**Projection of Thalassemics in Khyber Pakhtunkhwa**

Saima Wazir, Zafar Zaheer

Institute of Management Sciences, Peshawar, Pakistan

**Abstract:**

*Thalassemia is a rising genetic disease in Khyber Pakhtunkhwa. To be able to initiate policy making for controlling Thalassemia, projection of the prevalence of the disease in population is an important factor. Therefore, projection of the affected population in the Khyber Pakhtunkhwa is the aim of this research. In the first step, of estimating the patients and carriers of the disease, population projection is done with Cohort Component Method of Analysis. Pakistan Demographic Surveys conducted by Pakistan Bureau of Statistics is used to collect the data for projection of the total population. Then the output of population projection is used to analyze the number of Thalassemics in the total population. Hardy-Weinberg Analysis, a technique adopted from genetics, is used to identify the number of Thalassemia Major and Thalassemia Minor Patients in the population. It is observed that with the increasing population the number of Thalassemia affected individuals is increasing in the region. This calls for attention of policy makers for a preemptive strategy to timely control the disease.*

**1. Introduction:**

Thalassemia is a genetic disorder in which the patients are dependent on blood transfusions and regular treatment for the rest of their lives. To make the matters worse, the disease is inherited from the parents and cannot be treated easily. The rate of Thalassemia varies in different regions, broadly, it ranges from 2.2%-16% of the total population in different countries around the world (Thalassemia International Federation, 2012). However, the increasing prevalence indicates that the population will not be able to shoulder the economic cost or the psychological burden of the life-long suffering inflicted due to Thalassemia. The time has now come to probe the population for prevention of Thalassemia in the future and involve public itself in seeking medical help (Rahman and Lodhi, 2004). Reports state that in the coming years 95% of the total Thalassemics around the world will be hosted by Asian, Indian and Middle Eastern regions (Weatherall and Clegg, 2001, Lorey 2000).

Therefore, this research is conducted to project the segment of population that is affected by Thalassemia. For this purpose, population is projected on a whole with the help of Cohort Component Analysis. After the population is projected, Hardy Weinberg Analysis is applied on the output to calculate the number of the affected individuals in the coming years. The same procedure can be used to calculate the proportion of Thalassemia affected population in the future for policy making.

Projection of such population segments that have high disease prevalence is of importance in ensuring that they are provided with such services and support as required to enable them to live a healthier life (Marshall, 2009). Projecting such instances can put the government on alert as they could direct the resources and services in such sectors efficiently and effectively (Siegel, 2002).

## 2. Projection of Population

There are a number of methods that are used around the world for projecting population for different purposes. Some of them are growth rate methods, time series analysis, microsimulation and cohort component method. One of the determining factors for choosing any of these methods is the purpose of projection.

The focus of this study is to project the thalassemia affected population of Khyber Pakhtunkhwa. For this purpose, first of all the total population of KP is calculated and then within that, the number of Thalassemics is estimated.

Cohort Component Model (CCM) is chosen to project the overall population. This model allows for cohorts (characteristics based groups) such as age and gender etc. while projecting the future population (Jiang, 2007). Once the projection is done with the help of CCM, a categorized population in subgroups of age-sex cohorts is achieved. Moreover, in this study the division of population into rural and urban residence is also done to facilitate the policy making in the future if required, for Thalassemic patients. Cohort Component method is also suitable here because it helps to project population in the future for short intervals of time i.e. in five year segments. This data can then be targeted for future population planning and other health policies.

The next step of analysis is to estimate the number of Thalassemia affected individuals in the population.. To ensure that updated information on the number of Thalassemic patients is available, a recent estimate of populations is required. However, in Pakistan the census data is available only until 1998, whereas Demographic Survey data is present up to 2007. Both these databases are out dated. Therefore, population is projected for 2012 and then prevalence of Thalassemia is measured among them. This is done to compare the projected data with the government estimate of the total population. It helps in validation of the data projected through the model.

Thalassemia, unlike other diseases, is genetically inherited. It cannot be adopted as a result of a harmful lifestyle, nor can it be avoided by a person’s own precautions. A person is born with the disease and is destined to a lifetime of treatment to avoid fatal consequences. Hence, the usual disease projection models cannot be applied here. Genetic model of estimation is most suitable here. One of the models that specifically focus on identification of genetic disease inheritance from parents to children in case of two alleles presence is Hardy Weinberg Equilibrium. It is suitable for the purpose Thalassemia can be inherited only in case double alleles are present in parents of the future Thalassemia patients. In order to calculate the number of Thalassemics in the projected population, Hardy Weinberg Analysis is used. Both the models i.e. Cohort Component Method for population projection and Hardy Weinberg Analysis for Thalassemic identification are explained below.

