****

**Research Proposal**

**The Role of Literacy Rate and Urban Population on Tuberculosis**

Prathamesh Sirsikar

Research Methods in Intelligence

Dr. Danzell

12 December 2016

**Department of Intelligence Studies**

****

***Introduction***

The most recent published estimates of global burden of tuberculosis (TB) are based on data available up to 2015, give figure for individual countries of all the world. The magnitude of the Tuberculosis has changed since 1990, due to changing control practices spread of Human Immunodeficiency Virus (HIV) and population growth. To estimate the risk and prevalence of Mycobacterium Tuberculosis (MTB) infections and Tuberculosis incidence, prevalence included Tuberculosis disease data comparing to Human Literacy Rate and Urban population, with all the countries of the world from 1992 to 2012. Some of the other variables like Unemployment, Education index, Population of each country, Population growth of that countries are also included to compare with tuberculosis data. With the increase in Literacy rate and urban population there is likely to be an effect on Tuberculosis cases in the country. Traditionally, tuberculosis and literacy rate of that country are linked through a handful of variables. This paper will try to find the results by using a hypothesis-generating case study to evaluate what sort of effect Literacy rate, Urban population, Unemployment have on Tuberculosis on all the countries.

At the time of doing this research, several problems across. There are many researches about Tuberculosis. Using different approaches and epidemiological apparatuses and instruments in determining the cause-effect relationship of the disease and to come up with effective alternative and ways to control and prevent the proliferation of the disease yet there is no ample study on finding the relationship between literacy rate and Tuberculosis. This study is significant in the sense that the new technological apparatus which is Literacy rate and Urban population will be the tool in identifying and supporting the claim of previous studies declaiming the direct relationship between urbanization and education with Tuberculosis. This will help the health sector create more effective programs to prevent the rampant spread of the disease. Almost all studies on Tuberculosis cases have focused on their relation to the HIV epidemic, medical facilities, medical research, Population effects on tuberculosis rather than Public education, Literacy rate, urban population, Unemployment; thus, studies on worldwide literacy rate, Unemployment from specifically were largely absent in the previous literature. To counter this absence, the scope of the research was broadened and worldwide research on Tuberculosis from several disciplines. There is also very little research focused on the qualitative evaluation of Tuberculosis background, Mortality rate and some medical research from 1882 to 2015. This study intends to remedy the lack of literature on the direct connection between literacy rate and urban population by showing that there is a need to empirically examine this issue utilizing a quantitative approach to evaluate the Tuberculosis cases.

There are some terms need to be defined to properly understand this paper. Literacy rate are defined as the total percentage of population of limited period who can read and write with understanding. Literacy rate is the important factor to represent the tuberculosis data because it effects the tuberculosis cases. Like high literacy rate countries has low number of tuberculosis cases and with the low literacy rate countries have the high amount of tuberculosis cases. Urban population defined as human settlement with high population and good infrastructure with all the facilities. Urban population is also important factor to represent the Tuberculosis data. With high Urbanization countries have low number of tuberculosis cases. Just because in urban areas, it is easy to get all the medical facilities as early as possible. With more than 3 billion of the world's population living in cities, and accelerating urban expansion arising from globalization, population growth and migration, today's urban settings are redefining the field of public health. The complex dynamics of cities, with their concentration of the poorest and most vulnerable, pose an urgent challenge to the health community. While retaining fidelity to the core principles of disease prevention and control, major adjustments are needed in systems and approaches to effectively reach those with the greatest health risks within today's urban environment.

This research proposal will be broken down into four parts: (1) a review of the current literature; (2) an overview of the theoretical framework of this proposal, (3) an examination of the proposed methodological approach and, (4) the conclusions and future expectations for this research. The literature review will cover the global Tuberculosis report: first, it will focus in on background of tuberculosis; it will then broaden to cover the links between tuberculosis cases on a worldwide level; it will conclude with the generation of hypotheses based on literacy rate and urban population. The methodology section will cover research design and methods. The final section will evaluate the methods used and the expected results as well as conclusions that can be made as well as propose future topics for research.

