UNIVERSITY OF NAIROBI

"A BAYESIAN APPROACH TO THE SPATIAL ANALYSIS OF TUBERCULOSIS TREATMENT OUTCOMES IN KENYA"

DR. ELIZABETH NAUKUSI WANGIA

REG NO: W62/69035/2011

UNIVERSITY OF NAIROBI INSTITUTE OF TROPICAL AND INFECTIOUS DISEASES (UNITID).

A Research Project Submitted In Partial Fulfilment of the Requirements For The Award of Master of Science in Medical Statistics.

SEPTEMBER 2015

DECLARATION

I, **ELIZABETH NAUKUSI WANGIA REG.NO. W62/69035/2011** hereby declare that the work on which this project is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature …….

Date …….

APPROVAL

This project paper has been submitted for review with our authority as university supervisors:

Dr Thomas Achia

Signature ……..

Date ……..

Anne Wangombe

Senior Lecturer

University of Nairobi

Institute of Tropical and Infectious Diseases

Signature ……..

Date …….

DEDICATION

This project is dedicated to my dear husband Charles Wilson Okola who has supported me, throughout my studies and my children Hellen, Lakeesha and Nathan who believed in me and never failed to pray for me. I thank my father Eng. Jared Waudo Wangia for nurturing my love for Mathematics since childhood.

ACKNOWLEDGEMENT

I would like to acknowledge God, for giving me the strength, hope and determination.

Supervisors Dr. Thomas Achia and Anne Wangombe, for their guidance.

The Ministry of Health for sponsoring this course.

Thank you and God bless you.

ABSTRACT

Introduction: Tuberculosis is second only to HIV as the greater killer worldwide due to a single infectious agent. Improving the treatment outcome of tuberculosis is part of the Millennium Development Goals. Given the infectious nature of tuberculosis, its distribution and treatment outcomes should consider spatial patterning. Information on the distribution of tuberculosis treatment outcomes in Kenya is scarce, yet treatment outcome is an important indicator of tuberculosis management. Spatial analysis tools can be used to characterize spatial patterns of these treatment outcomes, thereby identifying areas at risk of the given outcomes. This study examined the spatial distribution of the tuberculosis treatment outcomes across the counties in Kenya.

Objective: To model the spatial distribution of tuberculosis treatment outcomes using Bayesian techniques.

Methods: Study area was Kenya, a country in the East Africa region. Secondary data was obtained from the national tuberculosis registers from January 2014 to March 2014 with incorporation of data from the Kenya Demographic and Health Survey 2014 and Census 2009. Treatment outcomes were categorized as cured, dead, defaulted, failure and treatment complete. Exploratory data analysis was done to estimate the proportions of the various covariates, and tests for global and local spatial auto correlation done to assess the relationship of the various outcomes per county. Covariates were selected using purposeful selection of variables, and variables with a significant univariate test were selected as candidates for the multivariate analysis. Augmentation of the linear predictors with a set of spatially correlated random effects was done, using conditional autoregressive prior distributions, specified by a set univariate full conditional distributions. Inference was based on obtaining the posterior distribution, of the different TB treatment outcomes, using the Integrated Nested Laplace Approximation Methodology (INLA) as a way of approximating the posterior marginals as proposed by Besag et al.

Results: A total number of 23,488 records were analysed comprising of 60.38% male, 70.32% patients between the age of 15-45, patients with pulmonary tuberculosis at 82.30%, and HIV positive patients were 59.03%. There was significant global spatial autocorrelation seen for the patients who were cured, failed treatment, died and those who completed treatment. However, most of the covariates did not show significant

spatial dependence. The fitted data also showed uniform distribution of the outcomes of tuberculosis treatment across the counties with occasional high risk spots.

Conclusion: The spatial effect of the tuberculosis treatment outcomes appeared weak across the various counties. This may imply that appropriate risk factors were adjusted for in the model such that the spatial random effect became less important.

Recommendation: Future related studies involving various TB outcomes should be traced at household level to minimize mismatch between risk factors and the TB outcomes.

Keywords: Bayesian statistics, Conditional Autoregressive Models, Tuberculosis treatment outcome

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

MDR TB	Multidrug resistant tuberculosis
WHO	World Health Organization
NTLD	National Tuberculosis and Lung Disease Program
SMR	Standardized Mortality Ratio
CAR	Conditional Autoregressive Model
INLA	Integrated Nested Laplace Approximation

CHAPTER ONE

1. INTRODUCTION

1.1 Background

The relationship between space and health dates back to Hippocrates who stated that "airs, waters, places" all played significant roles impacting human health and history (Derek Gregory, 2009). Concentration of a disease or outcome in a given area implies unusual presence of factors that cause that disease and spatial localization can provide hints as to why the disease occurs in that particular geographical region. Spatial analysis tools are used to characterize spatial patterns of diseases, thereby facilitating cost effective targeting of intervention measures by visualization and exploration of disease patterns.

Tuberculosis (TB) is an infectious disease caused by a bacillus belonging to a group of bacteria grouped in the Mycobacterium tuberculosis complex. World Health Organization estimates one third of the humans on earth are infected with TB, but are asymptomatic and cannot transmit it to others, with a 10 percent chance that their infection will develop into TB (Hayes, 2013). People infected with TB bacteria have a 10% lifetime risk of falling ill with TB, but in the presence of compromised immune systems (as with people living with HIV, malnutrition or diabetes, or people who use tobacco) the risk of TB is much higher.

TB is second only to HIV/AIDS as the greatest killer worldwide due to a single infectious agent (WHO, 2015). In 1993, the World Health Organization (WHO) declared tuberculosis a global emergency. Improving TB treatment outcomes was part of the Millennium Development Goals. There has been continued occurrence of high TB mortality despite effective treatment. 95% of TB deaths occur in low and middle income countries (WHO, Global Tuberculosis Report, 2015). In 2010, Africa alone contributed 26% of the global burden with nine out of the 22 high burden countries contributing 81% of the global burden coming from Africa. During this time, Kenya was ranked position 10 amongst the high burden countries (WHO, Global Tuberculosis Report, 2011). TB is the leading killer of HIV patients accounting for 25% of deaths. TB has a great impact on quality of life, contributing to a high estimate of disability adjusted life years (DALY) at 706.9/100,000, age standardized (Christopher J L Murray, 2015). One disability adjusted life year

(DALY) can be thought of as one lost year of healthy life. The sum of these DALYs across the population, which is also known as the burden of disease, is the measure of the gap between current health status and an ideal health situation where the entire population lives to an advanced age, free of disease and disability (WHO, Health statistics and information systems, 2015). Because of the long duration and associated side effects of standard TB drug treatment, patients often do not complete the full course of therapy. This fosters emergence of single- and multi-drug resistant TB (MDR-TB) strains, which has made diagnosis difficult and costly.

1.2 Statement of the problem

Information on the spatial distribution of tuberculosis treatment outcomes in Kenya is scarce, yet treatment outcome is an important indicator of the tuberculosis control program. Spatial analysis tools can be used to characterize spatial patterns of the various outcomes of TB, thereby identifying areas at risk of unfavourable treatment outcome (failure, death and default), and areas with successful TB outcomes can have their success strategies emulated.

1.3 Objective

The specific objective of this study is

To use Bayesian technique to determine the spatial distribution of TB treatment outcomes.

1.4 Justification

Most studies on TB outcome have used non-spatial methods to analyse and assess various factors associated with the disease (Bruce J. Kirenga, 2014) (Daniel Chukwunweolu Oshi, 2014) (Gebretsadik Berhe, 2012) (Belete Getahuna, 2013) (Cui Hua Liu, 2011). Given the infectious nature of tuberculosis, analysis of the determinants of tuberculosis, should consider spatial patterning.

1.5 Hypothesis

The primary hypothesis in this study is:

Treatment outcomes of TB show no significant spatial distribution.

CHAPTER TWO

2. LITERATURE REVIEW

2.1 History of Tuberculosis

It has been hypothesized that the genus Mycobacterium originated more than 150 million years ago (Hayman, 1984), with archaeological evidence in Egypt, documented more than 5000 years ago, with typical skeletal abnormalities of tuberculosis, including characteristic Pott's deformities (Zimmerman, 1979) (Cave, 1939), (D. Morse, 1964), and in America (Daniel T. , 2000). Hippocrates understood that TB attacked mainly people between the age of eighteen and thirty-five (Hippocrates, 1982). Mycobacterium tuberculosis was originally isolated by Robert Koch in 1882. Chemotherapeutic agents were discovered which revolutionized TB management between 1943 and 1957 (A. Schatz, 1944) (Daniel T. M., 2006).

2.2 Incidence

An estimated two billion people worldwide are infected with Mycobacterium tuberculosis (CDC, 2006) and approximately two million die from TB annually (Dermot Maher, Global epidemiology of tuberculosis, 2005). In 1993, the WHO declared tuberculosis a global emergency because of the scale of the epidemic and the urgent need to improve global tuberculosis control. (Dermot Maher, Global epidemiology of tuberculosis, 2005). TB incidence in Kenya is uncertain and therefore not measurable but estimated (Borgd, 2004). The current estimated TB incidence in Kenya stands at 109/100,000 for those with TB and HIV, and 268/100,000 for the TB cases with and without HIV with a case detection rate of 75%. Mortality rate from TB with HIV is 21/100,000, while TB without HIV is 20/100,000 (WHO, Global Tuberculosis Report, 2015). Notifications of TB cases provides a good indicator of TB incidence in countries that have both high-performance surveillance systems with little underreporting of diagnosed cases and access to health care means that few cases are not diagnosed. Prevalence rate for TB in Kenya stands at 299/100,000 population, (DLTLD, TB Treatment Guidelines, 2013).

TB is a major cause of morbidity and mortality in Kenya, affecting all age groups, but has its greatest toll in the most productive age group of 15 to 44 years, a fact that has not changed much since the time of Hippocrates (Hippocrates, 1982).

Prevalence is much higher among men than women. The sex ratio (M: F) in Kenya is 1.4:1, and most of the notified cases were adults. (WHO, Global Tuberculosis Report, 2015)

Currently, Kenya is ranked 15th among the 22 high burden countries that collectively contribute about 80% of the world's TB cases (DLTLD, TB Treatment Guidelines, 2013). In 2013, Kenya reported a case detection rate of 75% (WHO, Global Tuberculosis Report, 2015), compared to the global case detection rate estimated at 61% in 2008. Current WHO target for case detection rates and cure rates, stand at 70% and 80% respectively.

2.3 Natural history of TB

Infection refers to the presence of TB bacilli in the body without any clinical signs and symptoms of the disease. During the first 2 years after infection, people with TB infection are at high risk of developing TB disease. When an infected person coughs, droplet nuclei containing tubercle bacilli are inhaled, then they enter the lungs and travel to the alveoli. In the alveoli, they multiply, with a small number of tubercle bacilli entering the bloodstream, thus spreading throughout the body. Within 2-10 weeks the immune system produces macrophages that surround the tubercle bacilli. These cells form a hard shell that keeps the bacilli contained and under control. If the immune system cannot keep the bacilli under control, the bacilli begin to multiply rapidly moving from infection to TB disease.

If TB disease is left untreated, 50-60% of the cases die, 20-25% are spontaneously cured and 20-25% of the cases remain as chronic coughers within 5 years.

In studies of the natural history of the disease among sputum smear-positive and HIV-negative cases of pulmonary TB, around 70% died within 10 years; among culture-positive (but smear-negative) cases, 20% died within 10 years. The duration of tuberculosis from onset to cure or death is approximately 3 years and appears to be similar for smear-positive and smear-negative tuberculosis (Edine W. Tiemersma, 2011).

2.4 Kenyan progress

Data on TB prevalence and mortality in Kenya is sparse, since there has not been a national TB prevalence survey since 1956. There are plans to carry out a prevalence

survey in 2015. These results would provide data on the current status of the TB burden and the effectiveness of TB control interventions.

However Kenya has made tremendous progress in TB management, through the National Tuberculosis and Lung Disease (NTLD) program by reversing the upward trend of the burden of TB disease evidenced by a sustained decline in prevalence, incidence and mortality (DLTLD, Annual Report, 2012). Despite all the progress made, it is estimated that 9,500 (5,400-15,000) deaths in Kenya are due to TB making it the fourth leading cause of mortality in the country.

The government is working with its development and technical partners, community-based organizations (CBOs), non-governmental organizations (NGOs), Stop TB Partnership, and the private sector (USAID, Nutrition and Tuberculosis: a review of the literature and considerations for TB control programs, 2008) to ensure a TB free nation.

With an already declining rate of TB case notification, the government, through the NTLD program, is building on the program's solid foundation with a focus on the prevention of transmission, by actively engaging the private health actors, at both national and county levels, to support interventions covered within its plan. This National Strategic Plan (NSP) for TB, leprosy and lung disease represents a transition to program implementation through the newly established 47 counties and intentional acceleration of declining incidence

2.5 Challenges

The emergence of drug resistant TB is of increasing concern, as it poses major challenges in the fight against TB in resource limited countries like Kenya. Treatment of drug resistant TB is expensive, prolonged and associated with poor treatment outcomes compared with drug susceptible TB (NTLD, 2013). HIV has contributed to the resurgence of TB in Africa (Laurent X. Nouvel, 2006).

2.6 Variables

Patient related variables

Sex of an individual was seen to affect development of TB, with males having a higher prevalence compared to women. People aged between 15 and 45 years, had a higher chance of developing TB (Hippocrates, 1982). Those who were more likely to

die of TB were the males, individuals who had extremes of age and immunosuppression (Tuula Vasankari, 2007) (MacIntyre C. R., 1997).

Malnutrition was observed in over half of the TB patients at the onset of treatment, with 17% being severely malnourished and a further 22% being moderately malnourished (USAID, Nutrition and Tuberculosis: a review of the literature and considerations for TB control programs, 2008). Nutritional status goes hand in hand with immunity. In this case, the Body Mass Index (BMI) was used to determine the nutritional status. An underweight individual could have poor nutrition, and poor immunity.

Household related variables

Crowded living conditions were seen to increase the rate of TB spread through increased contact between infectious and susceptible individuals (Padmanesan Narasimhan, 2013). Poverty which is associated with unsatisfied basic needs led to a higher risk of nonadherence (Goodarz Kolifarhood, 2015). Individuals living in the outskirts of towns were shown to have higher spatial clusters of TB (Guy Harling, 2014) (Dye C) (María Belén Herrero, 2015).

Low levels of education, limited social support and high unemployment, have been found to increase the chances of getting TB and reduce adherence to treatment (Bhatti, 1995) (Parslow, 2001) (Spence, 1993) (K Tocque, 1999) (Melo, 2012) & (Penna, 2009).

Ecological variables

These are the factors, when observed at a group level, influence development of an outcome of a disease. For example proximity to water body, climate of a region and pollution. The effects of these factors cannot be quantified, and therefore are modelled as random effects.

Developing a better appreciation of the importance of these variables would enable a more informed understanding of the factors affecting TB treatment outcomes. A number of previous investigations of the spatial distribution of TB have been conducted within the context of socioeconomic factors, at various levels of aggregation. City-based studies have found tuberculosis to cluster around drinking establishments (Munch, 2003), in high-deprivation areas (Alvarez-Hernandez, 2010), (Randremanana, 2009), amongst migrant populations (Jia, 2008) and (Li, 2011), and

in areas with high migrant levels (Kistemann, 2002). Marginalized populations including prisoners have a higher chance of getting infected with TB (J. O'Grady, 2011).

Other covariates that determine TB treatment outcomes but could not be assessed in this study were substance abuse which was a strong predictor of increased prevalence, non-adherence and default (Gelmanova, 2007) (MacIntyre C. R., 1997), and alcoholism and late presentation which contributed substantially to mortality (Guy Harling, 2014)

Hospitalized adherent patients on treatment were seen to have a higher incidence of multidrug-resistant TB raising the possibility that treatment for drug-sensitive disease unmasks a pre-existing population of drug-resistant TB, or that these patients were re-infected with a drug-resistant strain of TB (Gelmanova, 2007).

2.7 Spatial Analysis

GIS and Spatial Scan Statistics have been used to describe the spatial distribution of diseases (Naus, 1965) (Cromley & McLafferty, 2002) (Knox, 1989) (Kullford & Nagarwalla, 1995) (Julian Besag, 1991). Detection of clusters has been useful in the regular surveillance of diseases, by identifying factors behind the spread of the diseases (Tango & Takahashi, 2005). Use of Spatial Scan Statistics is a common method for identification and interpretation of clusters (Kulldorff, Gregory, Samociuk, & DeChello, 2006) (Chen, Roth, Naito, Lengerich, & MacEachren, 2008) . Combination of GIS and spatial analysis has been used to analyse the spatial distribution of diseases in endemic areas, thereby identifying clusters (José Wilton Queiroz, 2010) (Fischer E, 2008) (Bakker MI H. M., 2006) (Hoeven TA, 2008).