**2.1. Cohort Component Method**

Population projection with Cohort component Method is done with the help of the following formula:

Pt = Pt-1 + Bt-1,t - Dt-1,t + Mt-1,t

where:

Pt = population at time t;

Pt-1 = population at time t-1;

Bt-1,t = births, in the interval from time t-1 to time t;

Dt-1,t = deaths, in the interval from time t-1 to time t; and

Mt-1,t = net migration, in the interval from time t-1 to time t.

With the help of these variables the formula can be applied to find the population of 2012 and onwards with the population of 2007, as suggested by CCM guidelines. It is this model that is used to find the overall projected population. Within this, the estimated number of Thalassemic patients is calculated with the help of Hardy Weinberg equilibrium.

**2.2. Hardy Weinberg Equilibrium**

The equilibrium can be understood with the help of the following equation, which is fit for a population that has random mating patterns.

p2 + q2 + 2pq = 1

p in this equation report the normal population, q stands for the homozygotes (also referred to as Thalassemia Major patients), whereas, 2pq report the number of heterozygotes i.e. the carriers of a genetic disease.

In Pakistan however, random mating is not observed. This can measured with the help of coefficient of inbreeding that is 0.0455, for the overall country (Ain et al., 2011).

(p2 + Fpq) + (q2 + Fpq) + 2pq (1-F) = 1

With involvement of F i.e. coefficient of inbreeding, the measures for inheritance patterns varies. In this case the first factor (p2 + Fpq) measures the number of Thalassemia Major patients, whereas the last component i.e. 2Fpq gives results for Thalassemia carriers.

This is based on the theory that both the parents have to be Thalassemia carriers to pass on the disease, Thalassemia Major, to their children. Unless both the alleles are combined, the genetic mutation is not expressed in the offspring of the carriers (Ahmed, 1998).

**2.3. Data Collection for Population Projection**

For projection of population secondary data is collected for application of CCM. Data on following variables is collected:

* Stationary population divided per age-sex cohort
* Number of deaths
* Number of children born
* Number of women in the reproductive age
* Sex ratio

With the help of these variables, mortality rate, survival rate, fertility rate and migration rates are calculated according to World Bank definitions of these variable (World Ban Developent Indicators):

Survival Rate = 1- Mortality rate

Mortality rate= age-sex specific deaths/ age-sex specific cohort population

Migration Rate= Observed growth – natural growth

Age specific Fertility rate = number of live births per cohort/age specific female population of childbearing age

To calculate these rates, the data for these variables is collected from Pakistan Demographic Survey published by Pakistan Bureau of Statistics for 2001, 2003, 2006 and 2007.

Information required for application of Hardy Weinberg Equilibrium, is adopted from previous research studies. Calculating them in itself is out of the scope of this research study therefore, their adoption from others’ findings is a more sutiable option.

# 3. Results

## 3.1. Population Projection

To project the number of Thalassemia patients, recent population estimates are required. However, the most recent census data available for Pakistan is of 1998. Since then the same data has been used to project the future demands for population planning and policymaking. However, the data is now a decade older and there are expected changes in the demographics of the population. Thus researchers have been using data from Pakistan Demographic Survey to be able to view a more realistic picture of the population. The most recent data available as a result of Pakistan Demographic Survey is from 2007. This data is more recent and takes into account the changes that may have occurred during the decade long period between 1998 and 2007. Moreover, the survey has been conducted repeatedly i.e. in 2001, 2003, 2006, and 2007. This provides enough information for the projection techniques to predict the future with higher certainty. However, even in the Demographic Survey the latest information available is up to 2007. To verify the projections from Cohort Component Method, the first five year projection is done till 2012. These projected values can be compared with the official estimates of the provincial population. According to estimates Khyber Pakhtunkhwa hosted 22-23 million of individuals in 2012 (KP Government, 2013). According to our model, a total of 22.96 million individuals are projected in 2012. This synchronizes with the official estimates, confirming that results for succeeding years can be relied upon.