***Research and Literature Review***

Global databases consistently demonstrate the higher incidence, prevalence, mortality and burden of tuberculosis and other communicable diseases in developing countries. According to the work of Pillay (2004) entitled: Using GIS to spatially portray the prevalence of Tuberculosis and its demographic consequences in selected countries in Sub-Saharan Africa, TB cases was recorded in South Africa from the estimated 180 507 TB cases against a TB incidence of 419/100,000 population. The analysis goes with more numbers and statistical data, but what the author found out was that the spatial dynamics of the problem can be analyzed using the GIS techniques via choropleth mapping of HIV/AIDS and TB prevalence data. There are significant impacts which these diseases have to society. The diseases like this will have cultural, social and economic effects to one country with great number of victims.

*1882 TB bacilli identified by Robert Koch*

Robert Koch, a German physician and scientist, presented his discovery of *Mycobacterium tuberculosis*, the bacterium that causes tuberculosis (TB), on the evening of March 24, 1882. Today, more than a century after Koch discovered *Mycobacterium tuberculosis*, there are still many infected individuals and around two million deaths annually resulting from the disease. Koch continued his studies on tuberculosis, hoping to find a cure. In 1890, he announced the discovery of tuberculin, a substance derived from tubercle bacilli, which he thought was capable of arresting bacterial development in\_vitro and in animals. This news gave rise to tremendous hope throughout the world, which was soon replaced by disillusionment when the product turned out to be an ineffective therapeutic agent. Tuberculin later proved to be a valuable diagnostic tool.(Research on tb – Koch 1882)

*1907 TST - tuberculin skin test (von Pirquet)*

Diagnosis of tuberculosis was aided by discovery of the acid-fast nature of the bacillus by Ehrlich in 1882, discovery of X rays by Roentgen in 1895, development of the tuberculin skin test by Von Pirquet and Mantoux in 1907-1908, and preparation of purified protein derivative (PPD) of tuberculin by Seibert in 1931. In the 1930s, the epidemiologic work of Wade Hampton Frost led to a better understanding of the epidemiology of tuberculosis. In the 1940s, using Seibert's PPD administered by the Mantoux method and chest X-ray examinations, the United States Public Health Service began a series of studies that elucidated further the epidemiology of tuberculosis and made apparent the distinction between tuberculous infection without disease (a positive skin test in the absence of signs and symptoms) and overt clinical tuberculosis. .(Research on tb - koch)

*1948 BMRC trial of streptomycin vs bed rest and 1952 Development of isoniazid*

Treatment has progressed from bed rest, special diets and fresh air, through pneumothorax and other lung-collapse procedures and surgical resection, to specific chemotherapy (streptomycin in 1947, para-aminosalicylic acid in 1949, isoniazid in 1952, and drugs such as rifampin in recent years). With combinations of modern drugs properly administered, tuberculosis is now virtually 100% curable. .(Research on tb - koch)

In the century since Koch's discovery, advances in prevention, diagnosis, and treatment of tuberculosis--especially treatment--have produced a spectacular decline in tuberculosis mortality and a striking decline in tuberculosis morbidity--primarily in technically advanced countries (Figure 1). Progress has been less dramatic in developing countries. Tuberculosis stubbornly persists as a major worldwide health problem. It is estimated that as many as 10 million cases of tuberculosis may occur throughout the world each year--4-5 million of them highly infectious, and 2-3 million resulting in death. Eradication of tuberculosis, although possibly attainable in technical terms, remains an elusive goal. Reported by Tuberculosis Control Div, Center for Prevention Svcs, Mycobacteriology Br, Bacterial Diseases Div, Center for Infectious Diseases, CDC.

The World Health Organization defines Tuberculosis as “an infectious disease caused by the bacillus Mycobacterium tuberculosis” (WHO Report). The disease generally affects the lungs (known as pulmonary tuberculosis), but M. tuberculosis can infect any organ within the body (Ringold et al. 464). Pulmonary TB is the most common form of the disease, with 80-85% of all deaths from TB in the twentieth century caused by the pulmonary TB. TB is an airborne disease, and when those with an active TB infection “cough, sneeze, talk or spit, they propel [TB bacilli in water droplets] into the air. Symptoms of an active pulmonary tuberculosis infection include fever, night sweats, weight loss, fatigue, and a “persistent cough that may be associated with bloody sputum” as well as “difficulty breathing or chest pain with breathing” (Ringold et al. 464).