Spatial analysis using Bayesian models

Bayesian inference is a process of learning from data, and is based on computing the posterior distribution of a vector of model parameters x conditioned on the vector of observed data y (Bivand, Gomez-Rubio, & Rue, 2015).

$$\pi\ \boldsymbol{x}\ \boldsymbol{y}\ \propto \pi\ \boldsymbol{x}\ \pi\ \boldsymbol{y}\ \boldsymbol{x} \qquad (1)$$

Bayesian inference has become very popular in spatial statistics in recent years, due to the availability of computation methods to fit spatial models. Bayesian methods have been used by Archie C et al, (Archie C. A. Clements, 2006) to predict the spatial distributions of Schistosoma haematobium and S. mansoni infections in Tanzania.

Integrated Nested Laplace Approximation (INLA) enables fitting of a large range of complex statistical models by reducing computation time. It focuses on the posterior marginals for latent Gaussian models, therefore it is useful when only marginal inference on the model parameters is needed (Rue, Martino, & Chopin, 2009). INLA allows different forms for the likelihood of the observations. Prior distributions $\pi(\theta)$ are also very flexible, depending on the latent effect.

CHAPTER THREE

3. METHODOLOGY

3.1 Setting and study design

The study area was carried out in Kenya, a country in East Africa, with an area of 581,313.2 Km^2 and a population of 38,610,097 (KNBS, 2009). Kenya is divided into 47 counties, which are geographical units representing a devolved government. The observation unit considered in this study was the county. TB in Kenya is managed by the NTLD program, which oversees diagnosis, treatment and reporting of patients with TB. Suspects undergo sputum smear microscopy and culture at the time of diagnosis. Those who are culture-positive also undergo drug sensitivity testing to isoniazid, rifampicin, ethambutol, streptomycin and kanamycin. Susceptibility is determined using the absolute concentration method on Lowenstein-Jensen medium, based on the following drug concentrations: isoniazid 1 mg/ml, rifampicin 40 mg/ml, ethambutol 5 mg/ml and streptomycin 10 mg/ml. Patients diagnosed with active TB are treated according to WHO recommendations (WHO, Global Tuberculosis Report, 2015). Those with multidrug resistant TB (MDR-TB) are switched to an individualized regimen based on the drug resistance profile. Patients undergoing TB treatment are assessed with repeat sputum smear, culture and drug-sensitivity testing in months 2, 3 and 5 as well as at the end of treatment and at six-month intervals thereafter. This information is kept in case registration books.

This study was a retrospective study of the newly detected TB cases registered between January 2014 and March 2014.

Figure 1: Map of Kenya by the counties

3.2 Data collection

For the current study, we used secondary data collected from an existing patient database from the national TB registers with permission from the NTLD-P. NTLD-P aims at accelerating the reduction of TB burden through provision of people-centred, universally accessible, acceptable and affordable quality services. The following information had been collected routinely, at county level, for all patients undergoing TB therapy under the NTLD-P: the demographic data (age, sex, county, sector, weight, height, BMI) and clinical information (type of TB, type of patient, regimen given, HIV status and outcome of treatment).

To investigate determinants of treatment outcome of TB, we included data from the Kenya Demographic and Health Survey (KDHS) 2014 and the Kenya Population and Housing Census 2009. KDHS provides information to monitor and evaluate population and health status in Kenya. The Kenya National Bureau of Statistics conducts this survey every five years. This was the most recent survey conducted, and was the first to provide county information. Details on how KDHS 2014 sampling, data collection, data quality control and analysis was done is available in the KDHS 2014 report (KNBS, Kenya Demographic and Health Survey, 2014). Each TB outcome was matched with the demographic characteristics, county information and socioeconomic characteristics.

The dataset comprised 23488 individuals who registered at the TB clinics across the country in the first quarter of 2014.

3.3 Ethics statement

This study used secondary data and did not involve any experiment or interaction with human or animal subjects. The process of testing for specimen collection and testing for TB is approved by Kenyatta National Hospital Ethics and Research Committee and Kenya Medical Research Institute (KEMRI) national ethical review committee.

3.4 Variable Selection

Dependent variable. This is the outcome of treatment of a patient enrolled at a TB centre. These outcomes were either cured, completed treatment, defaulted, died or

had treatment failure. There were other outcomes like transferred out and on treatment that were not considered.

Covariates. Tuberculosis is an airborne disease and therefore its risk factors will mainly be associated with how close people are to each other. This applies to people living in crowded conditions, where poverty prevails. Factors determining TB incidence and specific treatment outcomes were identified from the literature, which integrated ecological, socioeconomic, and biological variables in the analysis. This is due to the fact that these determinants operate through a common set of either indirect or direct variables that impact on TB treatment outcome. The observed variables used were: patient specific, clinical variables and household related.

The patient related variables were age, sex, type of TB and nutritional status represented by the patients BMI. Household related factors were economic status, literacy of the patient and residential areas in terms of rural versus urban. Clinical related variables were type of TB, type of patient (new or retreatment), type of health facility patient attended, HIV status and cotrimoxazole use for the HIV positive patients. Ecological variables are the factors observed at a group level that influence development and outcome of a disease, for example proximity to water bodies, climate of a region and pollution. These variables enabled a more informed understanding of the factors affecting TB treatment outcomes. These guided the selection of covariates. From the KDHS 2014, we picked the national TB prevalence by the counties, literacy levels, poverty levels and residence as either urban or rural. From the Census 2009, we picked the population per county, and from the data from NTLD-P, we included the age, sex, Body Mass Index (BMI), type of TB, type of patient as newly diagnosed or repeat patient, sector of treatment (private, public, prisons or faith based organizations) and the HIV status of the patient. Cotrimoxazole therapy was assessed among the patients who were HIV positive. Counties of residence were included to look for any form of clustering of the given outcomes and for assessment of any form of spatial relationship with neighbouring counties.

Since inclusion of all clinical and other relevant variables in the model regardless of their significance in order to control for confounding may lead to numerically unstable estimates and large standard errors (Zoran Bursac, 2008), purposeful selection of variables was done which began with a univariate analysis of each variable using a generalized linear model (GLM). This univariate model was used to

test for the association of each single covariate with the outcome variable. The outcome variable was in form of proportions. GLM allowed us to model our data using other distributions than the Normal (Olsson, 2002). The general linear model is as indicated in equation (2), while the generalized model is in equation (3).

$$\mathbf{y} = \mathbf{X}\beta + \mathrm{e} \qquad (2)$$

where $\boldsymbol{\eta} = \boldsymbol{X}\beta$

$$\mathrm{g}\ \boldsymbol{\mu} = \boldsymbol{\eta} = \mathbf{X}\beta \qquad (3)$$

$g(.)$ is the canonical link, which is that function $g(.)$ for which $g\ \boldsymbol{\mu} = \theta$. For Poisson, $\theta = \log\ \boldsymbol{\mu}$.

$\boldsymbol{X}\beta$ is the linear predictor where $\boldsymbol{X}$ contains values of the independent variables which include age, sex, HIV status among others and the parameter vector

$\beta = (\beta_0, \beta_1, \ldots \beta_\delta)$ is the vector of linear fixed-effects of categorical covariates.

In the model, some function of the mean of **y** (outcome variables) is a linear function of the predictors: $\boldsymbol{\eta} = \mathbf{X}\beta$, with **X** being a design matrix. Estimation of the parameters of generalized linear models is often done using the Maximum Likelihood method. The estimates are those parameter values that maximize the log likelihood, which for a single observation can be written.

Variables having a significant univariate test at some arbitrary level were selected as candidates for the multivariate analysis. This was based on a p-value cut-off point of 0.25 since levels such as 0.05 could fail to identify variables known to be important (Bendel RB, 1977) (Mickey RM, 1989). The association was considered statistically significant at 5% level of significance. Covariates were removed from the model when seen to be non-significant and not confounding. At the end of this process of deleting, refitting, and verifying, the model contained significant covariates, according to Hosmer and Lemeshow (Hosmer & and Lemeshow, 2000) (Hosmer & and Lemeshow, 1999). This data management, univariate analysis and multivariable analysis were carried out using R studio for Windows Version 0.99.486.

3.5 Case definitions

Treatment outcomes were classified according to WHO guidelines (WHO, Global Tuberculosis Report, 2015) as follows

Cured: A pulmonary TB patient with bacteriologically-confirmed TB at the beginning of treatment who was smear negative or culture negative in the last month of treatment and on at least one previous occasion.

Completed treatment: A TB patient who completed treatment without evidence of failure but with no record to show that sputum smear or culture results in the last month of treatment and on at least one previous occasion were negative, either because tests were not done or because results are unavailable.

Died: A TB patient who died from any cause during treatment.

Failed: A TB patient whose sputum smear or culture is positive at month five or later during treatment.

Lost to follow-up: A TB patient who did not start treatment or whose treatment was interrupted for two consecutive months or more.

Not evaluated: A TB patient for whom no treatment outcome is assigned. This includes cases 'transferred out' to another treatment unit as well as cases for whom the treatment outcome is unknown to the reporting unit.

Successfully treated: A patient who was cured or who completed treatment.

Cohort: A group of patients in whom TB has been diagnosed, and who were registered for treatment during a specified time period. This group forms the denominator for calculating treatment outcomes.

3.6 Statistical analysis

Data was received in an excel format, where cleaning was done, with patient identifiers like names and cell phone number omitted for confidentiality. 2009 Census data and data from Kenya Demographic and Health Survey 2014 was incorporated to this data to show the various variables including population densities, and other clinical indicators. The residential status at county level of each patient with TB was determined from the registers since coordinates of the households or specific health facilities were not available.

Exploratory data analysis

The demographic characteristics of the individuals were analysed and tabulated, by frequency of occurrence.

Observed probabilities of TB treatment outcomes, were calculated by dividing the specific treatment outcome, say total patients cured, by the sum of all possible outcomes. This gave an incidence rate for the given outcome.

The expected number of patients with a specific treatment outcome was calculated from the TB prevalence record obtained from KDHS. These expected numbers of cases were denoted by $E= (E_1 \ldots., E_n)$, were based on the size and demographic structure of the population living within each county. They were calculated by multiplying the population living in each county by the incidence rate for that county, for the specific TB treatment outcome.

Relative risk of TB treatment outcomes was calculated using the standardized mortality ratio (SMR), which compared the outcomes observed with those expected. This showed areas that experienced more than, or less than expected, of the treatment outcomes, with an aim of detecting which areas exhibit elevated risks. The observed numbers of TB treatment outcome cases per county were collectively denoted by $y= (y_1 \ldots., y_n)$, where y_k denoted the number of cases in county k. To assess which counties exhibited elevated levels of risk, the numbers of cases expected to occur per county were calculated. SMR was calculated for county k as $SMR_k = {}^{y_k}\ E_k$.Values above 1 represented areas with elevated levels of risk. However, elevated risks were likely to happen by chance if E_k was small, which could occur if the outcome in question was rare and/or the population at risk was small (P. Elliott, 2000) (S. Banerjee, 2004) (Lawson, 2008) .

To overcome this problem the Bayesian model-based approach was adopted, which estimated the set of disease risks using covariate information and a set of random effects. The random effects borrowed strength from values in neighbouring areas, which reduced the likelihood of excesses in risk occurring by chance.

The general formulation of the Bayesian model is given by

$$Y_k | E_k, R_k \sim \text{Poisson}\ E_k R_k \quad \text{for } k = 1 \ldots., n \qquad (4)$$

$$\ln R_k = \mu + x_k^T \beta + \emptyset_k$$

$\beta_i \sim N(0, 10)$ for i=1,…, p

$\mu \sim N(0,10)$

where R_k denoted the risk of disease in county k, which was modelled by an intercept term μ, a set of p covariates $x_k^T = x_{k1}, \ldots, x_{kp}$ and a random effect $\emptyset_k$. The regression parameters $\beta=\beta_1, \ldots\ldots, \beta_p$ and the intercept term μ were assigned weakly informative Gaussian prior distributions, with mean zero and variance 10. The random effects were included to model any over dispersion or spatial correlation in the data, which could persist after adjusting for the available covariate information. Over dispersion could occur because the Poisson likelihood enforces the restriction that $Var[Y_k] = E[Y_k]$, whereas in most studies of this type, $Var[Y_k] > E[Y_k]$. Unmeasured confounding can occur when an important spatially correlated covariate is either unmeasured or unknown, presence of neighbourhood effects, where subjects' behaviour is influenced by that of neighbouring subjects, and grouping effects, where subjects choose to be close to similar subjects can induce spatial autocorrelation into the response, which cannot be accounted for in a regression model. The random effects therefore exhibited a single global level of spatial autocorrelation.

Test for Autocorrelation

Global spatial autocorrelation of the data was assessed using Moran's I test with Monte Carlo simulations.

The Moran's I statistic is given by:

$$I = \frac{n}{S_0} \frac{\sum_{k=1}^{n} \sum_{j=1}^{n} W_{kj} (z_k - \bar{z})(z_j - \bar{z})}{\sum_{k=1}^{n} (z_k - \bar{z})^2}, \quad (5)$$

where

$S_0 = 2(\sum_k \sum_j W_{kj})$,

z_k may be either residuals $O_k - E_k$ or relative risks and W is a matrix which measures vicinity between regions.

The Moran's I statistic was used to tests the null hypothesis that the attribute being analysed was randomly distributed in the study area, that is, the spatial process promoting the observed pattern was random.

A Moran's scatterplot, which is a correlation between x and the spatial lag of x, formed by averaging all the values of x for neighbouring polygons, was also plotted to illustrate the types of spatial autocorrelation. The slope of the regression line is I statistic.

Local spatial autocorrelation testing is important since a significant Moran global statistic at a given spatial lag may hide large spatial patches of no autocorrelation, and an insignificant global result may hide patches of autocorrelation. This local spatial autocorrelation was tested using the Local Moran test and plotted on maps thereby showing the extent to which points that are close to a given point had similar values. A high positive Z score indicated that the surrounding features had similar values (either high values or low value).

Spatial analysis

Univariate GLM was used to identify factors that were individually associated with TB treatment outcomes. The spatial effects were investigated first where the county effect represented the spatial effect of county specific random effects. Thereafter, the estimation of more complex models was performed according to the clusters of the selected risk factors. The categorical covariates were dummy coded and the first factor levels were considered as reference categories.

Bayesian Modelling

If the outcome variable y is a random variable with probability density function $f(\theta|y)$ then, according to Bayes Theorem, the posterior density is proportional to the product of the prior density and the likelihood as shown in equation (6).

$$f\ \theta\ y\ \propto f\ \theta\ f\ y\ \theta\ , \qquad (6)$$

where $f(\theta|y)$ is the posterior density, $f(y|\theta)$ is the probability distribution for the data, conditional on the parameter θ (likelihood function), and the prior density being $f\ \theta$.

Before fitting the spatial model, the neighbourhood structure was defined, with two areas considered neighbours if they shared a common boundary. This was done using the function poly2nb () in R-INLA, and an adjacency neighbourhood matrix was created.

Model 1 was the unadjusted spatial model, which did not include the covariates.

Model 2 incorporated the various covariates which were all categorical, to assess autocorrelation in a regression model using a Poisson distribution. Augmentation of the linear predictors with a set of spatially correlated random effects, $\emptyset = \emptyset_1, \ldots, \emptyset_n$ (Lee, 2013) was done, using conditional autoregressive prior distributions, specified

by a set of n univariate full conditional distributions f $\emptyset_k|\emptyset_{-k}$ (where $\emptyset_{-k}= \emptyset_1,\ldots.,\emptyset_{k-1},\emptyset_{k+1},\ldots.,\emptyset_n$, for k=1,…,n. If two areas were defined to be neighbours their random effects were correlated, while random effects in non-neighbouring areas are modelled as being conditionally independent given the remaining elements of ϕ. Therefore, spatial autocorrelation is induced through adjacency structure of the aerial units; W which is a matrix in which element W_{kk} is n_k (the number of neighbours of area k) and element W_{kj} (with $k \neq j$) is 1 if areas k and j are neighbours and 0 otherwise. Areas (j,k) are neighbours if and only if they share a common border. The CAR described here is the intrinsic autoregressive model (Besag, York, & Mollie, 1991).

$$\emptyset_k|\emptyset_{-k,}W\tau_l^2 \sim N(\frac{1}{n_k} \quad \emptyset_{j_{j\sim k}}, \frac{\tau_l^2}{n_k}) \qquad (7)$$

The conditional expectation of $\emptyset_k$ was equal to the mean of the random effects in neighbouring areas, while the conditional variance was inversely proportional to the number of neighbours n_k. In the presence of strong spatial correlation, the more neighbours an area has, the more information there is in the data about the value of its random effect. The variance parameter τ_l^2 controls the amount of variation between the random effects, and the choice of hyperprior.

Inference

Inference was Bayesian and was based on obtaining the posterior distribution, of the different TB treatment outcomes, using the INLA methodology as a way of approximating the posterior marginals as proposed by Besag et al (Håvard Rue, 2009) with default prior probability distributions. Bayesian estimates were assessed as the posterior median, mean, mode and the interval. The assessment of the importance of a covariate in determination of spatial autocorrelation was indicated by the estimated value of its coefficient and its associated probability intervals. If the 95% credible intervals did not contain the value 0, we assumed that the coefficient was significant, and a value greater than zero indicated a positive relationship between the dependant variable and the response variable. Maps of the summary statistics of the posterior distributions of predicted outcomes of TB treatment and the covariate effects were then constructed.