|  |  |  |  |
| --- | --- | --- | --- |
| Table 3.1: Projected Population of Khyber Pakhtunkhwa | | | |
|  | 2012 | 2017 | 2022 |
| 0-4 | 3202317.605 | 3431408.458 | 3784758.297 |
| 5-9 | 3854242.317 | 4102402.945 | 4399327.201 |
| 10-14 | 3498570.912 | 3651206.131 | 3936778.383 |
| 15-19 | 2810401.509 | 3116844.434 | 3269112.835 |
| 20-24 | 1967824.029 | 2202276.092 | 2462742.664 |
| 25-29 | 1464077.514 | 1605954.482 | 1796110.531 |
| 30-34 | 976874.7045 | 1119514.464 | 1244232.185 |
| 35-39 | 957940.3149 | 968692.3546 | 1117047.362 |
| 40-44 | 849632.4955 | 870815.0928 | 900746.8577 |
| 45-49 | 837434.6506 | 878054.3751 | 914006.5822 |
| 50-54 | 645771.8329 | 700586.6831 | 741562.6337 |
| 55-59 | 513487.4192 | 578999.5429 | 630194.9239 |
| 60-64 | 471389.5786 | 458693.8549 | 513450.1338 |
| 65-69 | 370975.5967 | 390830.0818 | 384552.8914 |
| 70-74 | 239864.2368 | 280388.1007 | 296969.6705 |
| 75-79 | 115931.7353 | 121860.2816 | 144302.7439 |
| 80-84 | 88860.63447 | 114859.646 | 121029.3068 |
| 85+ | 97464.57973 | 159108.0928 | 258930.616 |
| TOTAL | **22963061.67** | **24752495.11** | **26915855.82** |

The other two columns in the table indicate the population projected in the next 10 years for Khyber Pakhtunkhwa. This projection is done in five-year intervals and presents the total population divided in the age cohorts of the population. Taking longer intervals for projection is neither feasible for Cohort Component method of projection, nor is it advisable for the purpose of future planning. One should have the projections for the intervals in which different plans cannot only be made, but implemented and monitored as well. Any length of interval longer than this will give rise to inaccuracies whereas a shorter interval is not feasible for policymaking.

Table 3.1 here shows the projections for the coming 10 years divided into age cohorts of the population. The table above shows detailed division of age-sex cohort. Adolescents and younger population is significantly higher than the older members of the region.

However, the age cohort of interest in this research is the population between 0-20 years of age. It is in this age group that prevalence of Thalassemic patients is observed. Life expectancy of Thalassemics in Pakistan is expected up to 30 years of age. On the other hand Khyber Pakhtunkhwa does not have standard health facilities required for management of the disease. It may be this reason that life expectancy in this region falls as compared to the rest of the country.

It has been reported that fertility and mortality rates both have been declining in the recent past in Pakistan. This means lesser births with longer life spans in the population. As a result lag has been created in the population delaying the population transition in Pakistan as compared to the rest of the countries around the globe. As a result Pakistan has a large proportion of the population going to school and work (Mahmood, 2011). This not only sheds light on the opportunities available in the country in case the Human Resource is properly utilized but also indicate lower number of births in the future, decreasing the Population burden the country sustains.

The province does not have any facility to conduct Prenatal Diagnosis of the pregnant women, who wish to diagnose the foetus for presence of Thalassemia Major. If a mother wishes to undergo the diagnosis, she is referred to Islamabad to avail the facility.

After projecting the population broadly, the number of Thalassemics that are added to the population is estimated with the help of Hardy Weinberg Analysis. The findings are reported in table 3.2.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3.2: Number of Thalassemia Patients and Carriers | | | | | | | | | | | |
| Calculating the number of Thalassemics | | | | | | | | | | | |
|  |  |  |  |  |  |  |  | |  | |  |
| Thalassemia  Major | Population | Carrier rate | p | q | F | p2+ Fpq | | Annual Births | | Thalassemic Patients | |
|  | 22963061 | 7 | 0.035 | 0.965 | 0.0165 | 1.225 | | 527634 | | 647 | |
| Thalassemia Carrier | Heterozygous | Carrier rate | p | q | F | 2pq(1-F) | | Annual Births | | Thalassemia Carriers | |
|  | 22963061 | 7 | 0.035 | 0.965 | 0.0165 | 66.435 | | 527634 | | 35054 | |

