of TB patients.

The population density in Nepal varies substantially by district; therefore, we computed the TB incidence rates for combine districts called "super-districts",

Table 1
Definitions and populations of super-districts.

Code	Super-district	Population	Code	Super-district	Population
1	Darchula	143,148	33	Rupandehi	857,291
2	Baitadi	273,219	34	Chitwan	568,495
3	Dadeldhura	148,217	35	Makwanpur	657,220
4	Kanchanpur	465,912	36	Parsa	599,199
5	Bajhang+ Bajura	323,202	37	Bara	675,072
6	Doti	244,907	38	Rautahat	654,723
7	Achham	269,504	39	Rasuwa+ Sindupalchowk	409,518
8	Kailali	761,652	40	Dhading	398,915
9	Karnali (zone)	975,186	41	Nuwakot	336,873
10	Dailekh	460,026	42	Kathmandu	1,363,512
11	Jajarkot	156,744	43	Kavre	451,595
12	Surkhet	344,237	44	Bhaktapur	270,107
13	Bardiya	466,702	45	Lalitpur	405,469
14	Banke	250,148	46	Dolkha	238,628
15	Rukum	221,859	47	Ramechhap	244,534
16	Salyan	249,442	48	Sindhuli	331,736
17	Rolpa	245,082	49	Mahottari	465,292
18	Pyuthan	245,503	50	Dhanusha	793,609
19	Dang	264,616	51	Sarlahi	759,631
20	Mustang+ Myagdi	149,599	52	Solukhumbu+ Okhaldunga	304,969
21	Baglung	312,830	53	Khotang	264,129
22	Parbat	181,277	54	Udaypur	344,588
23	Manang+ Lamjung	218,294	55	Siraha	677,957
24	Gorkha	334,022	56	Saptari	673,056
25	Kaski	455,559	57	Sankhuwasabha+ Tehrathum	313,791
26	Syangja	362,929	58	Bhojpur	228,983
27	Tanahu	368,194	59	Dhankuta	192,889
28	Gulmi	341,828	60	Sunsari	756,321
29	Arghakhanchi	242,159	61	Morang	1,000,114
30	Palpa	311,021	62	Taplejung+ Panchthar	390,466
31	Nawalparasi	672,760	63	Illam	334,376
32	Kapilvastu	576,769	64	Jhapa	801,041

defined as regions comprising contiguous districts in the same zone with a 2010 total population greater than 140,000, (Table 1). We divided the country into 64 super-districts listed in geographical order from west to east (keeping districts within the same zone together).

Gender, residential area (by super-district) and year were selected as explanatory variables in studying the incidence rates of TB. Gender and year were combined together to form sixteen gender-year groups. Age was not recorded.

Statistical methods

The additive log-linear model used for incidence rates with normally distributed

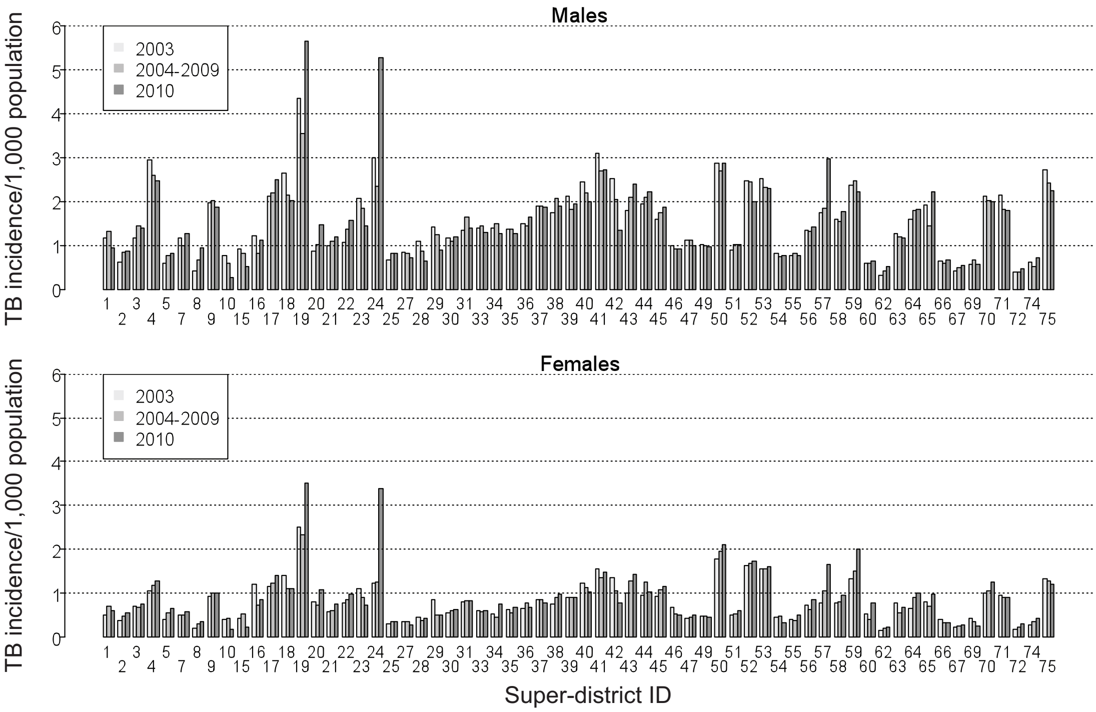

Fig 1–TB incidence/1,000 population in Nepal by super-district 2003-2010.