CHAPTER FOUR

4. RESULTS AND DISCUSSION

This study involved several variables that were expected to have impacts on the TB treatment outcomes as shown in **Table [1].**

4.1 Exploratory Data Analysis

Demographic characteristics

As outlined in **Table [1],** a total of 23488 patient records were analysed, and 60.38% of were male. The most affected age group was that between 15 and 45 at 70.32%. Patients attended to at public facilities were 78.54%, while those tested negative of HIV were 59.03% compared to those positive at 34.07, 6.9% did not have a HIV test done. Pulmonary TB was the most common type of TB at 82.30%. Those found to be underweight were 58.11%.

Observed probabilities of the specific variables for TB were calculated and tabulated **Annex [4-5].** Despite the fact that most TB patients were aged between 15 and 4 5, counties like Lamu, Isiolo, Nyandarua and Marsabit had more children aged less than 15 years affected by TB at 100%, 90.9%, 91% and 80.7% respectively. In Turkana, only 37.8% of patients attended public facilities for TB treatment, while Garissa had 54%. The county with the highest number of TB patients with HIV were Siaya, Homabay and Kisumu at 69.18%, 66.19% and 63.7%. Garissa, Wajir, Mandera and Marsabit had the highest number of patients who were HIV negative at 94.9%, 96.9%, 96.7 and 90.8%. Observed number of the TB treatment success per county was calculated, tabulated **Annex [1]** and mapped **Figure [2].** The expected number of TB treatment outcomes was calculated, tabulated **Annex [2]** and mapped **Figure [3].** Most of the counties had low expected numbers of the various treatment outcomes except for Machakos which had a high expected number of defaulters, cured and those who complete treatment. Laikipia had a high expected number of treatment failure, while Bomet had a high expected number of deaths. The relative risk of TB treatment success was thereafter calculated, tabulated **Annex [3]** and mapped **Figure [4]**. Counties like Kisumu, Marsabit, Nairobi and Nandi had a high relative risk of greater than 20 for the various treatment outcomes. Of note was Tana River which had a high SMR for those cured and those who had completed treatment.

Nairobi had the highest death relative risk, while Uasin Gishu had a high relative risk of defaulters. Kwale, Lamu, Kiambu and Meru had a high relative risk for the failed treatment outcome. Since there was a generally low probability of occurrence of outcomes of interest (cured, dead, defaulted, failure and treatment completed), a Poisson probability distribution was used for the likelihood distribution.

Table 1: Demographic Characteristics of Patients with TB from January 2014 to March 2014

		Freq.	Percent
HIV Status	Not Done	1,621	6.9
	Negative	13,865	59.03
	Positive	8,002	34.07
Sex	Female	9,307	39.62
	Male	14,181	60.38
Sector	Faith Based	60	0.26
	Prisons	340	1.45
	Private	4,640	19.75
	Public	18,448	78.54
TB Type	ExtraPulmonary	4,158	17.7
	Pulmonary	19,330	82.3
Age	<15 years	2,381	10.14
	15-45 year	16,502	70.26
	Above 45	4,605	19.61
Outcome	Cured	3,163	13.47
	Dead	779	3.32
	Failure	57	0.24
	Still on treatment	14,332	61.02
	Defaulted	365	1.55
	Completed Treatment	4,416	18.8
	Transferred out	376	1.6
BMI	Underweight	1,441	6.14
	Normal	16,383	69.75
	Overweight	5,664	24.11

Autocorrelation

1. Global autocorrelation: Moran's I statistic

Test for global spatial autocorrelation was done using Moran's I test that used Monte Carlo simulations. The results showed significant positive spatial autocorrelation as expressed by P-values <0.05 for all outcomes except defaulters, as shown in **Table [2].**

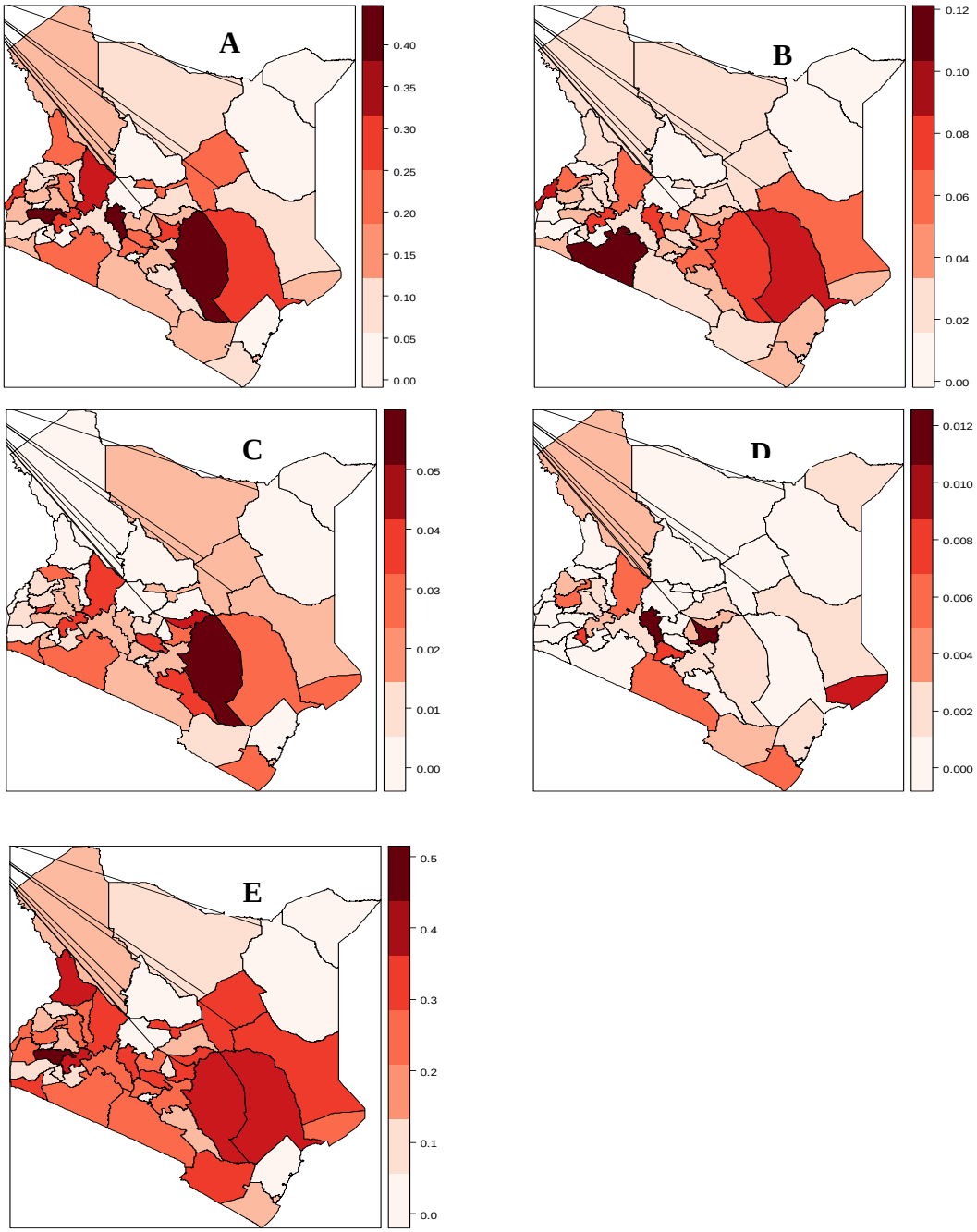

Figure 2: Shows the observed probability of success for the various tuberculosis treatment outcomes in Kenya observed between January and March 2015 with A=Cured, B=Dead, C=Defaulted, D=Failure, E=Treatment Complete

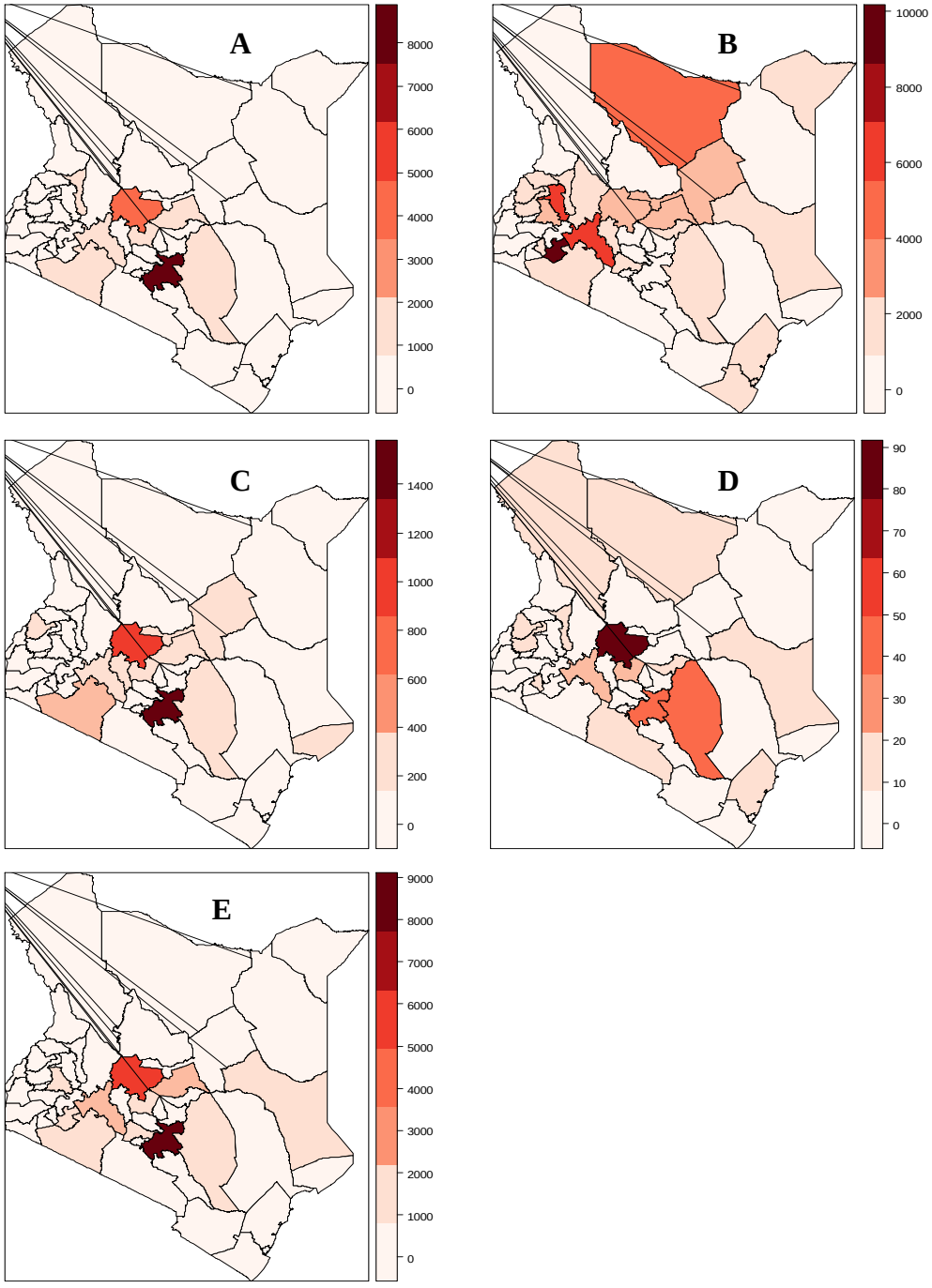

Figure 3: Shows the expected numbers of the various tuberculosis treatment outcomes in Kenya according to Kenya Demographic and Health Survey 2014 data with A=cured, B=Dead, C=Defaulted, D=Failure, E=Treatment Completed

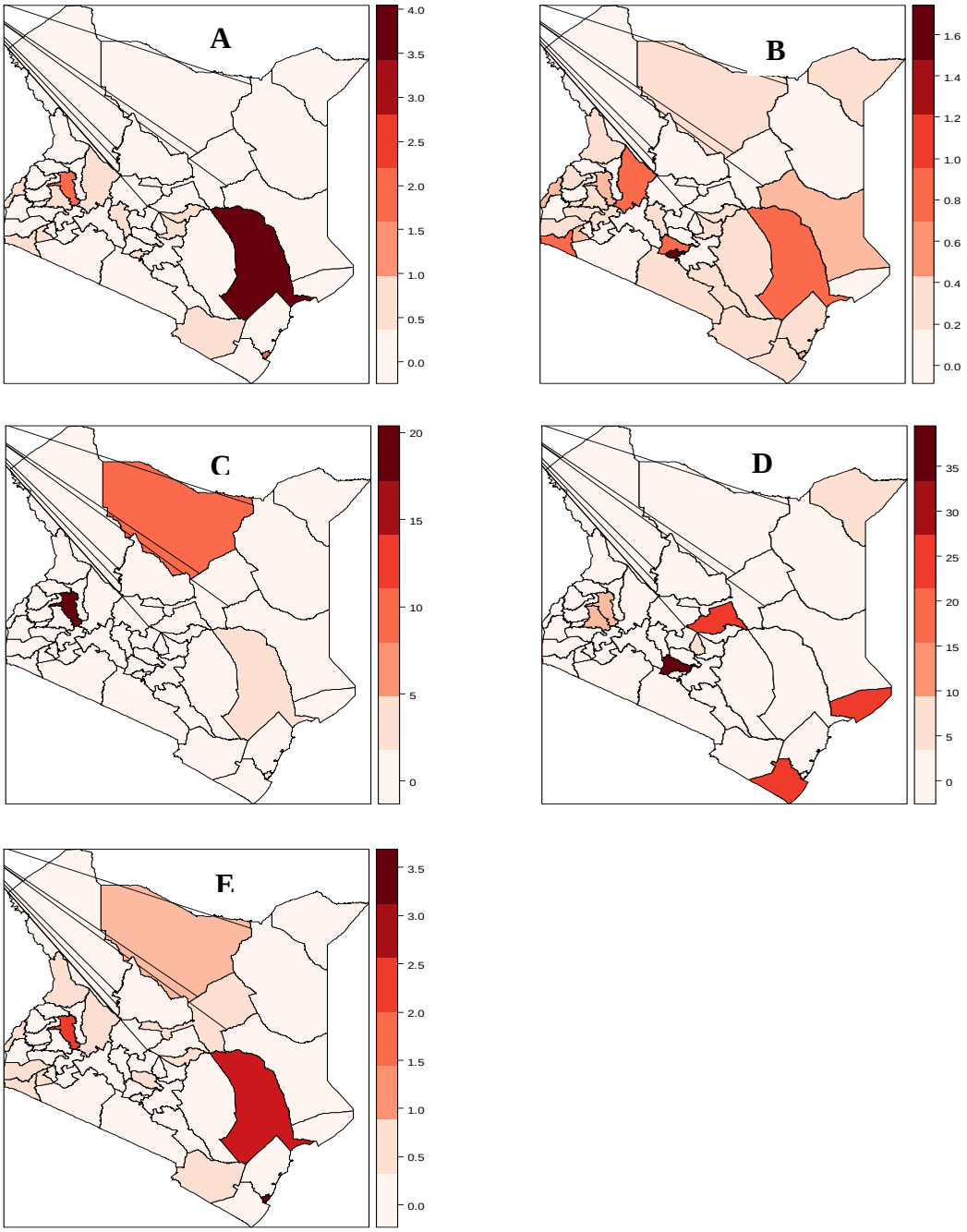

Figure 4: Shows the standard mortality ratio for the various Tuberculosis outcomes in Kenya observed between January 2014 and March 2014 with A=Cured, B=Dead, C=Defaulted, D=Failure, E=Treatment Complete

Table 2 : Shows the outputs of the Moran's I statistic

	Statistic Moran MC	**P value Moran MC**
Cured	0.1211	0.048
Dead	0.2478	0.004
Defaulted	-0.04	0.555
Failure	0.3644	0.001
Treatment Completed	0.1491	0.031

Moran's Scatterplot

There was evidence of positive spatial autocorrelation for all the treatment outcomes as seen in **Figure [5].**

2. Local spatial autocorrelation: Local Moran Test

In as much as there was no significant global spatial autocorrelation among patients who had defaulted, **Figure [6]** showed a high positive Z score in Kitui. The outcomes that had shown significant global spatial autocorrelation, had a number of counties with low Z scores for the Local Moran's, indicating that locally, there existed areas with no spatial autocorrelation, despite the significant global autocorrelation. This would have been missed if the local test was not done. Counties close to each other were seen to share the same local spatial autocorrelation pattern.