*Global Burden of Tuberculosis*

Most of the estimated number of cases in 2015 occurred in Asia (61%)2 and the WHO African Region (26%); smaller proportions of cases occurred in the Eastern Mediterranean Region (7%), the European Region (3%) and the Region of the Americas (3%). The 30 high TB burden countries3 accounted for 87% of all estimated incident cases worldwide. The six countries that stood out as having the largest number of incident cases in 2015 were (in descending order) India, Indonesia, China, Nigeria, Pakistan and South Africa (combined, 60% of the global total). Of these, China, India and Indonesia alone accounted for 45% of global cases in 2015. (WHO – 2016 report)

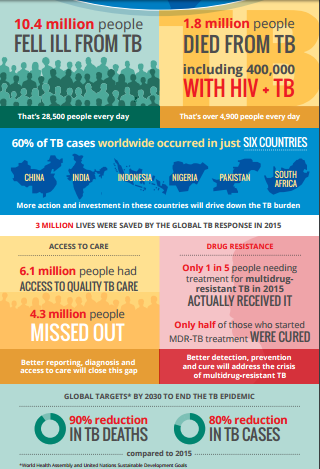


Figure 1 WHO report on Tuberculosis - 2015 Statistics

***Methodology***

Using the qualitative and quantitative approaches in research will help uncover the hidden clues about the relationship between Literacy rate and urban population with TB. The purpose of hypothesis-generating case studies is theory-building. They start with a vague notion of possible hypothesis, and attempt to formulate definite hypotheses to be tested subequently among a larger number of cases (Lijphart, 1971). Hypothesis-generating case studies aim to generalize beyond the data. They examine multiple cases to develop theoretical propositions that can be tested through other methods,(Levy, 2008). The theory built from this particular study is that Literacy rate and Urban population have an effect on Tuberculosis cases. This study tries to focus on Tuberculosis data for all the countries from 1992 to 2012. For that we have used Cross-national time-series analysis method to examine that data. The hypotheses generated will be able to be evaluated through Linear regression methods using statistical data collected from Official sources like World bank, World health Organization, Alesina et al. All estimates are generated by using Stata 14.1.

Hypothesis H1 will be evaluated individually for the occurrence of seven variables: Literacy rate, Population, Per Urban population, Unemployement, GDP Growth, Edu index, Country. The hypotheses will be tested using XTPCSE (Linear regression with panel-corrected standard errors) to look for statistical significance in the relationship between the hypotheses and the variables. XTPCSE calculates panel-corrected standard error (PCSE) estimates for linear cross-sectional time series models where the parameters are estimated by either OLS or Prais–Winsten regression. When computing the standard errors and the variance–covariance estimates, xtpcse assumes that the disturbances are, by default, heteroskedastic and contemporaneously correlated across panels. [ xtpcse tbdata literacy\_rate population per\_urbanpop unemployment gdp\_growth edu\_index, pairwise ]