errors was:

$$\ln\left(\frac{n_{ij}}{P_i}\right) = y_{ij} = \mu + \alpha_i + \beta_i \qquad (1)$$

In this model, P_i is the corresponding population at risk in 1000s and the terms α_i and β_i represents the super-district and gender-year effects that sum to zero so that μ is a constant encapsulating the overall incidence. The model fit is then assessed by the linearity in the plot of deviance residuals against normal quantiles.

The model also gives adjusted incidence rates for each factor of interest, obtained by replacing the parameters corresponding to the other factors by constants chosen to ensure the total expected number of cases equals the observed number.

Sum contrasts (Venable and Ripley, 2002) were used to obtain confidence intervals for comparing the adjusted incidence rates within each factor with the overall incidence rate. Since the confidence intervals for factor-specific incidence rates obtained from the model divide naturally into three groups according to their location entirely above the mean, around the mean, or entirely below the mean, we used this trichotomy to create schematic maps of super-districts according to their estimated TB incidence rates.

The R program was used for all statistical analysis, graphs and maps (R Development Core Team, 2008).

RESULTS

Preliminary analysis

A total of 271,873 TB cases were reported to the NTC during the eight-year studied period (2003- 2010); 175,365 were males and 96,508 were females. Fig 1 shows the distribution of TB in Nepal by super-

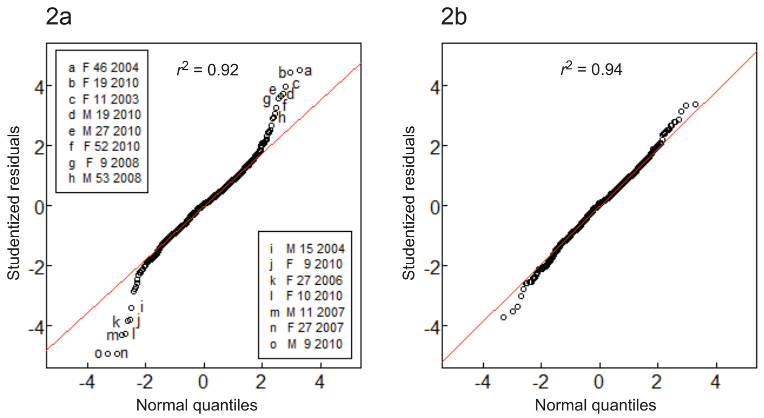

Fig 2–Diagnostic residual plots for log linear model before (left) and after (right) omitting outliers. Key to Fig 2a: F-female, M-male, the third column is the district ID and the fourth column is the year.

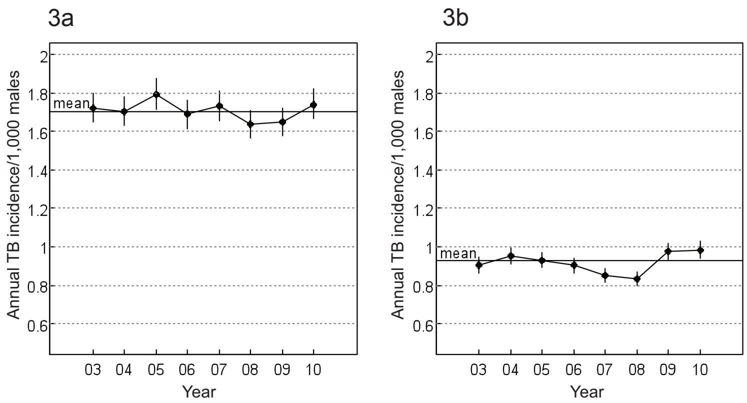

Fig 3–Annual TB incidence/1,000 population by year.

district in order of increasing incidence rate. Super-districts 53 and 14 had the lowest and highest TB incidence, respectively. Males had a proportionately higher incidence rates than females.

Statistical analysis

While fitting the model, we detected outliers in the data. The residual plots before (Fig 2a) and after (Fig 2b) omitting 13 residuals greater than 3 in magnitude are shown in Fig 2. Seven of these outliers occurred in 2010. The omission of outliers makes the normality assumption plausible.

