5.1: A PROBABILISTIC MODEL FOR ESTIMATING DIFFERENTIAL MORTALITY IN SICKLE CELL ANAEMIA

5.1 Introduction: It is well known that sickle cell anaemia is a genetic disorder which is transferred from one generation to the next. Here we have a locus with two alleles A and S, sickle cell allele being autosomal recessive, the homozygotes AA are normal, SS suffer from sickle cell disease and the heterozygotes are only carriers and they do not exhibit any gross disadvantages when compared with the normal homozygous AA. Indeed in falsiparum malaria prone area they are supposed to have some advantage over the normal homozygous AA and for this reason the allele S does not get eliminated from the population.

Till recently it was commonly believed that the homozygous SS suffering from sickle cell anaemia are physically weak and have grossly reduced longevity. However, some recent development shows that this belief needs modification. There are some forms of sickle cell anaemia distributed in certain parts of the world which are somewhat milder because of the concurrent presence of alpha thalassemia of raised level of foetal hemoglobin. (59,60)

In order to make comparisons of different extents of severity of sickle cell anaemia in different parts, it is necessary to develop a suitable model, which suits the theoretical understanding of this disorder, which fits real life data and which has parameters providing measures of extent of severity.

The excess selective mortality due to any
Disease is termed as differential mortality due to that particular disease. Testing for the presence of differential mortality and if present, estimating its magnitude is of considerable importance.

A model is designed to test the presence of and to estimate the magnitude of differential mortality due to any genetically inherited (single gene factor) disease. In particular, hemoglobin disorders are considered and the method is illustrated with data on Sickle cell gene that has been collected during this work.

During this study, number of surveys were arranged to assess the distribution of sickle cell hemoglobin in different caste and tribal groups from different geographical areas of Maharashatra. In this 18 different caste and tribal groups were studied. They were spread from Gadchiroli district from east through Vidarbha, Western Maharashtra to Ratnagiri district from west.

6.2: METHODOLOGY

Let us denote a normal hemoglobin gene as A and abnormal hemoglobin gene as S. Then if the gene proportions are \( p(A) = p \) and \( p(S) = q \), the genotype array according to Hardy-Weinberg (H-W) Equilibrium Law is \( \{p AA, 2pqAS, q SS\} \).

With the knowledge of gene frequencies this formula may be used to predict the incidence of the different genotypes at birth, although their prevalence at later ages will be influenced by the relative fitness of the different genotypes.
TESTING FOR PRESENCE OF DIFFERENTIAL MORTALITY-

If we denote observed frequencies of genotype AA, AS, SS as $N_1, N_2, N_3$ respectively and let $N_1 + N_2 + N_3 = N$, then

$$N_1 = pN, \quad N_2 = 2pqN, \quad N_3 = qN.$$ 

$E(N) = \frac{2}{p} = \frac{2}{p^*} = \frac{2}{q^*} = 2pq,$ $E(N) = \frac{2}{q} = \frac{2}{2q} = 2q,$ $E(N) = \frac{2}{1-q} = \frac{2}{2(1-q)} = 2-2p$ 

$E(N) = \frac{2}{2(1-q)} = \frac{2}{2(1-q)} = 2-2p$

$$p^* = \frac{2}{2(1-q)} = \frac{2}{2(1-q)} = 2-2p$$

Expected frequencies under H-W law are,

$$E(N_1) = p^* \cdot N, \quad E(N_2) = 2p^*q^*N, \quad E(N_3) = q^* \cdot N$$

The goodness of fit of H-W law can be tested applying chi-square goodness of fit test with one degree of freedom. If there is differential mortality in abnormal homozygotes then difference between observed $N_3$ and expected $N_3$ will be large and will yield a significant chi-square value. This test will provide us a test for the presence of differential mortality.
Estimation of Differential Mortality Rate

If the chi-square value is significantly large indicating the presence of differential mortality, then estimating its magnitude will be meaningful.

Let \( a \) be the survival factor and so \((1-a)\) be the differential mortality factor in abnormal homozygous.

Let \( N \) be the total sample size if there is no differential mortality i.e. if \( a=1 \).

Then,

\[
\begin{align*}
N &= p N \quad \text{(1)} \\
N &= 2pqN \quad \text{(2)} \\
N &= q N \quad \text{(3)}
\end{align*}
\]

\[
p^* = \frac{2N}{2N + N} \quad , \quad q^* = 1 - q^*
\]

\[
N = \frac{2N}{p} \quad \text{(eq-1)}
\]

\[
N = \frac{2N}{qN} \quad \text{(eq-3)}
\]

These are the moment estimators. Maximum likelihood Estimator (MEL) of \( p \) is same as moment estimator.
MLE of \( a \) takes different form. It can be shown,

\[
a = \frac{4N n}{31}
\]

Although moment estimator and MLE of \( a \) are not mathematically equivalent, they yield the same value of upto seven digits in few situations and upto four digits in many situations.

A probabilistic model for such problem was suggested recently by Kharshikar et al (1990). Application of their model becomes cumbersome as it involves interactive procedure and so a computer facility. On the other hand this method is very simple for understanding and application.

Result - The above proposed model for estimating the differential mortality is illustrated using the data collected on sickle cell genotype for various caste and tribal groups during surveys at various places are shown in table-1. All these data sets were analysed separately using the above method. In case of Pardhan tribal group we expect 6 to 7 cases of abnormal homozygotes (SS type) but in actual survey only 2 cases were found. Chi-squared value of this data set is 4.19. The corresponding probability of type I error is around (0.03) which is significantly higher than a conventional level (0.05) of significance.