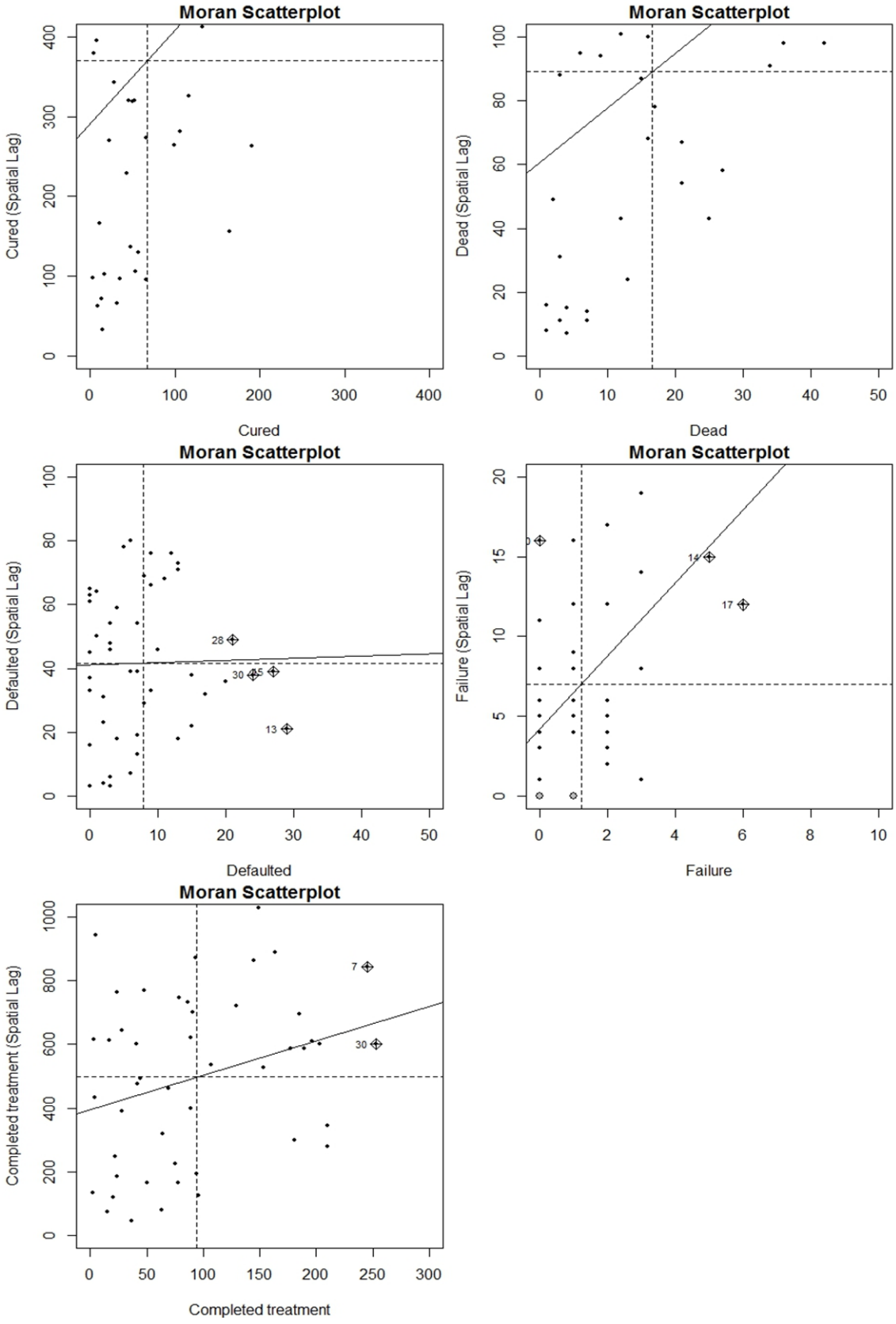

Figure 5: Shows the Moran Scatterplot for Tuberculosis treatment outcomes.

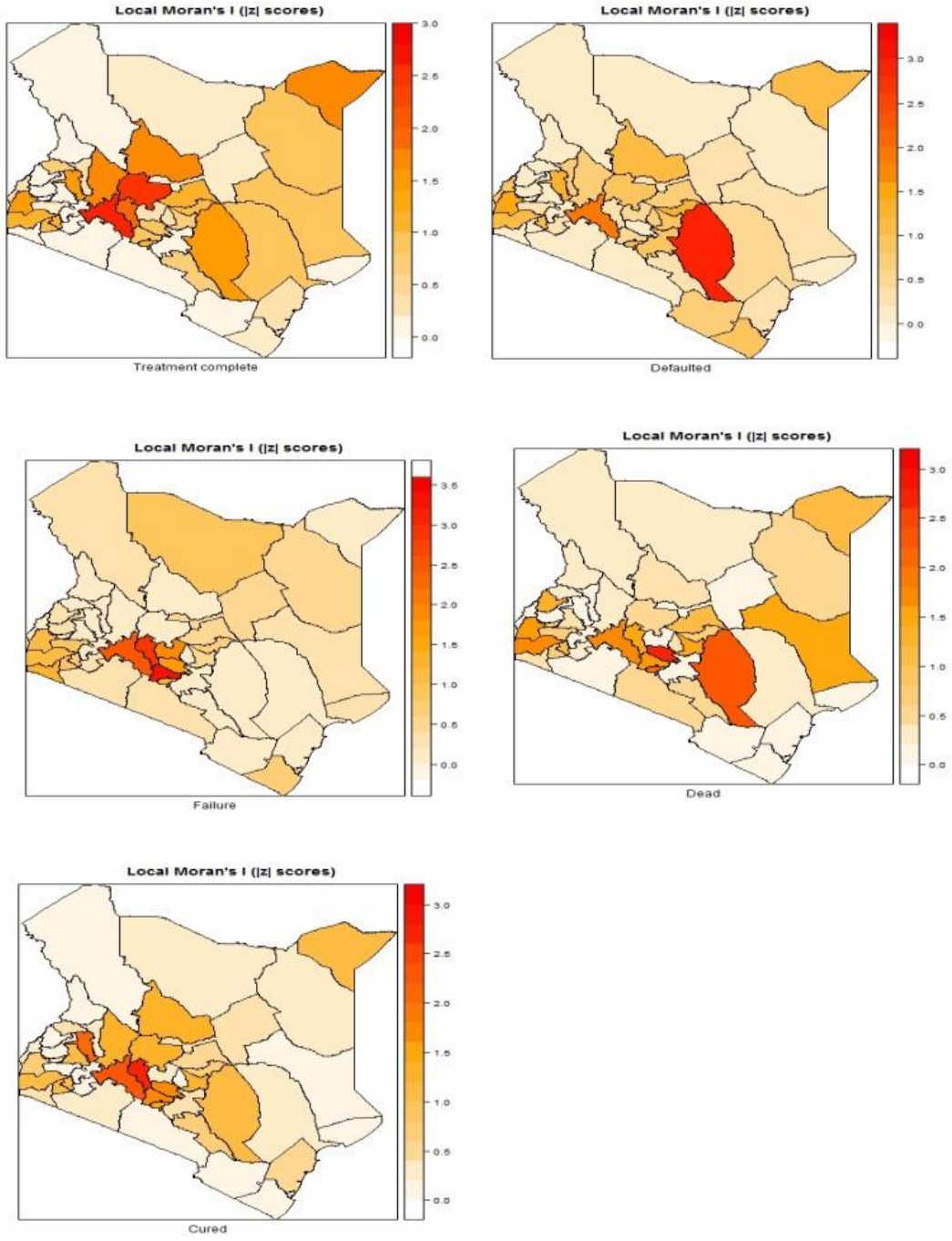

Figure 6: Shows the local Moran test statistic for the Tuberculosis treatment outcomes.

Variable selection

The univariate variable selection done for the 5 treatment outcomes are as seen in **Tables [3-7]**, with their respective estimates and significance level. These variables had a significant association with the respective treatment outcome, and were therefore included in the model.

Table 3: Estimated coefficients of the multivariate model for the outcome cured

	Estimate	**Std Error**	**z value**	**Pr(>|z)**
HIV pos	-7.93E-02	1.22E-01	-6.472	9.70E-11
Age >45	-7.03E-02	1.23E-02	-2.046	0.04071
Age <15	-7.15E-02	3.44E-02	-2.151	0.03149
Female	-1.18E-02	1.96E-03	-6.015	1.80E-09
Normal Weight	2.35E-02	5.25E-03	4.487	7.24E-06
Underweight	1.04E-02	9.12E-04	11.436	<2E-16
Age 15-45	-7.36E-02	3.36E-04	-2.187	0.02876
Private sector	6.98E-02	3.36E-02	2.075	0.03796
Public sector	7.01E-02	3.38E-02	2.076	0.03786
Urban	-2.94E-06	2.14E-07	-13.729	<2E-16
FBO	5.55E-02	3.35E-02	1.658	0.0973
Prisons	9.09E-02	3.37E-02	2.699	0.00696
Poverty	8.70E-07	1.37E-07	6.351	2.14E-10
Literate	-1.03E-02	1.71E-03	-5.99	2.10E-09
Cotrimox therapy	7.63E-02	1.23E-02	6.227	4.76E-10

Table 4: Estimated coefficients of the multivariate model for the outcome dead

	Estimate	**Std Error**	**z value**	**Pr(>|z)**
Female	-1.39E-02	1.46E-03	-9.481	< 2e-16
Normal Weight	2.06E-02	7.73E-03	2.668	0.007625
Underweight	7.06E-03	1.04E-03	6.805	1.01E-11
Age 15-45	-1.61E-02	1.88E-03	-8.607	< 2e-16
Urban	-1.40E-06	3.59E-03	-3.883	0.000103
Prisons	1.64E-02	4.86E-03	3.366	0.000764
New Patient	1.31E-02	1.87E-03	7.042	1.89E-12
Literate	-1.02E-02	2.95E-03	-3.445	0.000571

Table 5: Estimated coefficients of the multivariate model for the outcome defaulted

	Estimate	Std Error	z value	Pr(>|z)
HIV pos	1.63E-02	4.20E-03	3.871	0.0000108
HIV neg	1.58E-01	2.80E-02	5.638	1.72E-08
Age >45	3.57E-01	1.25E-01	2.858	0.00427
Age <15	3.30E-01	1.20E-01	2.74	0.006141
Female	3.55E-02	7.64E-03	4.644	3.42E-06
Normal Weight	2.96E-02	1.80E-02	1.645	0.0999973
Underweight	1.31E-02	3.22E-03	4.059	4.92E-05
Age 15-45	3.91E-01	1.23E-01	3.176	0.001492
Private sector	-4.24E-01	1.25E-01	-3.403	6.67E-04
Public sector	-4.24E-01	1.25E-01	-3.392	0.000695
Urban	-1.28E-05	1.34E-06	-9.539	<2e-16
FBO	-4.29E-01	1.20E-01	-3.583	0.00034
Prisons	-3.35E-01	1.23E-01	-2.731	0.006319
New patient	2.17E-02	5.91E-03	3.676	0.000237
Poverty	2.30E-06	4.44E-07	5.184	2.18E-07
Literate	-7.35E-03	6.23E-03	-1.18	0.238054
Cotrimox therapy	-1.67E-01	2.70E-02	-6.174	6.67E-10

Table 6: Estimated coefficients of the multivariate model for the outcome failure

	Estimate	Std Error	z value	Pr(>|z)
HIV neg	6.98E-02	1.47E-02	4.762	1.92E-06
HIV pos	4.75E-01	1.82E-01	2.609	0.00907
Age >45	1.34E+00	2.74E-01	4.901	9.53E+07
Age <15	1.51E+00	2.76E-01	5.464	4.66E-08
Female	-1.34E-01	3.03E-02	-4.419	9.09E-03
Normal Weight	-4.33E-01	6.43E-02	-6.729	1.70E-11
Underweight	1.57E-02	9.06E-03	1.734	0.08295
Age 15-45	1.32E+00	2.66E-01	4.937	7.93E-07
Private sector	-1.35E+00	2.57E-01	-5.242	1.59E-07
Public sector	-1.33E+00	2.60E-01	-5.133	2.85E-07
Urban	2.49E-05	1.84E-06	13.51	< 2e-16
FBO	-1.27E+00	2.65E-01	-4.768	1.86E-06
Prisons	-1.38E+00	2.70E-01	-5.099	3.41E-07
New patient	-2.63E-02	2.04E-02	-1.292	0.19645
Poverty	-6.98E-02	1.29E-06	-5.43	5.64E-08
Literate	5.02E+02	2.61E-02	1.975	0.04826
Cotrimox therapy	-3.98E-01	1.79E-01	-2.221	0.02635

Table 7: Estimated coefficients of the multivariate model for the outcome treatment complete

	Estimate	Std Error	z value	Pr(>|z)
HIV pos	1.59E-02	7.73E-03	2.062	0.039232
Age >45	8.67E-02	2.81E-02	3.086	0.002025
Age <15	7.54E-02	2.74E-02	2.753	0.005908
Normal Weight	3.28E-02	4.07E-03	8.05	8.29E-16
Underweight	1.00E-02	7.27E-04	13.787	< 2e-16
Age 15-45	8.69E-02	2.75E-02	3.166	0.001548
Private sector	-1.03E-01	2.74E-02	-3.743	0.000181
Public sector	-1.04E-01	2.76E-02	-3.774	0.00016
Urban	-4.36E-06	2.01E-07	-21.753	< 2e-16
FBO	-9.39E-02	2.74E-02	-3.426	0.000612
Prisons	-7.88E-02	2.76E-02	-2.859	0.004248
New patient	1.34E-02	1.36E-03	9.865	< 2e-16
Poverty	9.29E-07	1.00E-07	9.261	< 2e-16
Literate	-6.59E-03	1.43E-03	-4.62	3.84E-06
Cotrimox therapy	-2.11E-02	7.78E-03	-2.71	0.006729

4.2 Spatial analysis

An adjacency matrix to illustrate the neighbourhood structure of the counties was constructed, **Figure [7].**

Modelling

The covariates' estimates, standard deviation and 95% credibility intervals were as shown in **Table [8-12]** for individual TB treatment outcomes**.** Literacy showed a statistically significant spatial dependence for the patients who were cured. The less literate one was, the higher the chances of them being cured **Table [8].** Patients aged 15-45 years showed a statistically significant spatial dependence on the patients who died **Table [9].** Female patients were more likely to fail in their treatment than their male counterparts **Table [10].** Patients who had treatment failure had several variables showing spatial dependence. These were HIV status, age, literacy, female, sector, and normal weight **Table [11].** There were no variables that showed spatial dependence for those who had completed treatment **Table [12].** This widely non-significant spatial dependence may be partly due to the fact that in region based analysis, covariates may not directly be linked to the cases. **Figure [8]** shows the maps of both the unadjusted spatial effects, and total spatial effects. The maps of the

total spatial effect pattern are very similar to the unadjusted effect map for those patients who were cured, dead and who had completed treatment. Without taking into account the covariates, the risk of defaulters was highest in Vihiga county only, while the adjusted (total) spatial effects were significant in a number of counties especially in the western and coastal counties. Treatment failure showed significant unadjusted posterior means in the northern counties, while the adjusted means were significant in the southern and central counties.

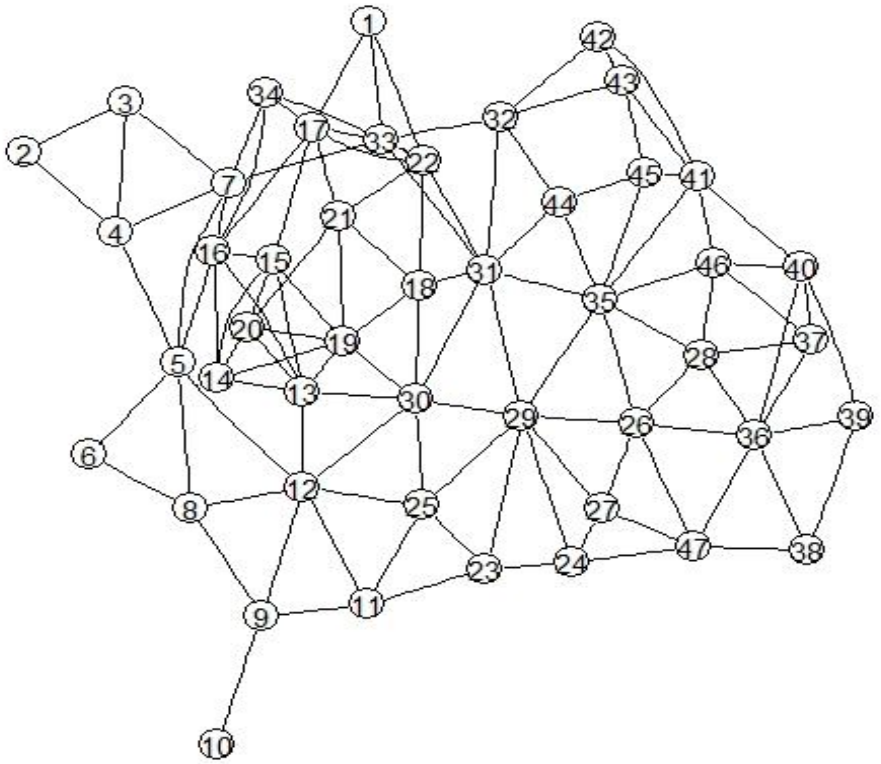

Figure 7: Displays the adjacency matrix with neighbourhood structure

KEY:

1. NAIROBI	2. MOMBASA	3. KWALE
4. KILIFI	5. TANA RIVER	6. LAMU
7. TAITA TAVETA	8. GARISSA	9. WAJIR
10. MANDERA	11. MARSABIT	12. ISIOLO
13. MERU	14. T'NITHI	15. EMBU
16. KITUI	17. MACHAKOS	18. NYANDARUA
19. NYERI	20. KIRINYAGA	21. MURANG'A
22. KIAMBU	23. TURKANA	24. WEST POKOT
25. SAMBURU	26. UASIN GISHU	27. E'MARAKWET
28. NANDI	29. BARINGO	30. LAIKIPIA
31. NAKURU	32. NAROK	33. KAJIADO
34. MAKUENI	35. KERICHO	36. KAKAMEGA
37. VIHIGA	38. BUNGOMA	39. BUSIA
40. SIAYA	41. HOMA BAY	42.MIGORI
43. KISII	44. BOMET	45. NYAMIRA
46. KISUMU	47.TRANS NZOIA	