Fig 3 shows 95% confidence intervals for TB incidence rates by year for males (Fig 3a) and females (Fig 3b). The mean TB incidence among males was 1.70/1,000 population. The mean TB incidence among females was 0.92/1000 population. Among both males and females there was an overall decline in TB incidence between 2005 and 2008, followed by a sharp increase in TB incidence between 2008 and 2010.

Fig 4 shows 95% confidence intervals for annual TB incidence rates by super-district, and regions. The mean incidence of TB was 1.3/1,000 population. Super-districts with confidence intervals above the mean were categorized as having a higher than average incidence, super-districts with confidence intervals below the mean were categorized as having a lower than average incidence and super-districts with confidence intervals overlapping the mean were categorized as having an average incidence.

A high incidence of TB was seen in the Terai region (Fig 5).

DISCUSSION

We studied spatial and temporal variations in TB incidence for Nepal from 2003 to 2010. Linear regression models containing gender, year and super-districts as factors

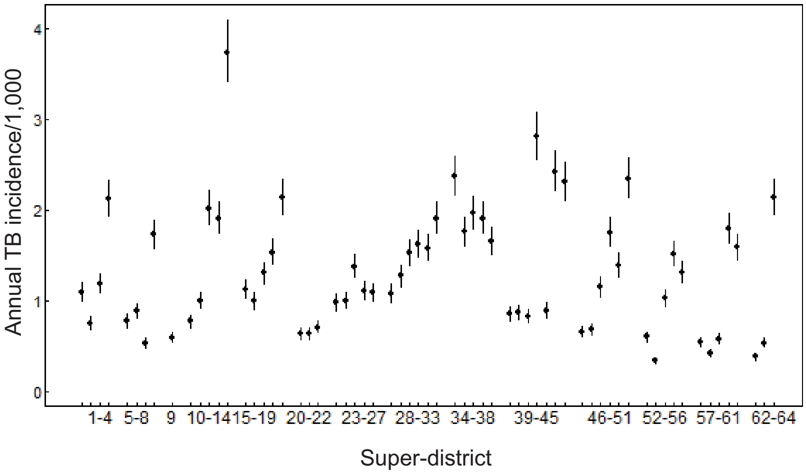

Fig 4–Annual TB incidence/1,000 by super-district.

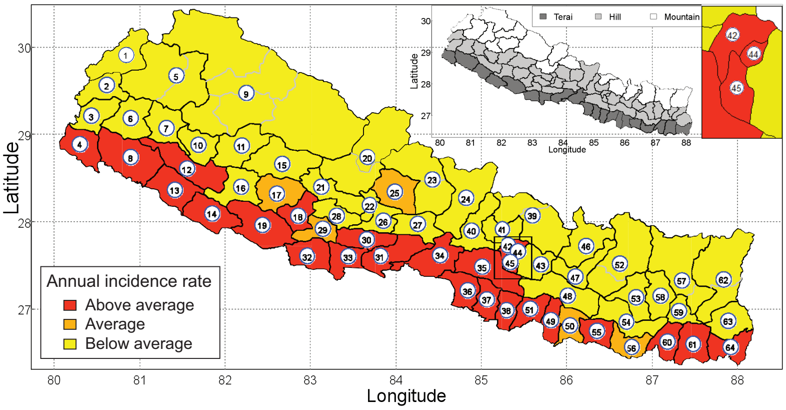

Fig 5–Map of annual TB incidence in Nepal, 2003-2010.

were fitted to the log-transformed disease incidences. After omitting 13 outliers, this model provided an acceptable fit with an adjusted *r*-square of 0.94. The overall incidence of TB was found to be 1.31 cases per 1,000 population. There was a marked gender difference in the incidence of TB: males had a higher incidence rate than females. This is consistent with the findings of a recent study from Nepal (Kakchapati *et al*, 2010) and from other countries (Borgdorff *et al*, 2000; Martinez *et al*, 2000). Epidemiological studies have found gender differences in TB incidence begin to appear between age 10 and 16 years and remain higher among males than females thereafter (Uplekar *et al*, 2001; WHO, 2009b).

The incidence of TB in Nepal increased sharply during 2008-2010. The reasons for the increase were not investigated in the study. It could be due to an increase in HIV infection, MDR-TB or the emergence of extensively drug-resistant TB (National Tuberculosis Center, 2010). Further research is needed to identify factors responsible for this increase. TB was distributed unevenly in Nepal: higher incidence rates were found in the Terai regions which could be attributed to medical, social and environmental factors. The Terai region has higher temperatures, a lower socio-economic status, a high population, malnutrition, more poverty and social deprivation, all contributing to TB infection (National Tuberculosis Center, 2010). The greater number of TB cases in Terai may have been influenced by people from neighboring India crossing the border into Nepal seeking treatment. India occupies the first position on the World Health Organization's list of high TB burden countries (WHO, 2009a).