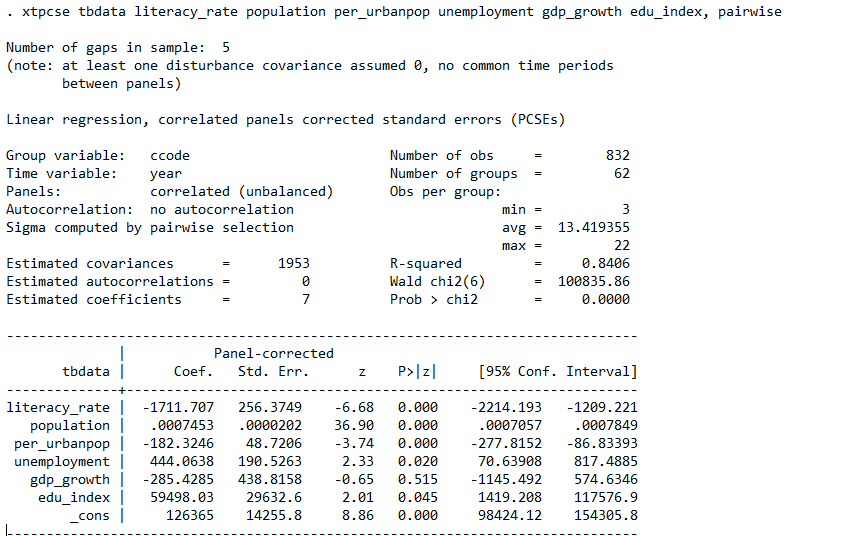


Figure 2 Tuberculosis Data compared with Literacy Rate

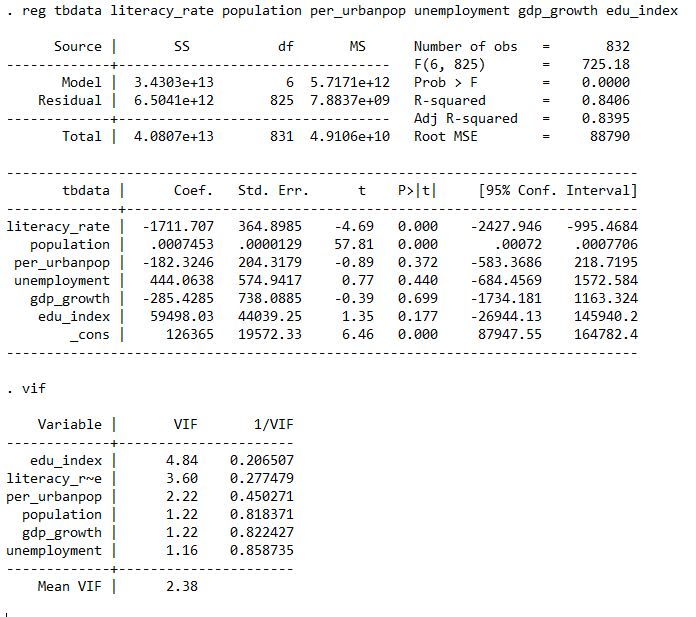


Figure 3 Hypothesis 1 by Using OLS linear regression

Findings of Hypotheses model 1 by using OLS regression:

Use the regress command for OLS regression (you can abbreviate it as reg). Specify the DV first followed by the IVs. By default, Stata will report the unstandardized (metric) coefficients. Use the vif command to get the variance inflation factors (VIFs) and the tolerances (1/VIF). vif is one of many post-estimation commands. You run it AFTER running a regression. It uses information Stata has stored internally.

Let's focus on the three predictors, whether they are statistically significant and, if so, the direction of the relationship. Literacy rate, Coef. = -1711.707 The coefficient is negative which would indicate that larger Literacy rate is related to lower Tuberculosis cases -- which is what we would expect.   Next, the Urban population (per\_urbanpop = -182.3246 , p=.372) is significant and its coefficient is negative indicating that the greater the Urban population, the lower the Tuberculosis cases.  Please note, that we are not saying that Literacy rate or Urban population causing Tuberculosis. Thus, higher levels of literacy rate and urbanpopulation are associated with lower tuberculosis. This result also makes sense.  Finally, the Unemployment (**coef.. =444.0638** , p=.440) this would seem to indicate that the unemployment is also an important factor in predicting Tuberculosis cases -- this result was somewhat unexpected.

For Hypothesis H2, we have generated one new veriable by using Literacy rate and Urban population and named that new veriable as urban\_literacy. Now hypothesis H2 will be examined individually for the occurrence of these sevan variables: Urban\_literacy, Population, Per Urban population, Unemployement, GDP Growth, Edu index, Country. The hypotheses will be tested using XTPCSE (Linear regression with panel-corrected standard errors) to look for statistical significance in the relationship between the hypotheses and the variables. XTPCSE calculates panel-corrected standard error (PCSE) estimates for linear cross-sectional time series models where the parameters are estimated by either OLS or Prais–Winsten regression. When computing the standard errors and the variance–covariance estimates, xtpcse assumes that the disturbances are, by default, heteroskedastic and contemporaneously correlated across panels.