Table 8: Estimated coefficients of the fixed effects for the outcome cured

	mean	**sd**	**0.025 quant**	**0.5 quant**	**0.975 quant**	**Estimated mean**
HIV pos	-0.0447	0.0634	-0.1699	-0.0447	0.0801	0.9563
Age >45	0.3774	0.2723	-0.1582	0.3770	0.9148	1.4585
Age <15	0.3535	0.2647	-0.1672	0.3530	0.8759	1.4240
Female	0.0059	0.0132	-0.0201	0.0059	0.0318	1.0059
Normal weight	0.0099	0.0385	-0.0659	0.0099	0.0858	1.0100
Underweight	0.0152	0.0084	-0.0014	0.0152	0.0318	1.0153
Age 15-45	0.3760	0.2685	-0.1520	0.3756	0.9058	1.4565
Private sector	-0.3889	0.2695	-0.9216	-0.3884	0.1408	0.6778
Public sector	-0.3887	0.2698	-0.9219	-0.3883	0.1414	0.6779
Urban	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
FBO	-0.3719	0.2676	-0.9009	-0.3715	0.1539	0.6894
Prisons	-0.3488	0.2662	-0.8750	-0.3484	0.1741	0.7055
Poor	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
Literate	-0.0255	0.0059	-0.0371	-0.0255	-0.0140	0.9748
Cotrimox therapy	0.0402	0.0644	-0.0867	0.0401	0.1671	1.0410

Table 9: Estimated coefficients of the fixed effects for the outcome dead

	mean	sd	0.025quant	0.5quant	0.975quant	Estimated mean
Female	-0.0094	0.0058	-0.0207	-0.0094	0.0020	0.9906
Normal weight	-0.0034	0.0253	-0.0536	-0.0034	0.0462	0.9966
Underweight	0.0073	0.0045	-0.0016	0.0073	0.0161	1.0073
Age 15-45	-0.0115	0.0058	-0.0229	-0.0116	-0.0001	0.9885
Urban	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
Prisons	0.0095	0.0201	-0.0302	0.0095	0.0490	1.0095
New patient	0.0069	0.0063	-0.0055	0.0069	0.0191	1.0069
Literate	-0.0113	0.0084	-0.0280	-0.0113	0.0053	0.9887

Table 10: Estimated coefficients of the fixed effects for the outcome failure

	mean	sd	0.025 quant	0.5 quant	0.975 quant	Estimated mean
HIV neg	0.0668	0.0147	0.0389	0.0665	0.0966	1.0691
HIV pos	0.4428	0.1820	0.0415	0.4597	0.7545	1.5571
Age >45	1.3466	0.2741	0.8108	1.3457	1.8865	3.8442
Age <15	1.5151	0.2760	0.9748	1.5146	2.0581	4.5500
Female	-0.1389	0.0303	-0.2003	-0.1383	-0.0811	0.8703
Normal weight	-0.4444	0.0643	-0.5763	-0.4425	-0.3236	0.6412
Underweight	0.0156	0.0091	-0.0012	0.0153	0.0344	1.0157
Age 15-45	1.3132	0.2664	0.7917	1.3127	1.8371	3.7182
Private sector	-1.3401	0.2574	-1.8479	-1.3393	-0.8373	0.2618
Public sector	-1.3270	0.2596	-1.8395	-1.3260	-0.8200	0.2653
Urban	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
FBO	-1.2671	0.2653	-1.7950	-1.2648	-0.7527	0.2816
Prisons	-1.3808	0.2696	-1.9116	-1.3803	-0.8531	0.2514
New patient	-0.0290	0.0204	-0.0681	-0.0293	0.0119	0.9715
Poor	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
Literate	0.0537	0.0261	0.0042	0.0531	0.1068	1.0552
Cotrimoxazole therapy	-0.3686	0.1793	-0.6746	-0.3858	0.0273	0.6917

Table 11: Estimated coefficients of the fixed effects for the outcome treatment complete

	mean	sd	0.025quant	0.5quant	0.975quant	Estimated mean
HIV pos	-0.0027	0.0670	-0.1348	-0.0027	0.1293	0.9973
Age >45	0.5291	0.2898	-0.0415	0.5287	1.1008	1.6974
Age <15	0.5027	0.2823	-0.0529	0.5023	1.0595	1.6531
Age 15-45	0.5218	0.2851	-0.0395	0.5214	1.0843	1.6851
Normal weight	0.0140	0.0417	-0.0681	0.0139	0.0961	1.0141
Under weight	0.0081	0.0095	-0.0105	0.0081	0.0268	1.0082
Private sector	-0.5318	0.2873	-1.0996	-0.5315	0.0331	0.5875
Public sector	-0.5372	0.2882	-1.1067	-0.5369	0.0294	0.5844
Urban	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
FBO	-0.5168	0.2880	-1.0859	-0.5165	0.0494	0.5964
Prisons	-0.5169	0.2871	-1.0842	-0.5166	0.0477	0.5964
New patient	0.0127	0.0160	-0.0188	0.0126	0.0443	1.0127
Poor	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
Literate	-0.0037	0.0149	-0.0332	-0.0037	0.0257	0.9963
Cotrim therapy	-0.0029	0.0675	-0.1360	-0.0029	0.1301	0.9971

Table 12: Estimated coefficients of the fixed effects for the outcome defaulted

	mean	sd	0.025 quant	0.5 quant	0.975 quant	Estimated mean
HIV neg	0.0300	0.0433	-0.0557	0.0301	0.1151	1.0305
HIV pos	0.1494	0.2352	-0.3143	0.1491	0.6142	1.1612
Age >45	1.6648	0.9241	-0.1533	1.6624	3.4938	5.2847
Age <15	1.5574	0.8908	-0.1959	1.5553	3.3198	4.7464
Female	0.1030	0.0487	0.0071	0.1028	0.1993	1.1085
Normal weight	-0.0574	0.1306	-0.3147	-0.0576	0.2007	0.9443
Underweight	0.0065	0.0305	-0.0536	0.0064	0.0669	1.0065
Age 15-45	1.6995	0.9063	-0.0843	1.6973	3.4925	5.4710
Private sector	-1.7264	0.9161	-3.5431	-1.7239	0.0736	0.1779
Public sector	-1.7281	0.9180	-3.5487	-1.7256	0.0758	0.1776
Urban	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
FBO	-1.6734	0.9106	-3.4773	-1.6715	0.1176	0.1876
Prisons	-1.6886	0.9031	-3.4780	-1.6867	0.0874	0.1848
New patient	0.0028	0.0589	-0.1116	0.0022	0.1208	1.0028
Poor	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000
Literate	0.0033	0.0483	-0.0928	0.0035	0.0981	1.0033
Cotrim therapy	-0.1683	0.2231	-0.6104	-0.1679	0.2705	0.8451

Maps of the fitted values (estimated risk) by the mean of the various TB outcomes are provided in **Figure [8].** The estimated risk for the patients who were cured was high in Meru and its surrounding counties. Other areas were Murang'a, Machakos and Kisii. Estimated risk for the patients who died, defaulted, failed treatment and those who completed treatment were highest in Kisumu, Kisii, Kitui and Nakuru respectively. Neighbouring counties had a relatively high estimated risk. Corresponding analysis of the estimated risk by fitting the standard deviation was done as shown in **Figure [9]** and showed a similar pattern as the mean. There was a significant spatial effect across the various TB outcomes, though most of the covariates did not explain the spatial autocorrelation seen in this analysis. This could mean that there observed spatial autocorrelation could be explained by other unobserved variables.

Spatial dependence structure can be modelled in different ways using available statistical models. It is convenient to ensure that observation are independent and identically distributed, which may not always be the case when working with data that exhibits some correlation between areas. This study follows the conditional autoregressive processes (CAR) model as was proposed by Besag et al 1991 (Besag, York, & Mollie, 1991) and application based on the method proposed by Havard Rue (Håvard Rue, 2009) to express the spatial dependence. INLA uses numerical methods to approximate posterior marginal distributions since it is robust and fast for large scale complex spatial and spatiotemporal analysis. However, due to its recent inspection, INLA is less established than MCMC methods. The CAR model is the most commonly used in practice for region based spatial analysis. In this study, the weak spatial dependence may imply that appropriate risk factors were adjusted for in the model such that the spatial random effect became less important.

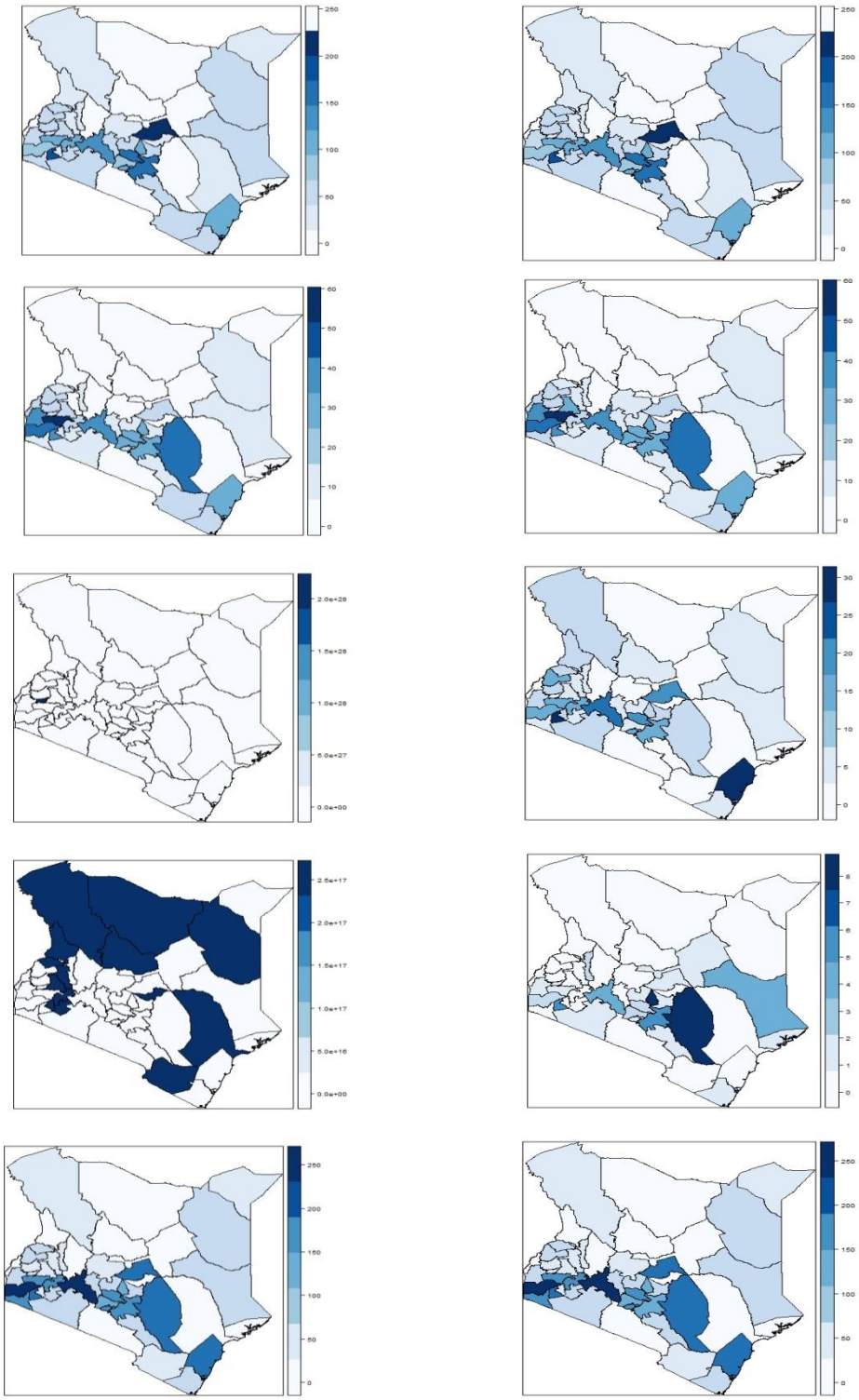

Figure 8: Shows the fitted model by the mean with A=Cured, B=Dead, C=Defaulted, D=Failure, E=Treatment complete. Left: Unadjusted, Right: Adjusted spatial effects.

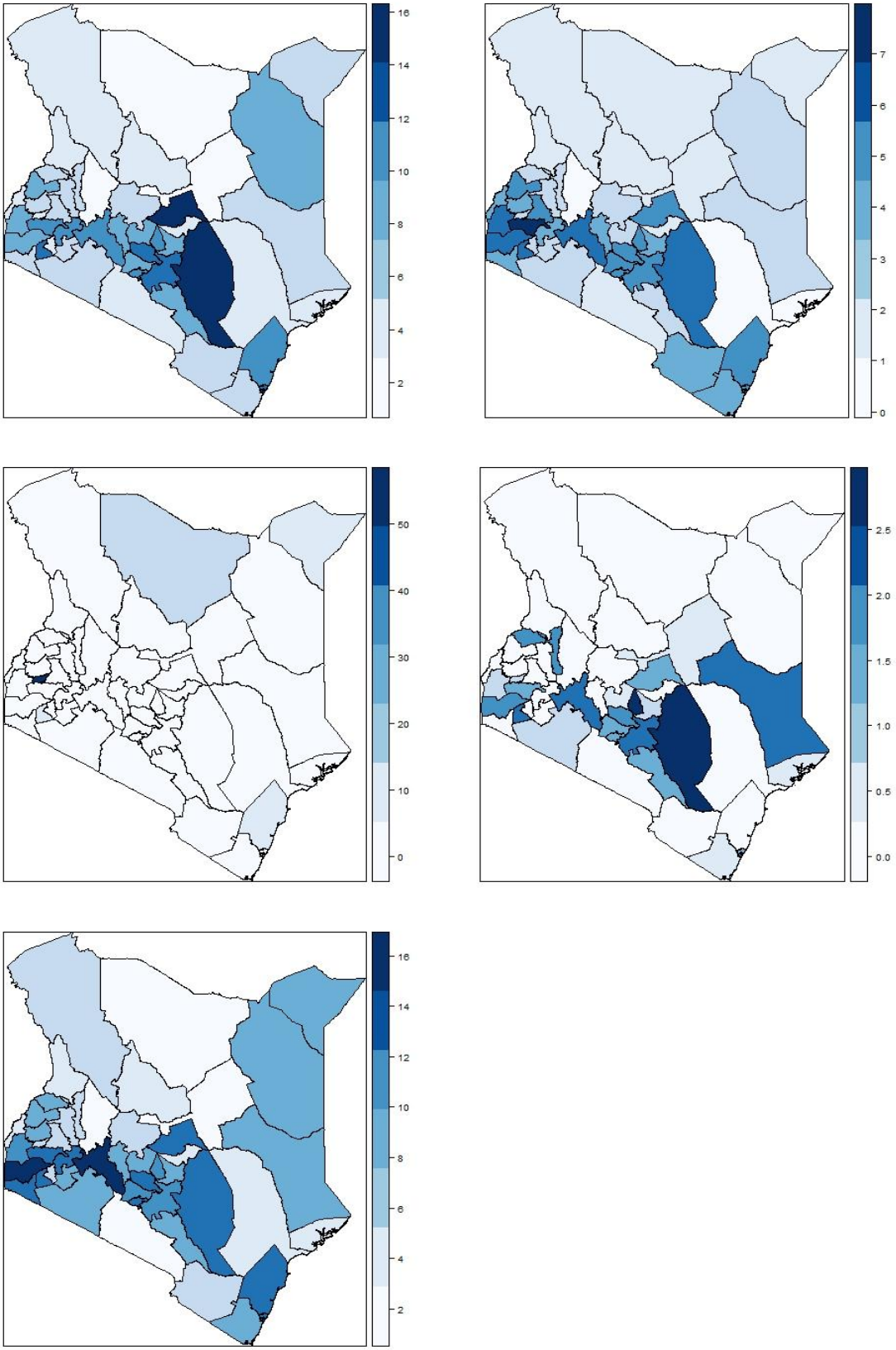

Figure 9: Shows the fitted model by the standard deviation with A=Cured, B=Dead, C=Defaulted, D=Failure, E=Treatment complete

Conclusion and Recommendation

There is increased availability of spatial data has fuelled a growth in modelling in this area.

We recommend that future related studies involving various TB outcomes be traced at household level (point level data). This will enable minimal mismatch between risk factors and the TB outcomes.

Other statistical techniques could be used to develop more robust spatial models. This could also be done over time to further adjust for the temporal autocorrelation in the model.

However, other methods for classical statistical analysis are also available like generalized mixed effect models and generalized additive models (GAM). A specific difficulty with most such methods is their use of point support for spatial dependency pattern rather than polygon so that the relationship between observations is distance based. Due to the complexity in the matrix algebra (covariance matrices) involved, such methods tend to focus on interpolation than modelling. Nevertheless, Bayesian approach is natural when technical inference with nonlinear terms is involved.