For this data set, \( a \) (survival factor) is 0.31 and \( 1-a \) (mortality factor) is 0.69 or 69% which is indeed very high.

Above data suggest that from 18 different communities studied only Pardhan tribal community have high mortality

69
due to sickle cell disease, while in all other caste and tribal communities there is no statistically significant mortality. This also suggest that the sickle cell disease is more severe in Pardhan tribal community compared with all other caste or communities.
Table -1: Table showing observed and expected frequencies by phenotype and caste groups.

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Name of the caste/Tribe and district</th>
<th>Observed and Expected Frequencies</th>
<th>Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Phenotype</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AA</td>
<td>AS</td>
</tr>
<tr>
<td>1.</td>
<td>Andha Nanded</td>
<td>0</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>148.9</td>
</tr>
<tr>
<td>2.</td>
<td>Banjara Jalgaon</td>
<td>0</td>
<td>972</td>
</tr>
<tr>
<td></td>
<td>Osmanabad E</td>
<td></td>
<td>971.2</td>
</tr>
<tr>
<td></td>
<td>Dhule</td>
<td></td>
<td>0.93</td>
</tr>
<tr>
<td>3.</td>
<td>Bhill Dhule</td>
<td>0</td>
<td>1259</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>1258.14</td>
</tr>
<tr>
<td>4.</td>
<td>Gond Nanded</td>
<td>0</td>
<td>350</td>
</tr>
<tr>
<td></td>
<td>Yavatmal E</td>
<td></td>
<td>349.38</td>
</tr>
<tr>
<td></td>
<td>Gadchiroli</td>
<td></td>
<td>0.93</td>
</tr>
<tr>
<td>5.</td>
<td>Katkari Ratnagiri</td>
<td>0</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>97.73</td>
</tr>
<tr>
<td>6.</td>
<td>Kolam Yavatmal</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>32.94</td>
</tr>
<tr>
<td>7.</td>
<td>Kokana Dhule</td>
<td>0</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>102.39</td>
</tr>
<tr>
<td>8.</td>
<td>Korku Amaravati</td>
<td>0</td>
<td>644</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>646.05</td>
</tr>
<tr>
<td>9.</td>
<td>Mahadeo- Koli Raigad</td>
<td>0</td>
<td>366</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>365.98</td>
</tr>
<tr>
<td>10.</td>
<td>Madia Gadchiroli</td>
<td>0</td>
<td>196</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>195.06</td>
</tr>
<tr>
<td>11.</td>
<td>Malhar-koli Thane</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>31.83</td>
</tr>
<tr>
<td>12.</td>
<td>Pawara Dhule</td>
<td>0</td>
<td>778</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>778.38</td>
</tr>
<tr>
<td>13.</td>
<td>Pardhan Yavatmal</td>
<td>0</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>92.68</td>
</tr>
<tr>
<td>14.</td>
<td>Rajgond Gadchiroli</td>
<td>0</td>
<td>256</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>256.08</td>
</tr>
<tr>
<td>15.</td>
<td>Scheduled Osmanabad caste</td>
<td>0</td>
<td>926</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td></td>
<td>925.98</td>
</tr>
</tbody>
</table>

71
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16. Tadwi</td>
<td>Jalgaon</td>
<td>06</td>
<td>66</td>
<td>00</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>65.87</td>
<td>5.99</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>17. Thakur</td>
<td>Pune</td>
<td>00</td>
<td>358</td>
<td>00</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>Thane</td>
<td>00</td>
<td>358</td>
<td>00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A'Nagar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Warli</td>
<td>Thane</td>
<td>00</td>
<td>343</td>
<td>00</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>342.40</td>
<td>29.95</td>
<td>0.65</td>
<td></td>
</tr>
</tbody>
</table>
Discussion-

A probabilistic model for such problem was suggested recently by Kharshikar et al (1990). Application of their model becomes cumbersome as it involves interactive procedure and so a computer facility. On the other hand this method is very simple for understanding and application. The model suggested by Kharshikar et al was also applied to determine the differential mortality due to sickle cell hemoglobin and it also showed that out of the tribal and caste groups studied there is differential mortality in Pardhan tribal community only and showed the mortality factor to be 72%. Our findings (mortality factor 69%) are quite consistent with the reported one.

The effect of non-random mating lead to a preferential association of heterozygotes and disturb the equilibrium. This is shown by Al-Awamy et al in the Eastern Province of Saudi Arabia where cousin marriages are frequent and the numbers of homozygotes are greater than would be expected from the observed gene frequency.

G.R. Serjeant have shown variable spectrum of clinical involvement in SS disease. If the different patients with an apparently identical molecular abnormality have clinical courses so variable that some die in early childhood, while others have a virtually unrecognized condition at the age of 50 years, it is clear that inheritance of SS disease alone is not the sole determinant of clinical severity.

Hence, to corelate the higher mortality in Pardhan
tribal community as compared to other tribal and non-tribal groups, it is essential to study in detail genetic as well as non-genetic factors causing variability. The genetic factors alapha thalassemia, heterocellular persistence of foetal hemoglobin, and G-6-PD deficiency and environmental factors such as socio-economic status, frequency of serious infections such as malaria, folate deficiency and deficiencies of iron, zinc, and other trace elements.

Interaction of alapha thalassemia 2 with SS (64,65) disease influences haematological indices, reduces (66) haemolytic rate and increases deformability of red (67) cells.