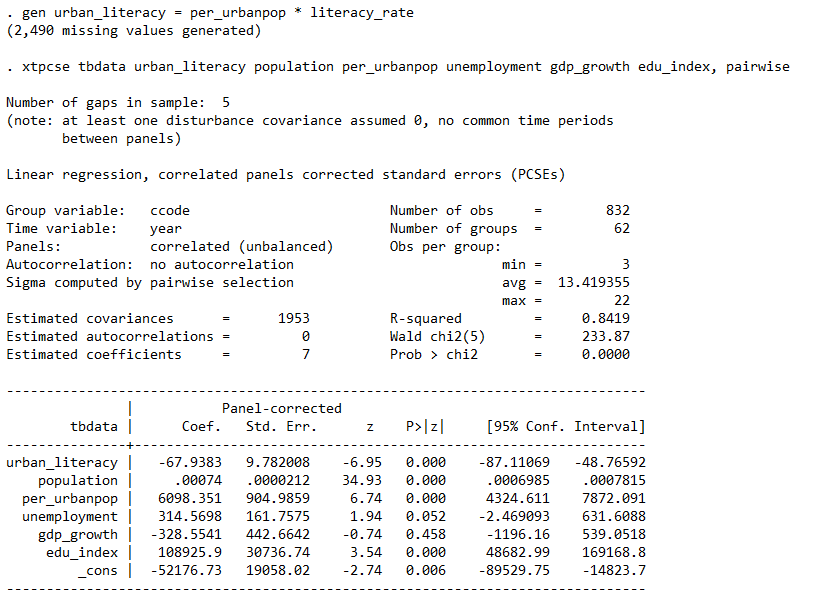


Figure 4 Tuberculosis Data Compared with Urban Population + Literacy Rate

The analysis will be done for two different types of case scenario. The first analysis will focus on the reported Tuberculosis cases because of the low literacy rate. The second analysis will use numbers generated from combining two veriables (Literacy rate + Urban population). The equation will be formed using the following factors: the number of Tuberculosis cases reported, Literacy rate of each country from 1992 to 2012, Urbanization percenatage of each country and how that each factor effects on each other.

Let's focus on the two predictors, whether they are statistically significant and, if so, the direction of the relationship. Urban\_literacy, Coef. = -67.9383 The coefficient is negative which would indicate that larger Literacy rate and urban population is related to lower Tuberculosis cases -- which is what we would expect.  Please note, that we are not saying that Literacy rate or Urban population causing Tuberculosis. Thus, higher levels of literacy rate + urbanpopulation are associated with lower tuberculosis. This result also makes sense.  Finally, the Unemployment (**coef.. = 314.5698** , p=.052) this would seem to indicate that the unemployment is also an important factor in predicting Tuberculosis cases.

***Expected Results and Conclusions***

The expected result of this study is that towards the end of the research, we may be able to underline and draw a clear connection between the Literacy rate, Urban population with Tuberculosis. The literature serves to generate two generalizable hypotheses as well as the potential for well a large number more specific hypotheses that could each have a case study dedicated to them. The purpose of this paper is to generate hypotheses for later quantitative testing; the generalized hypotheses are the best options. The specific hypotheses can be formed by looking at one aspect of literature rate, urban population and tuberculosis data, for example: What effects on tuberculosis more, literacy rate or Urban population? Or, Literacy rate and urbanization is really important to stop the tuberculosis? It is also possible to have intermediate hypotheses, for example: literacy rate and urban population are both as important to reduce the tuberculosis cases? In this regards, the research expected to fulfill its entire task from gathering data and information, to analysis and presentation and that all the objective and problems stated will be answered clearly.

Using multiple linear regression, these two hypotheses will be tested with the seven variables estimating the relationship between them. Statistical significance (p < .01) is expected between TBdata and Literacy rate, population, Unemployment, education index, urban population and country thereby supporting H1. H2 is expected to be accepted in regards to Urban\_literacy together, then unemployment, and education index, population and country.

Very important in this study is the fulfillment of providing a clear answer to the noble aim of this study. Moreover, to prove the claim of the study that there is a preexisting direct connection between Literacy rate, Urban population and how it effects on Tuberculosis and other health issues.

***References***

1. http://www.who.int/tb/country/en/
2. http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf?ua=1
3. Research - <http://www.stata.com/manuals13/xtxtpcse.pdf>
4. Koch R *- Research on tb . Die Atiologic der Tuberkulose. Berliner Klinische Wochenschrift* 1882; 15:221-30.

https://www.cdc.gov/mmwr/preview/mmwrhtml/00000222.htm

1. Wilsey, Ashley *“Half in Love with Easeful Death:” Tuberculosis in Literature , Pacific University, 5-2012*
2. <https://www.nobelprize.org/educational/medicine/tuberculosis/readmore.html> - 1882
3. <http://sampleresearchproposals.blogspot.com/2009/01/hivaids-and-tb-research-proposal.html>
4. Ringold, Sarah, Cassio Lynm, and Richard M. Glass*. "Tuberculosis the Journal of the American Medical Association”* 300.4 (2008): 464.