ANNEX

ANNEX 1

Shows the observed numbers of TB outcomes

County	Cured	Dead	Defaulted	Failure	Trtmnt compl
Baringo	8.6915	1.8437	1.1413	0.0878	15.8906
Bomet	30.5839	22.9379	0	0	22.9379
Bungoma	327.3236	71.4161	35.708	0	470.1557
Busia	201.0586	74.0742	52.9102	10.582	264.5508
Elgeyo-Marakwet	9.2734	14.3316	1.6861	0.843	18.5468
Embu	441.374	163.939	50.4427	0	630.5344
Garissa	143.5928	39.1617	23.497	7.8323	242.8024
Homa Bay	750.3591	99.104	42.4732	14.1577	1359.1409
Isiolo	70.9373	38.0639	11.2462	1.7302	211.9467
Kajiado	52.8718	13.2179	0	0	26.4359
Kakamega	130.1315	59.1507	0	11.8301	59.1507
Kericho	330.5533	146.058	69.1856	0	491.9863
Kiambu	1612.523	246.963	174.3268	0	2382.4669
Kilifi	289.9031	100.978	16.2867	9.772	472.3141
Kirinyaga	292.8019	74.5816	80.1062	5.5246	580.0792
Kisii	911.0656	194.672	85.6557	38.9344	1004.5082
Kisumu	8293.046	1476.25	1172.315	43.4191	8510.1411
Kitui	409.4856	197.683	52.9507	3.53	624.8185
Kwale	1087.943	193.617	46.0993	27.6596	935.8156
Laikipia	228.8217	73.2229	13.7293	0	315.7739
Lamu	42.4507	7.7183	0	1.2864	36.0188
Machakos	571.35	43.95	87.9	0	879
Makueni	370.4937	69.6079	29.1904	11.2271	334.567
Mandera	134.7743	18.3783	4.0841	4.0841	181.7412
Marsabit	149.3333	18.6667	0	0	294
Meru	129.8098	43.2699	0	0	86.5399
Migori	848.8888	91.664	79.7079	7.9708	753.2393
Mombasa	450.9922	130.094	26.0188	0	1326.9577
Murang'a	268.3507	40.9071	27.8168	3.2726	343.6198
Nairobi	4735.977	912.959	599.1296	85.5899	5278.0464
Nakuru	1142.047	311.467	207.6449	25.9556	2188.9234
Nandi	1282.135	598.33	170.9514	0	1196.6595
Narok	203.7008	38.1939	33.9501	8.4875	364.964
Nyamira	104.9873	44.4177	36.3418	0	193.8228
Nyandarua	295.1549	71.4085	33.3239	4.7606	433.2113
Nyeri	383.2818	80.6909	26.897	13.4485	598.4576
Samburu	51.6402	22.9512	34.4268	0	86.0671
Siaya	382.1735	220.039	40.5336	11.581	619.5843
Taita Taveta	99.1598	32.3787	0	0	82.9704
Tana River	127.1538	6.6923	6.6923	0	160.6154
Tharaka-Nithi	52.8175	8.8029	2.9343	0	49.8832
Trans Nzoia	60.7802	15.195	10.13	0	119.0279
Turkana	32.6271	6.5254	15.226	0	80.4802
Uasin Gishu	548.523	117.541	58.7703	0	861.9647
Vihiga	167.8216	61.026	0	15.2565	213.5911
Wajir	215.9259	22.9012	9.8148	0	255.1852
West Pokot	167.2525	29.5152	68.8687	0	236.1212

ANNEX 2

Shows the expected number of TB treatment outcomes

County	Cured	Dead	Defaulted	Failure	Trtmnt compl
Baringo	4	3	0	0	3
Bomet	55	12	6	0	79
Bungoma	57	21	15	3	75
Busia	11	17	2	1	22
Elgeyo-Marakwet	35	13	4	0	50
Embu	55	15	9	3	93
Garissa	53	7	3	1	96
Homa Bay	82	44	13	2	245
Isiolo	8	2	0	0	4
Kajiado	11	5	0	1	5
Kakamega	43	19	9	0	64
Kericho	111	17	12	0	164
Kiambu	89	31	5	3	145
Kilifi	106	27	29	2	210
Kirinyaga	117	25	11	5	129
Kisii	191	34	27	1	196
Kisumu	116	56	15	1	177
Kitui	236	42	10	6	203
Kwale	50	16	3	0	69
Laikipia	33	6	0	1	28
Lamu	13	1	2	0	20
Machakos	165	31	13	5	149
Makueni	66	9	2	2	89
Mandera	32	4	0	0	63
Marsabit	3	1	0	0	2
Meru	213	23	20	2	189
Migori	52	15	3	0	153
Mombasa	164	25	17	2	210
Murang'a	166	32	21	3	185
Nairobi	99	21	13	1	181
Nakuru	132	36	24	3	253
Nandi	45	21	6	0	42
Narok	48	9	8	2	86
Nyamira	26	11	9	0	48
Nyandarua	62	15	7	1	91
Nyeri	57	12	4	2	89
Samburu	9	4	6	0	15
Siaya	66	38	7	2	107
Taita Taveta	49	16	0	0	41
Tana River	19	1	1	0	24
Tharaka-Nithi	18	3	1	0	17
Trans Nzoia	48	12	8	0	94
Turkana	15	3	7	0	37
Uasin Gishu	28	6	3	0	44
Vihiga	22	8	0	2	28
Wajir	66	7	3	0	78
West Pokot	17	3	7	0	24

ANNEX 3

Shows the relative risk of the TB treatment outcome

County	Cured	Dead	Defaulted	Failure	Trtmnt compl
Baringo	2.1729	0.6146	-	-	5.2969
Bomet	0.5561	1.9115	0	-	0.2904
Bungoma	5.7425	3.4008	2.3805	0	6.2687
Busia	18.2781	4.3573	26.4551	10.582	12.025
Elgeyo-Marakwet	0.265	1.1024	0.4215	-	0.3709
Embu	8.025	10.9293	5.6047	0	6.7799
Garissa	2.7093	5.5945	7.8323	7.8323	2.5292
Homa Bay	9.1507	2.2524	3.2672	7.0789	5.5475
Isiolo	8.8672	19.032	-	-	52.9867
Kajiado	4.8065	2.6436	-	0	5.2872
Kakamega	3.0263	3.1132	0	-	0.9242
Kericho	2.978	8.5917	5.7655	-	2.9999
Kiambu	18.1182	7.9665	34.8654	0	16.4308
Kilifi	2.7349	3.7399	0.5616	4.886	2.2491
Kirinyaga	2.5026	2.9833	7.2824	1.1049	4.4967
Kisii	4.77	5.7257	3.1724	38.9344	5.125
Kisumu	71.4918	26.3616	78.1544	43.4191	48.0799
Kitui	1.7351	4.7067	5.2951	0.5883	3.0779
Kwale	21.7589	12.1011	15.3664	-	13.5625
Laikipia	6.934	12.2038	-	0	11.2776
Lamu	3.2654	7.7183	0	-	1.8009
Machakos	3.4627	1.4177	6.7615	0	5.8993
Makueni	5.6135	7.7342	14.5952	5.6135	3.7592
Mandera	4.2117	4.5946	-	-	2.8848
Marsabit	49.7778	18.6667	-	-	147
Meru	0.6094	1.8813	0	0	0.4579
Migori	16.3248	6.1109	26.5693	-	4.9231
Mombasa	2.75	5.2038	1.5305	0	6.3188
Murang'a	1.6166	1.2783	1.3246	1.0909	1.8574
Nairobi	47.8381	43.4743	46.0869	85.5899	29.1605
Nakuru	8.6519	8.6519	8.6519	8.6519	8.6519
Nandi	28.4919	28.4919	28.4919	-	28.4919
Narok	4.2438	4.2438	4.2438	4.2438	4.2438
Nyamira	4.038	4.038	4.038	-	4.038
Nyandarua	4.7606	4.7606	4.7606	4.7606	4.7606
Nyeri	6.7242	6.7242	6.7242	6.7242	6.7242
Samburu	5.7378	5.7378	5.7378	-	5.7378
Siaya	5.7905	5.7905	5.7905	5.7905	5.7905
Taita Taveta	2.0237	2.0237	-	-	2.0237
Tana River	6.6923	6.6923	6.6923	-	6.6923
Tharaka-Nithi	2.9343	2.9343	2.9343	-	2.9343
Trans Nzoia	1.2663	1.2663	1.2663	-	1.2663
Turkana	2.1751	2.1751	2.1751	-	2.1751
Uasin Gishu	19.5901	19.5901	19.5901	-	19.5901
Vihiga	7.6283	7.6283	-	7.6283	7.6283
Wajir	3.2716	3.2716	3.2716	-	3.2716
West Pokot	9.8384	9.8384	9.8384	-	9.8384

ANNEX 4

Shows the probabilities of the clinical variables

County	HIVpos	HIVneg	HIVND	newpt	public	private	prisons	FBO	PTB	cotri
Nairobi	0.3415	0.5178	0.1407	0.8542	0.6351	0.3501	0.0148	0	0.7501	0.9791
Baringo	0.2628	0.6667	0.0705	0.8758	0.9876	0.0124	0	0	0.7143	1
Bomet	0.2421	0.6479	0.11	0.9367	0.9538	0.0462	0	0	0.8224	0.9798
Bungoma	0.3549	0.5804	0.0647	0.9004	0.8086	0.1758	0.0156	0	0.8633	1
Busia	0.5089	0.4304	0.0608	0.8481	0.8709	0.1114	0.0177	0	0.8304	0.99
Elgeyo-Mara	0.3435	0.5038	0.1527	0.9008	0.8855	0.0076	0	0.1069	0.8015	1
Embu	0.2425	0.7335	0.024	0.9251	0.8772	0.0808	0.0419	0	0.9042	0.9753
Garissa	0.0503	0.9497	0	0.9195	0.5436	0.3893	0.0671	0	0.8154	1
Homa Bay	0.6619	0.3298	0.0083	0.9018	0.8355	0.1538	0.0107	0	0.8047	0.9928
Isiolo	0.2051	0.75	0.0449	0.9038	0.9936	0.0064	0	0	0.8526	1
Kajiado	0.3178	0.6082	0.074	0.9068	0.8137	0.1863	0	0	0.874	1
Kakamega	0.3735	0.5766	0.0499	0.8814	0.8419	0.1478	0.0103	0	0.8093	1
Kericho	0.3047	0.6387	0.0566	0.9261	0.8424	0.1342	0.0233	0	0.8541	1
Kiambu	0.2867	0.6262	0.0871	0.8933	0.7153	0.2642	0.0049	0.0157	0.8571	0.9863
Kilifi	0.3154	0.6165	0.0681	0.9144	0.7227	0.2583	0.019	0	0.8542	0.9849
Kirinyaga	0.2272	0.7635	0.0094	0.8923	0.8267	0.1663	0.007	0	0.8993	1
Kisii	0.3883	0.6033	0.0084	0.9212	0.8859	0.0871	0.027	0	0.9025	1
Kisumu	0.6369	0.3402	0.0229	0.8954	0.7019	0.2524	0.0457	0	0.8486	0.9924
Kitui	0.3121	0.6773	0.0106	0.8901	0.906	0.0922	0.0018	0	0.8493	0.9943
Kwale	0.2524	0.6166	0.131	0.8854	0.9268	0.0732	0	0	0.7962	0.9367
Laikipia	0.3679	0.6274	0.0047	0.8779	0.8685	0.108	0.0235	0	0.8498	1
Lamu	0.2167	0.7833	0	0.9	1	0	0	0	0.7667	1
Machakos	0.245	0.735	0.0199	0.897	0.952	0.0395	0.0085	0	0.8025	0.9942
Makueni	0.3193	0.6563	0.0244	0.8938	0.8783	0.0642	0	0.0575	0.9049	1
Mandera	0.0321	0.9679	0	0.9038	1	0	0	0	0.8526	1
Marsabit	0.0613	0.908	0.0307	0.9387	0.6687	0.3313	0	0	0.8098	1
Meru	0.1933	0.764	0.0427	0.9045	0.7831	0.1966	0.0202	0	0.8517	0.9884
Migori	0.5259	0.4647	0.0094	0.9437	0.8185	0.1815	0	0	0.8451	1
Mombasa	0.3177	0.6311	0.0512	0.855	0.6693	0.2847	0.046	0	0.8134	0.9918
Murang'a	0.2476	0.7253	0.0271	0.9091	0.8665	0.1044	0.029	0	0.8685	1
Nakuru	0.3522	0.551	0.0968	0.9069	0.8329	0.1567	0.0104	0	0.8251	1
Nandi	0.3224	0.623	0.0546	0.8703	0.9459	0.0216	0.0108	0.0216	0.8378	0.9492
Narok	0.3156	0.6201	0.0642	0.9197	0.8089	0.1884	0.0028	0	0.8615	0.9912
Nyamira	0.3882	0.5907	0.0211	0.9241	0.9156	0.0844	0	0	0.8861	1
Nyandarua	0.3602	0.6256	0.0142	0.8826	0.8028	0.1972	0	0	0.8498	0.9868
Nyeri	0.3313	0.614	0.0547	0.8636	0.7758	0.2152	0.0091	0	0.8121	0.9908
Samburu	0.311	0.622	0.0671	0.9024	0.7744	0.2195	0.0061	0	0.9146	0.9412
Siaya	0.6918	0.2984	0.0098	0.8887	0.8592	0.1211	0.0196	0	0.7791	0.9976
Taita Taveta	0.2952	0.5904	0.1145	0.8757	0.9704	0.0296	0	0	0.8876	0.9796
Tana River	0.125	0.8173	0.0577	0.9135	0.9712	0.0288	0	0	0.8654	1
Tharaka-Nith	0.2103	0.7491	0.0406	0.927	0.708	0.292	0	0	0.7482	1
Trans Nzoia	0.2693	0.5882	0.1424	0.9567	0.8173	0.1796	0.0031	0	0.8762	0.8046
Turkana	0.272	0.6402	0.0878	0.9718	0.3785	0.6215	0	0	0.8701	0.9688
Uasin Gishu	0.3752	0.5593	0.0655	0.917	0.8622	0.1254	0.0124	0	0.7809	0.9811
Vihiga	0.4349	0.5056	0.0595	0.8848	0.8625	0.1375	0	0	0.7546	1
Wajir	0.0185	0.9691	0.0123	0.9877	1	0	0	0	0.8704	1
West Pokot	0.0842	0.633	0.2828	0.9226	0.8249	0.1751	0	0	0.7609	0.88

ANNEX 5

Shows probability of patient specific variables

County	Child	Young adult	Old adult	Female	Underwgt	Normal wgh	Overwght
Nairobi	0.0826	0.7724	0.1449	0.3994	0.9409	0.0422	0.0169
Baringo	0.1366	0.6832	0.1801	0.3975	0.9938	0.0062	0
Bomet	0.1022	0.7251	0.1727	0.4209	0.9708	0.0195	0.0097
Bungoma	0.1191	0.6191	0.2617	0.4453	0.9473	0.0293	0.0234
Busia	0.1367	0.6025	0.2608	0.4835	0.9443	0.038	0.0177
Elgeyo-Marak	0.1221	0.5802	0.2977	0.3511	0.9847	0.0076	0.0076
Embu	0.1497	0.6856	0.1647	0.3293	0.9731	0.024	0.003
Garissa	0.1946	0.557	0.2483	0.396	0.9362	0.0503	0.0134
Homa Bay	0.1101	0.6805	0.2095	0.4627	0.9609	0.026	0.013
Isiolo	0.1282	0.6154	0.2564	0.3397	0.9359	0.0385	0.0256
Kajiado	0.1456	0.6758	0.1786	0.4	0.9288	0.0603	0.011
Kakamega	0.0842	0.6684	0.2474	0.4811	0.9553	0.0395	0.0052
Kericho	0.0858	0.7661	0.1481	0.3599	0.9689	0.0175	0.0136
Kiambu	0.0774	0.7649	0.1577	0.362	0.9384	0.044	0.0176
Kilifi	0.1065	0.6725	0.221	0.4263	0.9588	0.0285	0.0127
Kirinyaga	0.0965	0.6753	0.2282	0.3185	0.9602	0.0328	0.007
Kisii	0.0685	0.7386	0.1929	0.3797	0.9585	0.0332	0.0083
Kisumu	0.0758	0.7581	0.1661	0.4399	0.7632	0.024	0.2127
Kitui	0.0816	0.6755	0.2429	0.3741	0.9734	0.0213	0.0053
Kwale	0.1019	0.6306	0.2675	0.4236	0.9459	0.0318	0.0223
Laikipia	0.061	0.7653	0.1737	0.3474	0.9577	0.0282	0.0141
Lamu	0.0833	0.6333	0.2833	0.3167	1	0	0
Machakos	0.0552	0.7143	0.2306	0.3343	0.9676	0.0254	0.0071
Makueni	0.08	0.6933	0.2267	0.4004	0.9757	0.0199	0.0044
Mandera	0.0577	0.641	0.3013	0.4295	0.9679	0.0192	0.0128
Marsabit	0.1296	0.6049	0.2654	0.4847	0.9816	0.0123	0.0061
Meru	0.073	0.7921	0.1348	0.3079	0.982	0.0135	0.0045
Migori	0.0861	0.6792	0.2347	0.4585	0.939	0.0485	0.0125
Mombasa	0.073	0.7828	0.1442	0.3481	0.9323	0.0425	0.0252
Murang'a	0.0504	0.6957	0.2539	0.3056	0.9555	0.0329	0.0116
Nakuru	0.121	0.7232	0.1558	0.4291	0.9121	0.0409	0.047
Nandi	0.0865	0.6649	0.2486	0.3135	0.9892	0	0.0108
Narok	0.1247	0.6925	0.1828	0.3518	0.9612	0.0332	0.0055
Nyamira	0.0633	0.7046	0.2321	0.4726	0.9536	0.0295	0.0169
Nyandarua	0.0986	0.6901	0.2113	0.3474	0.8685	0.0986	0.0329
Nyeri	0.0697	0.7212	0.2091	0.3394	0.9303	0.0606	0.0091
Samburu	0.1402	0.6646	0.1951	0.5183	0.9878	0.0061	0.0061
Siaya	0.0917	0.7152	0.1931	0.4452	0.9656	0.0196	0.0147
Taita Taveta	0.0947	0.568	0.3373	0.3373	0.8994	0.0118	0.0888
Tana River	0.1538	0.5962	0.25	0.4423	0.9808	0.0192	0
Tharaka-Nithi	0.1752	0.635	0.1898	0.3759	0.9891	0.0073	0.0036
Trans Nzoia	0.1115	0.6842	0.2043	0.3437	0.9505	0.0402	0.0093
Turkana	0.1893	0.6073	0.2034	0.5056	0.9944	0	0.0056
Uasin Gishu	0.0656	0.7287	0.2057	0.4117	0.9558	0.0318	0.0124
Vihiga	0.0409	0.6654	0.2937	0.4275	0.9591	0.026	0.0149
Wajir	0.0988	0.5185	0.3827	0.4012	0.963	0.0309	0.0062
West Pokot	0.1751	0.6162	0.2088	0.3838	0.9966	0.0034	0