TB incidence is higher among densely populated metropolitan cities, such as Kathmandu, Banke, Bhakthapur, Chitwan, Dang, Jhapa and others. The highest incidence was in Banke followed by the capital city Kathmandu. Several studies have found urban centers have higher TB rates than rural areas (Tupasi *et al*, 2000; Barr *et al*, 2001). Population density, migration, poverty and overcrowding appear to be major factors for disease transmission in these areas (Acevedo-Garcia, 2001).

There were several limitations in our study. It was based on secondary data; therefore, we could not include some risk factors for TB, such as age, due to unavailability of age-specific incidence data. Additional analyses are needed to evaluate the trends of TB using data over a longer study period, or more detailed incidence data (monthly, quarterly). However, our findings are illustrated by the map showing the districts with higher incidence rates. Such maps can be used by public health authorities to prioritize preventive measures to control TB outbreaks by focusing on zones with a high or increasing incidence of TB.

ACKNOWLEDGEMENTS

This study was partially funded by the Graduate School, Prince of Songkla University, Thailand. We would like to express our gratitude to the NTC, Thimi, Bhaktapur, for permission to use their data. We are indebted to Prof Don McNeil for supervising our research.

REFERENCES

Acevedo-Garcia D. Zip-code level risk factors for tuberculosis: neighborhood environment and residential segregation in New Jersey, 1985-1992. *Am J Public Health* 2001; 91: 734-41.

Barr RG, Diez-Roux AV, Knirsch CA, Pablos-Mendez A. Neighborhood poverty and the resurgence of tuberculosis in New York City, 1984-1992. *Am J Public Health* 2001; 91: 1487-93.

Borgdorff MW, Nagelkerke NJD, Dye C, Nunn P. Gender and tuberculosis: a comparison of prevalence surveys with notification data to explore sex differences in case detection. *Int J Tuberc Lung Dis* 2000; 4: 123-32.

Dye C, Garnett GP, Sleeman K, Williams BG. Prospects for worldwide tuberculosis control under the WHO DOTS strategy: Directly Observed Short Course Therapy. *Lancet* 1998; 352: 1886-91.

Department of Health Services (DoHS). Annual report (2066-2067). Kathmandu: Government of Nepal, 2008.

Kakchapati S, Yotthanoo S, Choonpradupb C. Modeling tuberculosis incidence in Nepal. *Asian Biomed* 2010; 4: 355-60.

Kongchouy N, Kakchapati S, Choonpradub C. Modeling the incidence of tuberculosis in southern Thailand. *Southeast Asian J Trop Med Public Health* 2010; 48: 574-82.

Lopez AD, Mathers CD, Ezzati M, Murray CJ. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006; 367: 1747-57.

Legrand J, Sanchez A, Le Pont E, *et al*. Modeling the impact of tuberculosis control strategies in highly endemic overcrowded prisons. *PLoS One* 2008; 3: e2100.

Martinez AN, Rhee JT, Small PM, Behr MA. Sex differences in the epidemiology of tuberculosis in San Francisco. *Int J Tuberc Lung Dis* 2000; 4: 26-31.

National Tuberculosis Center (NTC). National Tuberculosis Control Programme, annual report (2066-2067). Bhaktapur: Government of Nepal, 2010.

R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2008. [Cited 2008 Feb 8]. Available from: URL: http://www.R-project.org

Tupasi TE, Radhakrishna S, Quelapio MID, *et al*. Tuberculosis in the urban poor settlements in the Philippines. *Int J Tuberc Lung Dis* 2000; 4: 4-11.

Uplekar MW, Rangan S, Weiss MG, Ogden J, Borgdorff MW, Hudelson P. Attention to gender issues in tuberculosis control. *Int J Tuberc Lung Dis* 2001; 5: 220-4.

Uthman OA. Spatial and temporal variations in incidence of tuberculosis in Africa, 1991 to 2005. *World Health Popul* 2008; 10: 5-15.

Venables WN, Ripley BD. Modern applied statistics with S. 4th ed. New York: Springer-Verlag, 2002: 139-208.

Williams BG, Granich R, Chauhan LS, Dharmshaktu NS, Dye C. The impact of HIV AIDS on the control of tuberculosis in India. *PNAS* 2005; 102: 9619-24.

World Health Organization (WHO). Tuberculosis control in the South East Asia Region:WHO annual report 2009. Geneva: WHO, 2009a.

World Health Organization (WHO). Global tuberculosis control 2009: epidemiology, strategy, financing: WHO report 2009. Geneva: WHO, 2009b.

World Atlas (on line). [Cited 2010 Jul 17]. Available from: URL: http://www.worldatlas.com/webimage/countrys/asia/np.htm