Table 3.2 reports the findings for the number of children born with Thalassemia in 2012. These children may be both Thalassemia Major (patients) and Thalassemia Minor (Carriers). The distinction of these two is important because it is the marriage of the carriers that result in the birth of Thalassemia Major (there is a 25% probability that a children born to Thalassemia Minor couples may be Thalassemia Major patients). The disease is genetic in nature and cannot be acquired from others, or avoided by adopting a certain life-style. Thus a genetic transmission model Hardy-Weinberg Equilibrium serves the purpose of identification of the number of people with Thalassemia trait in the population.   
Hardy-Weinberg Equilibrium deals with the genetic disease that is transmitted from parents to children with two gene alleles. Thalassemia is also such a disease in which both the parents as carriers, transmit the disease to their children only if they marry another carrier, with 25% probability.   
Table 3.2 shows results of the Analysis applied after projection of population of KP for 2012. Adopting the segregated population into age-sex cohorts is not necessary here because the disease may affect anyone whose ancestors may carry Thalassemia genes. As mentioned in the methodology information on carrier and prevalence rate of Thalassemia in KP is required for projection of Thalassemics. Carrier rate is reported as 7% in KP (Khattak et al. 2013). “p” represents the gene responsible for transmission of the disease. As mentioned above, the affected gene is transmitted in case both the parents are gene carriers, so p equals 3.5%. However, q on the other hand represents normality in transmission. Another matter of interest is the pattern of mating in the population. It may be that the carriers do not reproduce with other carriers, or it may be that mating is random. In Pakistan, especially Khyber Pakhtunkhwa, the trend is more inclined towards consanguineous marriages. This means that there is a higher possibility of marriage between off springs of the same ancestors, thus, carriers marrying each other. This is measured quantitatively with the help of inbreeding coefficient denoted by F. F-value is adopted from researches conducted previously in the province. It is reported to be 0.0165 (Wahab and Ahmad, 1996). With the help of Hardy Weinberg Analysis, the formulae for calculation of both Thalassemia Major patients and Thalassemia Minor carriers are calculated in the form of their prevalence (Ahmed, 1998).  
Once the prevalence is known, its application on the expected births (achieved with projection) can help us find the total number of Thalassemics born in the year. The table reports that there is an expected number of 647 children with Thalassemia Major. These children would require standard Thalassemia treatment to live a healthy life. Similarly, 35,054 individuals will be added to the population, each year, who may carry Thalassemia genes. If these individuals are not made aware of the challenges they may face during reproduction, they may have to face the misery of supporting a Thalassemia patient. The disease burdens a family not only financially but socially and psychologically as well.

**Conclusion:**

With the increasing population of Khyber Pakhtunkhwa the prevalence of Thalassemia is increasing. Projection of population shows that a total of 647 individuals with Thalassemia Major, and 35,054 with Thalassemia minor are born every year if the health and population demographics remain the same. Therefore, the government is required to take the required steps to ensure that awareness about the disease is created and parents are informed about prenatal tests which can help them in early detection of disease in their child. Further decision should be facilitated with the help of genetic counseling. However, most of these facilities are unavailable in the region and needs attention of the policy makers.

**References:**

1. Ahmed, S. (1998) An Approach for Prevention of Thalassemia in Pakistan, Thesis submitted to Department of Obstetrics and Gynaecology, University College London Medical School, University of London
2. Ain, Q. U., Ahmad, L., Hassan, M., Rana, S. M., Jabeen, F. (2011). Prevalence of β-thalassemic Patients Associated With Consanguinity and Anti-HCV - Antibody Positivity – A Cross Sectional Study, *Pakistan Journal of Zoology, 43(1)*, pp. 29-36, 2011.
3. Jiang, B. (2007). Better Long-Term Population Projection Method: A Case Study for Canberra Suburbs, Thesis submitted to The Department of Computer Science Australian National University
4. Khattak A. Z., Taj, A. S., Shah, S. M. A., Khattak, M. F. (2013). Clinical And Haematological Presentations Of Haemoglobin E Disorders From Northern Pakistan, Journal of Medical Sciences, Vol. 21, No. 3, 134-136
5. Khyber Pakhunkhwa Government (2013) Home Department Estimates.
6. Lorey, F. (2000). Asian immigration and public health in California: thalassemia in new- borns in California. J. Pediatr. Hematol. Oncol. 22: 564–566.
7. Mahmood, N. (2011). Demographic Dividend: Effects of Population Change on School Education in Pakistan, Working Paper, PIDE: 68
8. Marshall, A. (2009). Developing a methodology for the local estimation and projection of limiting long term illness and disability, Doctoral Desertation submitted to School of Social Sciences, University of Manchester
9. Rahman, M., Lodhi, Y. (2004). Prospects and Future of Conservative Management of Beta Thalassemia Major in a Developing Country, Pakistan Journal of Medical Sciences, Vol. 20, No. 2
10. Siegel, J. (2002). Applied Demography. London, Academic Press.
11. Thalassemia International Federation (2012). Activity Report 2011, 2012 Activity Plan
12. Weatherall, D.J. & Clegg, J.B. (2001). Inherited haemoglobin disorders: an increasing global health problem. Bull. World Health Organ. 79: 704–712