ANNEX 6

Shows the probability of the TB treatment outcome

County	Cured	Dead	Defaulted	Failure	Trtmnt compltd
Nairobi	0.0261	0.0055	0.0034	0.0003	0.0477
Baringo	0.0248	0.0186	0	0	0.0186
Bomet	0.1338	0.0292	0.0146	0	0.1922
Bungoma	0.1113	0.041	0.0293	0.0059	0.1465
Busia	0.0278	0.043	0.0051	0.0025	0.0557
Elgeyo-Marak	0.2672	0.0992	0.0305	0	0.3817
Embu	0.1647	0.0449	0.0269	0.009	0.2784
Garissa	0.1779	0.0235	0.0101	0.0034	0.3221
Homa Bay	0.097	0.0521	0.0154	0.0024	0.2899
Isiolo	0.0513	0.0128	0	0	0.0256
Kajiado	0.0301	0.0137	0	0.0027	0.0137
Kakamega	0.0739	0.0326	0.0155	0	0.11
Kericho	0.216	0.0331	0.0233	0	0.3191
Kiambu	0.0871	0.0303	0.0049	0.0029	0.1419
Kilifi	0.168	0.0428	0.046	0.0032	0.3328
Kirinyaga	0.274	0.0585	0.0258	0.0117	0.3021
Kisii	0.3963	0.0705	0.056	0.0021	0.4066
Kisumu	0.1394	0.0673	0.018	0.0012	0.2127
Kitui	0.4184	0.0745	0.0177	0.0106	0.3599
Kwale	0.1592	0.051	0.0096	0	0.2197
Laikipia	0.1549	0.0282	0	0.0047	0.1315
Lamu	0.2167	0.0167	0.0333	0	0.3333
Machakos	0.2327	0.0437	0.0183	0.0071	0.2102
Makueni	0.146	0.0199	0.0044	0.0044	0.1969
Mandera	0.2051	0.0256	0	0	0.4038
Marsabit	0.0184	0.0061	0	0	0.0123
Meru	0.2393	0.0258	0.0225	0.0022	0.2124
Migori	0.0814	0.0235	0.0047	0	0.2394
Mombasa	0.1424	0.0217	0.0148	0.0017	0.1823
Murang'a	0.3211	0.0619	0.0406	0.0058	0.3578
Nakuru	0.1149	0.0313	0.0209	0.0026	0.2202
Nandi	0.2432	0.1135	0.0324	0	0.227
Narok	0.133	0.0249	0.0222	0.0055	0.2382
Nyamira	0.1097	0.0464	0.038	0	0.2025
Nyandarua	0.2911	0.0704	0.0329	0.0047	0.4272
Nyeri	0.1727	0.0364	0.0121	0.0061	0.2697
Samburu	0.0549	0.0244	0.0366	0	0.0915
Siaya	0.108	0.0622	0.0115	0.0033	0.1751
Taita Taveta	0.2899	0.0947	0	0	0.2426
Tana River	0.1827	0.0096	0.0096	0	0.2308
Tharaka-Nithi	0.0657	0.0109	0.0036	0	0.062
Trans Nzoia	0.1486	0.0372	0.0248	0	0.291
Turkana	0.0424	0.0085	0.0198	0	0.1045
Uasin Gishu	0.0495	0.0106	0.0053	0	0.0777
Vihiga	0.0818	0.0297	0	0.0074	0.1041
Wajir	0.4074	0.0432	0.0185	0	0.4815
West Pokot	0.0572	0.0101	0.0236	0	0.0808

REFERENCES

A. Schatz, E. B. (1944). Streptomycin, a substance exhibiting antibiotic activity against Gram-positive and Gram-negative bacteria. *Proceedings of the Society for Experimental Biology and Medicine, 55*, 66-69.

Alencar CH, R. A. (2012, April). Clusters of leprosy transmission and of late diagnosis in a highly endemic area in Brazil: focus on different spatial analysis approaches. *Tropical Medicine and International Health, 4*, pp. 518-525.

Alvarez-Hernandez, G. L.-V.-C.-P. (2010). An analysis of spatial and socio-economic determinants of tuberculosis in Hermosillo, Mexica, 2000-2006. *International Journal of Tuberculosis and Lung Disease, 14*, 708-713.

Annesi, I., Moreau, T., & Lellouch, J. (1989). Efficiency of the Logistic Regression and Cox Proportional Hazards Models in Longitudinal Studies. *Statistics in Medicine*(12).

Archie C. A. Clements, N. J. (2006). Bayesian spatial analysis and disease mapping: tools to enhance planning and implementation of a schistosomiasis control programme in Tanzania. *Tropical Medicine and International Health, 11*(4), 490-503.

Bakker MI, H. M. (2004 , Dec). Population survey to determine risk factors for Mycobacterium leprae transmission and infection. *International Journal of Epidemiology, 6*, pp. 1329-1336.

Bakker MI, H. M. (2006). Risk factors for developing leprosy—a population-based cohort study in Indonesia. *Leprosy Review, 77*, 48-61.

Ballot, D. E., Chirwa, T. F., & Cooper , P. A. (2010). Determinants of survival in very low birth weight neonates in a public sector hospital in Johannesburg. *BMC Pediatrics*(30).

Belete Getahuna, G. A. (2013). Treatment outcome of tuberculosis patients under directly observed treatment in Addis Ababa, Ethiopia. *The Brazilian Journal of Infectious Diseases, 17*(5), 521-528-.

Bendel RB, A. A. (1977). Comparison of stopping rules in forward regression. *Journal of American Statistical Association, 72*, 46-53.

Besag, J., York, J., & Mollie, A. (1991). Bayesian Image Restoration with Two Applicationa in Spatial Statistics. *Annals of the Institute of Statistical Mathematics, 43*(1), 1-59.

Beyers N, G. R. (1996). The use of a geographical information system (GIS) to evaluate the distribution of tuberculosis in a high-incidence community. *South African Medical Journal, 1*, 40-44.

Beyers, N., Gie, R., Zietsman, H., Kunneke, M., Hauman, J., Donald, P., & M.Tatley. (1996). The use of a geographical information system (GIS) toevaluate the distribution of tuberculosis in a high-incidence community. *South African Medical Journal*(1), pp. 40-44.

Bhatti, N. L.-G. (1995). Increasing incidence of tuberculosis in England and Wales:a study of the likely causes. *British Medical Journal, 310*, 967-969.

Bivand, R. S., Gomez-Rubio, V., & Rue, H. (2015, January). Spatial Data Analysis with R-INLA with some extensions. *Journal of Statistical Software, 63*(20).

Borgd, M. W. (2004). New Measurable Indicator for Tuberculosis. *Emerging and infectious diseases journal, 10*(9).

Bottoms, S. F., Paul , R. H., Iams, J. D., Merce, B. M., Thom, E. A., Roberts, J. M., . . . McNellis, D. (1997). Obstetric determinants of neonatal survival: Influence of willingness to perform cesarean delivery on survival of extremely low-birth-weight infants. *American Journal of Obstetrics and Gynecology, 176*(5).

Brianti E, D. M. (2007). Use of a health information system (HIS) for the epidemiological surveillance of leishmaniasis in urban areas. *Veterinary Research Communications, 31*, 213–215.

Bruce J. Kirenga, J. L. (2014, March 7). Treatment Outcomes of New Tuberculosis Patients Hospitalized in Kampala, Uganda: A Prospective Cohort Study. *PLoS One, 9*(3).

Callas, P. W., Pastides, H., & Hosmer, D. W. (1998). Empirical Comprisons of Proportional Hazards, Poisson, and Logistic Regression Modelling of Occupational Cohort Data. *American Journal of Industrial Medicine*(1).

Carlos, A. N., & Marilia, B. G. (2013). Low birth weight: causes and consequences. *Diabetology & Metabolic Syndrome Journal*.

Cave, A. (1939, July). The evidence for the incidence of tuberculosis in ancient Egypt. *British Journal of Tuberculosis, 33*(3), 142-152.

CDC. (2006, March 24). Emergence of Mycobacterium tuberculosis with Extensive Resistance to Second-Line Drugs — Worldwide, 2000–2004. *Morbidity and Mortality Weekly Report, 55*(11), pp. 301-308. Retrieved from http://www.cdc.gov/mmwr/pdf/wk/mm5511.pdf

Chan-Yeung, M., Yeh, A. G., Tam, C. M., Kam, K. M., Leung, C. C., Yew, W. W., & Lam, C. W. (2005, December). Socio-demographic and geographic indicators and distribution of tuberculosis in Hong Kong: a spatial analysis. *The International Journal of Tuberculosis and Lung Disease, 9*(12), pp. 1320-1326.

Charlotte Lewden, F. R.-L. (2002). Factors Associated with Mortality in Human Immunodeficiency Virus Type 1–Infected Adults Initiating Protease Inhibitor–Containing Therapy: Role of Education Level and of Early Transaminase Level Elevation (APROCO–ANRS EP11 Study). *The Journal of Infectious Diseases*, 710-714.

Chen, J., Roth, R. E., Naito, A. T., Lengerich, E. J., & MacEachren, A. M. (2008). Geovisual Analytics to Enhance Spatial Scan Statistic Interpretation:An Analysis of U.S Cervical Cancer Mortality. *International Journal of Health Geographics, 7*(57).

Christopher J L Murray, R. M.-A. (2015, August 26). Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990–2013: quantifying the epidemiological transition. *The Lancet*. Retrieved September 1, 2015, from http://dx.doi.org/10.1016/S0140-6736(15)61340-X

Collaboration, C. (2006, March 21). Effective therapy has altered the spectrum of cause-specific mortality following HIV seroconversion. *Official Journal of the AIDS Society*, pp. 741-749.

Collett, D. (2003). *Modelling Survival Data in Medical Research.* Chapman and Hall/CRC.

Cromley, E., & McLafferty, S. (2002). *GIS and Public Health.* New York: Guilford Press.

Cui Hua Liu, L. L. (2011, April 29). Characteristics and Treatment Outcomes of Patients with MDR and XDR Tuberculosis in a TB Referral Hospital in Beijing: A 13-Year Experience. *PLoS One, 6*(4).

D. Morse, D. B. (1964). Tuberculosis in ancient Egypt. *The American review of respiratory disease, 90*, 5224–5541.

D.Y.Lin. (1997). Non-parametric inference for cumulative incidence functions in competing risks studies. *Stotistics in Medicine, 16*.

Daniel Chukwunweolu Oshi, S. N. (2014). Profile, Outcomes, and Determinants of Unsuccessful Tuberculosis Treatment Outcomes among HIV-Infected Tuberculosis Patients in a Nigerian State. *Tuberculosis research and treatment*.

Daniel, T. (2000). The origins and precolonial epidemiology of tuberculosis in the Americas: can we figure them out? *International Journal of Tuberculosis and Lung Disease, 4*, 395-400.

Daniel, T. M. (2006, November). The History of Tuberculosis. *Respiratory Medicine, 100*(11), 1862–1870.

Danielle A. Southern, P. D. (2006). Kaplane-Meier methods yielded misleading results. *Journal of Clinical Epidemiology, 59*(10).

David G. Kleinbaum, . K. (n.d.). *Survival analysis: A self learning text.* Springer.

Derek Gregory, R. J. (2009). Medical Geography. In *The Dictionary of Human Geography, 5th Edition* (FIFTH ed., pp. 451-453). Wiley and Blackwell.

Dermot Maher, M. R. (2005, June 26). Global epidemiology of tuberculosis. *Clinic in Medicine*, pp. 167-182.

Dermot Maher, M. R. (2005, June). Global Epidemiology of Tuberculosis. *Clinics in Chest Medicine, 26*(2), 167-182.

Despina Kotsanas, K. R. (2007). Risk factors for peritoneal dialysis-related peritonitis: Can we reduce the incidence and improve patient selection? *Nephrology,* 239-245.

DLTLD. (2012). *Annual Report.* NAIROBI: MINISTRY OF HEALTH.

DLTLD. (2013). *TB Treatment Guidelines.* Nairobi.

DLTLD, (. o. (2012). *Annual Report.* Nairobi: Ministry of Health Kenya.

Dreyfuss, M. L., Msamanga, G. I., Spiegelman, D., Hunter, D. J., Urassa, E. J., Hertzmark, E., & Fawze, W. W. (2001). Determinants of low birth weight among HIV-infected pregnant women in Tanzania. *The American Journal of Clinical Nutrition, 74*(6).

Dye C, F. K. (n.d.). Disease control priorities in developing countries. New York: Oxford University Press.

E. Rolland Dickson, P. M. (1989). Prognosis in primary biliiary cirrhosis:Model for decision making. *Hepatology*(10).

EAJ Fischer, P. S. (2008, Sep). The spatial distribution of leprosy cases during 15 years of a leprosy control program in Bangladesh: An observational study. *BMC Infectious Diseases, 26*.

Edine W. Tiemersma, M. J. (2011). Natural History of Tuberculosis: Duration and Fatality of Untreated Pulmonary Tuberculosis in HIV Negative Patients: A Systematic Review. *6*(4).

Elizabeth L. Corbett, C. J. (2003, May 12). The Growing Burden of Tuberculosis Global Trends and Interactions With the HIV Epidemic. *Archives of Internal Medicine,* pp. 1009-1021.

Fatoumata Dicko, S. D.-M. (2014). Reasons for hospitalization in HIV-infected children in West Africa. *Journal of the International AIDS Society*, 18818.

FC, T. (2002). The application of GIS technology to equitably distribute fieldworker workload in a large, rural South African health survey. *Tropical Medicine and International Health, 7*, 80-90.

Field, D. J., Dorling, S. J., Manktelow, B. N., & Draper, E. S. (2008). Survival of Extremely Premature Babies in a Geographically Defined Population: Prospective Cohort Study of 1994-9 Compared with 2000-5. *British Medical Journal*(1221).

Fischer E, P. D. (2008). The spatial distribution of leprosy cases during 15 years of a leprosy control program in Bangladesh: an observational study. *BMC Infectious Diseases, 8*, 126.

Gebretsadik Berhe, F. E. (2012). Treatment outcome of smear-positive pulmonary tuberculosis patients in Tigray Region, Northern Ethiopia. *BMC Public Health, 12*, 537.

Gelmanova, I. S. (2007). Barriers to successful tuberculosis treatment in Tomsk, Russian Federation: non-adherence, default and the acquisition of multidrug resistance. In S. K. IY Gelmanova, *Bulletin of the World Health Organization* (9 ed., Vol. 85, pp. 703-711). World Health Organization. Retrieved September 2015

Goodarz Kolifarhood, D. K.-Z. (2015, September). Spatial and non-spatial determinants of successful tuberculosis treatment outcomes: An implication of Geographical Information Systems in health policy-making in a developing country. *Journal of Epidemiology and Global Health, 5*(3), 221-230.

Guy Harling, M. C. (2014). A spatial analysis of socia and economical determinants of tuberculosis in Brazil. *Health and Place, 25*, 56-67.

Guyatt, H. L., & Snow, R. W. (2001). Malaria in Pregnancy as an Indirect Cause of Infant Mortality in Sub-Saharan Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*(6).

Guyatt, H. L., & Snow, R. W. (2004). Impact of Malaria during Pregnancy on Low Birth Weight in Sub-Saharan Africa. *Clinical Microbiology Reviews*(4).

Håvard Rue, S. M. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society, 71*, 319-392.

Hayes, J. (2013, March 22). Tuberculosis: A Persistent Cough. Retrieved from http://www.healthmap.org/site/diseasedaily/article/tuberculosis-persistent-cough-32213#sthash.VrRx40PY.dpuf

Hayman, J. (1984). Mycobacterium ulcerans: an infection from Jurassic time? *The Lancet*, 1015-1016.

Health, M. o. (2008). *National manual for the management of HIV-related opportunistic infections and conditions. A healthcare workers manual.* Nairobi: Ministry of Health.

Hempel, S. (2006). *The Medical Detective: John Snow and the Mystery of Cholera.* London: Granta Books.

Hippocrates. (1982). *The aphorisms of Hippocrates with a Translation into Latin, and English* (Special ed edition ed.). Birmingham: Classics of Medicine Library.

Hoeven TA, F. E. (2008). Social distance and spatial distance are not the same, observations on the use of GIS in leprosy epidemiology. *Epidemiology and Infection, 136*, 1624–1627.

Hosmer, S., & and Lemeshow, D. (1999). *Applied Survival Analysis: Regression Modelling of Time to Event Data.* New York: Wiley.

Hosmer, S., & and Lemeshow, D. (2000). *Applied Logistic Regression.* New York: Wiley.

Huang L, K. M. (2007). A spatial scan statistic for survival data. *iometrics, 63*, 109-118.

Hyun J Lim, X. Z. (2010). Methods of competing risks analysis of end-stage renal disease and mortality among people with diabetes. *BMC Medical Reasearch Methodology*, 97.

Hyun J Lim, X. Z. (2010). Methods of competing risks analysis of end-stage renal disease and mortality among people with diabetes. *BMC Medical, 10*(97).

Ingram, D. D., & Kleinman, J. C. (1989). Empirical Comparisons of Proportional Hazards and Logistic Regression Models. *Statistics in Medicine*(5).

J. O'Grady, M. M. (2011). Tuberulosis in prisons: anatomy of global neglect. *European Respiratory Journal, 38*(4), 752-754.

Jackson, M. I. (2006). *Why Do Low Borth Weight Children do Worse in School? Understanding the Link Between Infant Health and Education.* Montreal: American Sociological Association, Annual Meeting.

Jan Beyersmann, M. D. (2007). A competing risks analysis of bloodstream infection after. *Statistics in Medicine, 26*.

Jia, Z. J. (2008). Spatial analysis of tuberculosis cases in migrants and permanent residents, Beijing, 2000-2006. *Emerging Infectious Diseases, 14*, 1413-1419.

John P Klein, R. B. (2003). Inference for Competing Risks. In Elsevier, *Handbook of Statistics* (pp. 291-311). Elsevier.

Josafá Gonçalves Barreto, D. B.-P. (2014). Spatial Analysis Spotlighting Early Childhood Leprosy Transmission in a Hyperendemic Municipality of the Brazilian Amazon Region. *PLoS Neglected Tropical Diseases., 12*.

José Wilton Queiroz, G. H.-N. (2010). Geographic Information Systems and Applied Spatial Statistics Are Efficient Tools to Study Hansen's Disease (Leprosy) and to Determine Areas of Greater Risk of Disease. *The American Journal of Tropical Medicine and Hygiene, 2*, pp. 306-314.

Julian Besag, J. N. (1991). The Detection of Clusters in Rare Diseases. *Journal of the Royal Statistical Society. Series A (Statistics in Society), 154*(1), 143-155.

K Tocque, M. R. (1999, March 1). Social factors associated with increases in tuberculosis notifications. *European Respiratory Journal, 13*(3).

Kalimba, E. M., & Ballot, D. E. (2013). Survival of extremely low-birth-weight infants. *The South African Journal of Child Health, 7*(1).

Katz, J., Lee, A. C., Kozuki, N., Lawn, J. E., Cousens, S., Blencowe, H., & Ezzati, M. (2013). Mortality Risk in Preterm and Small-for-gestational-age Infants in Low-income and Middle-income Countries: A Pooled Country Analysis. *The Lancet, 382*(9890).

Kenya, M. o. (2013). *Guidelines for Management of Tuberculosis and Leprosy in Kenya.* Nairobi: Ministry of Health Kenya.

Kistemann, T. M. (2002). Spatial patterns of tuberculosis incidence in Cologne (Germany). *Social Science and Medicine, 55*, 7-19.

Klebanoff, M. A., Schulsinger, C., Mednick, B. R., & Secher, N. J. (1997). Preterm and Small for Gestational-age Birth Across Generations. *American Journal of Obstetrics and Gynecology*(3).

Klein, J. P. (2006). Modelling competing risks in cancer studies. *Statistics in Medicine, 25*(6).

KNBS. (2009). *KENYA POPULATION AND HOUSING CENSUS.* Nairobi.

KNBS. (2014). *Kenya Demographic and Health Survey.* Nairobi: KNBS Kenya National Bureau of Statistics.

Knox, E. (1989). Methodology of enquiries into Disease Clustering. In *Detection of Clusters* (pp. 17-20). London.

Kramer, M. (1987). Determinants of low birth weight: Methodological assessment and meta-analysis. *Bulletin of the World Health Organization, 65*(5), pp. 663-737.

Kulldorf M, N. N. (1995). Spatial Disease Clusters: Detection and Inference. *Statistics in Medicine, 8*, pp. 799-810.

Kulldorff M, F. E. (1997). Breast Cancer Clusters in the NorthEastern United States: A geographic analysis. *American Journal of Epidemiology, 2*, pp. 161-170.

Kulldorff M, M. F. (2007). Multivariate scan statistics for disease surveillance. *Statistics in Medicine, 26*, 1824-1833.

Kulldorff, M., Gregory, D., Samociuk, H., & DeChello, L. (2006). Cancer Map Patterns: Are they Random or Not. *American Journal of Preventive Medicine, 30*, 537-549.

Kullford, M., & Nagarwalla, N. (1995). Spatial Disease Clusters:Detection and Inference. *Statistics in Medicine*, 799-810.

Laetitia Teixeira, A. R. (2013). Modelling competing risks in nephrology. *14*(110).

Laetitia Teixeira, A. R. (2013). Modelling competing risks in nephrology research: an example in peritoneal dialysis. *BMC Nephrology*(14).

Lance A. Waller, C. A. (2004). *Applied Statistics for Public Health Data.* JOHN WILEY & SONS, INC.

Laurent X. Nouvel, E. K.-K. (2006, September). Multidrug-resistant Mycobacterium tuberculosis, Bangui, Central African Republic. *Emerging Infectious Diseases*, pp. 1454-1456.

Lawn, J. E., Gousons, S., & Zupan, J. (2005). 4 million neonatal deaths: when? Where? Why? *Lancet*(9462).

Lawson, A. (2008). *Bayesian disease mapping: hierarchical modelling in spatial epidemiology.* Chapman and Hall.

Lee, D. (2013). CARBayes: An R Package for Bayesian Spatial Modeling with Conditional Autoregressive Priors. *Journal of Statistical Software, 55*(13).

Lehana Thabane, L. M. (2013). A tutorial on sensitivity analyses in clinical trials. (13). Retrieved January 21, 2015, from http://www.biomedcentral.com/1471-2288/13/92

Lehana Thabane, L. M. (2013). A tutorial on sensitivity analyses in clinical trials:the what, why, when and how. *BMC Medical Research Methodology, 13*(92). Retrieved from http://www.biomedcentral.com/1471-2288/13/92

Li, T. H. (2011). Impact of new migrant populations on the spatial distribution of tuberculosis in Beijing. *International Journal of Tuberculosis, 15*, 163-168.

Liu, L., Cousens, S., Lawn, J. E., & Black, R. E. (2012). Global regional and national causes of child mortality. *Lancet, 380*(9853).

Lubchenco, L. O., Brazie, J. V., & Searls, D. T. (1972). Neonatal Mortality Rate: Relationship to Birth Weight and Gestational Age. *Pediatrics, 81*.

MacIntyre C. R., K. N. (1997). Impact of tuberculosis control measures and crowding on the incidence of tuberculous infection in Maryland prisons. *Clinical Infectious Diseases, 24*(6), 1060–1067.

Madeleine Okome-Nkoumou, V. G.-E. (2014). Opportunistic Diseases in HIV-Infected Patients in Gabon following the Administration of Highly Active Antiretroviral Therapy: A Retrospective Study. *American Journal of Tropical Medicine and Hygiene*, 211-215.

MARCOS A. ESPINAL, D. A. (2001, April). GLOBAL TRENDS IN RESISTANCE TO ANTITUBERCULOSIS DRUGS. *New EnglandJournal of Medicine, 344*(17), pp. 1294-1303.

María Belén Herrero, S. A. (2015, August). Spatial analysis of the tuberculosis treatment dropout, Buenos Aires, Argentina. *Rev Saúde Pública, 49*.

Marion Verduijn, D. C. (2011). The analysis of competing events like cause-specific mortality—beware of the Kaplan-Meier Method. *Nephrology Dialysis Transplant*, pp. 56-61.

Martin Kulldorff, R. H. (2005). A Space–Time Permutation Scan Statistic for Disease Outbreak Detection. *PLoS Medicine, 3*.

Mayer, J. D. (2000). Geography, ecology and emerging infectious diseases. *Social Science and Medicine*, 937-952.

McCowan, M. E., Dekker, A. G., Chan, E., Chappell, C. L., Hunter, M., & Moss-Morris, R. (2009). Spontaneous Preterm Birth and Small for Gestational Age Infants in Women who Stop Smoking Early in Pregnancy: Prospective Cohort Study. *British Medical Journal*(b1081).

Melo, T. R.-S. (2012). Cadernos de Saude Publica. *Spatial and temporal distribution of tuberculosis in indeginous of Rondonia State, Western Amozon, Brazil, 28*, 267-280.

Mickey RM, G. S. (1989). The impact of confounder selection criteria on the efffect estimation. *American Journal of Epidemiology, 129*(1), 125-137.

Moriguchi, S., Hayashi, Y., Maehara, Y., Nose, Y., Korenaga, D., & Sugimachi, K. (1993). A comparison of the logistic regression and the Cox proportional hazard models in retrospective studies of patients with gastric cancer. *Journal of Surgical Oncology*(1).

Munch, Z. V. (2003). Tuberculosis transmission patterns in a high incidence are:a spatial analysis. *International Journal of Tuberculosis and Lung Disease, 7*, 271-277.

Naus, J. (1965). The Distribution of the Size of Maximum Cluster of Points on a Line. *Journal of the American Statistical Association*, 532-538.

NTLD. (2013). *Guidelines for Management of Tuberculosis and Leprosy in Kenya.*

Olsson, U. (2002). *Generalized Linear Models An Applied Approach.* Sweden: Studentlitteratur.

P. Elliott, J. W. (2000). *Spatial epidemiology: methods and applications.* Oxford University Press.

Padmanesan Narasimhan, J. W. (2013). Risk Factors for Tuberculosis. *Pulmonary Medicine*. Retrieved from http://dx.doi.org/10.1155/2013/828939

Parslow, R. E.-S. (2001). Tuberculosis deprivation, and ethnicity in Leeds, UK1982-1997. *Archives of disease in childhood, 84*, 109-113.

Penna ML, d. O. (2009, September). The epidemiological behaviour of leprosy in Brazil. *Leprosy Review, 3*, pp. 332-344.

Penna, G. P. (2009). High incidence of diseases endemic to the Amazon region of Brazil, 2001–2006. *Emerging Infectious Diseases, 15*, 626-632.

Queiroz JW, D. G.-N. (2010 , Feb). Geographic information systems and applied spatial statistics are efficient tools to study Hansen's disease (leprosy) and to determine areas of greater risk of disease. *The American Journal of Tropical Medicine and Hygiene, 2*, pp. 306-314.

Ramalingaswami, V., Urban , J., & Jon , R. (1996). The Asian enigma. *The Progress of Nations Nutrition*. UNICEF.

Randremanana, R. S. (2009). Spatial clustering of pulmonary tuberculosis and impact of the care factors in Antananarivo City. *Tropical Medicine and International Health, 14*, 429-437.

Ravi Varadhan, C. O. (2010, June). Evaluating Health Outcomes in the Presence of Competing Risks A Review of Statistical Methods and Clinical Applications. *Medical care, 48*(6).

Ravi, V., Weiss, C. O., Segal, J. B., Wu, A. W., Scarfstein, D., & Boyd, C. (2010, June). Evaluating health outcomes in the presence of competing risks. *Medical Care, 48*(6).

Reichman, N. E. (2005). Low birth weight and school readiness. *The Future of Children*(1).

Ricketts, S. A., Murray, E. K., & Schwalberg, R. (2005). Reducing Low Birth Weight by Resolving Risks. *American Journal of Public Health, 95*(11).

Robin, G., Jean-Michel, P., & Christine, T.-M. (2010). Variable Selection using Random Forests. *HAL archives-ouvertes, 31*(14), 2225-2236.

Roger S. Bivand, V. G.-R. (2015). Journal of Statistical Software. *63*(20).

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian Inference for Latent Gaussian Models by Using Integrated Nested Laplace Approximations. *Journal of the Royal Statistical Society, 71*(2), 319-392.

Ryan, F. (1992). *The forgotten plague. How the battle against tuberculosis was won and lost.* Boston: Little, Brown and Company.

S. Banerjee, B. C. (2004). *Hierarchical modelling and analysis for spatial data.* Chapman and Hall.

Satagopan, J. M., Ben-Porat, L., Berwick, M., Robson, M., Kutler, D., & Auerbach, A. D. (2004). A note on competing risks in survival data analysis. *British Journal of Cancer, 91*.

Spence, D. H. (1993). Tuberculosis and poverty. *British Medical Journal, 307*, 759-761.

TA, G., W, L., J, C., & BE, S. (1999). Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Statistics in Medicine, 18*(6).

Talbot TO, K. M. (2000). Evaluation of spatial filters to create smoothed maps of health data. *Statistics in Medicine*, 2399-2408.

Talbot TO, K. S. (2008). Development of an interactive environmental public health tracking system for data analysis, visualization, and reporting. *Journal of Public Health Management and Practice, 14*, 526-532.

Tango, T., & Takahashi, K. (2005). A Flexibly Shaped Spatial Scan Statistic for Detecting Clusters. *International Journal of Health Geographics, 4*(11).

Taylor, S. L., Sen, S., Greenhalgh, D. G., Lawless, M., Curri, T., & Palmieri, T. L. (2015, March 11). A competing risk analysis for hospital length of stay in patients with burns. *JAMA Surgery*.

Tiwari N, A. C. (2006). Investigation of geo-spatial hotspots for the occurrence of tuberculosis in Almora district, India, using GIS and spatial scan statistic. *International Journal of Health Geographics, 5*, 33.

Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, pp. 234-240.

Tobler, W. (1977). Spatial interaction patterns. *Journal of Environmental Systems, 6*, 271–301.

Tuula Vasankari, P. H. (2007, October 14). Risk factors for poor tuberculosis treatment outcome in Finland: a cohort study. *BMC Public Health, 7*(291).

UNESCO. (2013). *UNESCO Institute for Statistics and data.*

USAID. (2008, April). Nutrition and Tuberculosis: a review of the literature and considerations for TB control programs. *Africas Health in 2010*.

Videlis Ndubaa, A. H. (2015, June). Prevalence of tuberculosis in adolescents, western Kenya: implications for control programs. *International Journal of Infectious Diseases, 35*, 11-17.

Vikse, B. E., Irgens, L. M., Leivastad, T., Hallan, S., & Iversen, B. M. (2008). Low Birth Weight Increases Risk for End-Stage Renal Disease. *Journal of the American Society of Nephrology, 9*.

Waller LA, G. C. (1995). *Applied Spatial Statistics for Public Health Data.* Hoboken, NJ: John Wiley & Sons.

WHO. (1992). *International statistical classification of diseases and related health problem, tenth revision.* Geneva: WHO.

WHO. (1997). *The world health report 1997 - conquering suffering, enriching humanity.* Geneva: WHO.

WHO. (2004). *Towards the development of a strategy for promoting optimal fetal growth.* Geneva: WHO.

WHO. (2004). *Towards the Development of a Strategy for Promoting Optimal Fetal Growth.* Geneva: WHO Technical Consultation.

WHO. (2011). *Global Tuberculosis Report.*

WHO. (2015, March). Retrieved August 27, 2015, from http://www.who.int/mediacentre/factsheets/fs104/en/

WHO. (2015). *Global Tuberculosis Report.* Geneva: WHO.

WHO. (2015). *Health statistics and information systems*. Retrieved September 01, 2015, from Metrics: Disability-Adjusted Life Year (DALY) Quantifying the Burden of Disease from mortality and morbidity:
http://www.who.int/healthinfo/global_burden_disease/metrics_daly/en/

Wilcox, A. J. (2001). On the importance and the unimportance of birthweight. *International Journal of Epidemiology*(6).

Zimmerman, M. (1979). Pulmonary and osseous tuberculosis in an Egyptian mummy. *Bulletin of the New York Academy of Medicine,* pp. 604-608.

Zoran Bursac, C. H. (2008). Purposeful selection of variables in logistic regression. *Source code for Biology and Medicine, 3*(17).