APPLICATION OF RENEWAL EQUATION AND MARKOV CHAIN TO THE INHERITANCE PATTERN OF SICKLE CELL ANAEMIA (SCA)

BY

AKANBI, Olumuyiwa.O

Department of Mathematics & Statistics,
Federal Polytechnic, Ilaro,
Ogun State, Nigeria.
Email: Olumuyiwaaaakanbi@yahoo.com

ABSTRACT

*Sickle cell anaemia happens to be one of the sickle cell diseases characterized by the predominance of haemoglobin **S**. This paper investigates the dynamics of the disease by Renewal equation and the probability of acquiring any of the alleles by Markov Chain. The Renewal equation reflects the aboundance of **AS** group(carrier) at the long run. Meanwhile, the absorbing state of the system is mainly for the mating between three pairs namely **AA&AA, AA&SS and SS&SS.** In the same vein the Markov process also reflects that **AS** group has greater potential to remain aboundant in the system. The study concludes that, with or without genotype screening before marriage, the population of **AS** group will be in aboundance, which implies that **AS&AS** relative activities cannot be permanently zero. Hence curative measure is the only permanent solution to the emergence of sickle cell anaemia.*

Keywords: *Markov Chain, Renewal equation, sickle cell anaemia and genotype.*

Introduction

Individual genotype **AA, AS, SS, SC** or **AC** differs amongst the world's population (Seeley, 1998) . Of interest however, is the fact that genetic mechanism on morphogenetic traits is still not clearly understood as it is seen to occur with variable frequency in different populations and thus useful in evaluating and analyzing evolutionary forces and classification (Das, 2003). Meanwhile, marked inter-individual variability in genetic and

non-genetic factors has been said to posses the ability to influence the disposition of many endobiotics and xenobiotic affecting health (Lamba et al, 2002)

Sickle Cell Anaemia which is a genetically transmitted disease is caused by a defective allele (mutant form) of the gene coding for a sub unit of the haemoglobin protein. The Sickle haemoglobin tends to precipitate or "clump together" within the red blood cells after releasing its oxygen. If the clumping is extensive the red blood cell assumes an abnormal sickle shape. These sickle red blood cells plug the blood vessels thus preventing normal red blood cell passage and consequently depriving the tissue of needed oxygen.

Each person has two copies of the gene that determines whether that person has Sickle Cell Anaemia. If both copies are "normal alleles" then only normal haemoglobin is produced **AA**. If one of the two alleles is defective then that person has a mixture of normal and Sickle haemoglobin: a condition known as Sickle Cell trait "**AS**" (Carrier). If both alleles are defective, essentially only sickle haemoglobin is made and the person has Sickle Cell Anaemia "**SS**".

The first case of **SCA** was reported in 1910 on a Jamaican student in the USA (Herrick,1910) The term **SCA** is a term first used by Mason in 1922 to describe the homozygous state (Sergeant, 1985). In 1949, Neel illustrated that **SCA** was transmitted as a recessive gene '**S**'. But it is well known by scientist now that the gene is neither dominant nor recessive but of intermediate penetrance.

The allele causing Sickle Cell Amaemia is found most often in people of African ancestry. It was traced to one family in Ghana (krobo people) in 1670. The "**S**" gene is found mainly where malaria is endemic (Kathleen et al, 2011). It also occurs in people of Mediterranean, Arab, East India, South and Central American ancestry. From the world population of about 7,058, 157,073 (US census bureau, 2013) (No of people officially counted). 5% of world population lives with SCD. Mathematically, about 352,907,854 people have haemoglobin disorder "officially counted" while every year 300,000 infants are born with SCD, including 200,000 cases in Africa(World health assembly,2006). Nigeria of about

150 million population with growth rate of 3.2% has prevailing rate of 150,000 offspring per year. Nigeria by the virtue of her population ranks first as a **SCA** endemic country in Africa with annual infant deaths totally around 100,000, 8% of infant mortality. This alarming infant death related to **SCA** made the authors to carryout mathematical transmission dynamics and its simulations.

Literature Review

Mathematical models of heredity are largely based on one-locus, two allele gene populations, where little or no attempt is made to consider the dynamics of the population and the analysis is somehow probabilistic, accounting of genotype composition of a diploid population. Nevertheless, success approaches have been developed for continuous-time model that laid more emphasis on the dynamics of population of SCA (Tchuenche,2002). K. H. Rosen 1983 presented a mathematical model for polygamous system, where he introduced the interaction function (KMF) where M, F, were the male & female population sizes. A new set of functional differential equations modeling heterosexual population dynamics where K was called the polygamy factor was developed by C.O.A Sowunmi in 1993. Here he proved that population renewal is the product of an interaction between the reproductive males and non-gestating reproductive females. C.O.A. Sowunmi and J. M. Tchuenche were the first to use the general interaction function Fij. where Fij contributes the major mathematical background for the formulation of our birth renewal equation. J.M Tchuenche has done a lot in the dynamics of SCA. He is the pioneer mathematician who made use of differential equation to model the disease (SCA). J.M Tchuenche (2000) demonstrated the dynamical behaviour of SCA by a set of first-order non-linear partial differential equation. He also developed the birth renewal equation the same year. General modeling approach which considered the selective advantage of Haemoglobin **(S)** over haemoglobin (**HbA**) was introduced by J. M Tcheunche. The age and character-dependent population dynamics model of genetically transmitted diseases where standard techniques of functional analysis were employed to further simplify and solve the basic equations implicitly in special cases of **SCA** (J. M. Tchuenche 2005) was introduced

Chelsea Liddell et al (2011) made it critically clear with his mathematical model of sickle cell genome frequency in response to selective pressure from malaria that selection pressure for the carrier gene in the presence of increasing malaria death for either adult or children has higher frequencies of the gene as well as shortened time to reach these frequencies.

Recently, differential impact of sickle cell trait on symptomatic and asymptomatic malaria was introduced where a decreased frequency of S-gene may eventually increase the overall prevalence of both symptomatic and asymptomatic malaria. Hence, the control of symptomatic malaria might result in evolutionary repercussion, despite short-term epidemic logical benefits Eunha Shim et al (2013).

A Markov chain is one of the basic methods in studies that involve random processes, it has also gained attention in the search for the dynamics of inheritance diseases. The method was named in honour of a Russian mathematician A.A. Markov in 1906. He derived this process while studying probabilities in playing-card games. This idea was propagated by Fisher et al. Fisher's work laid the foundation of the modern mathematical theory of genetics in plants and animals. The Markovian assumption is the probability of moving from one state to another and is independent of the history of before arriving in that state.

Since a model is a simplified representation of a complex system, designed to focus on a specific question. Our model is focused on answering questions related to the dynamics of sickle cell anaemia. In which we hope that our projection on the dynamics of SCA will help the government to identify this disease as a national priority and allocate resources in targeting intervention and treatment plans.

Markov Chain has been applied to numerous studies in the medical field, such as determination of degree of efficiency of noninsulin-dependent diabetes in a population of patients (Kuo et al ,1999). Markov chain has also been applied to analyse longitudinal disease progression for liver cancer (Kay, 1986). (Debanne et al,2000) introduced a multivariate Markov chain model to forecast tuberculosis trend in the U.S.A from 1980 to 2010 among different races in the country. (Cherry, et al ,2012) with the help of Markov chain studied the clinical effectiveness and cost-effectiveness of primary stroke prevention in children with sickle cell anaemia disease. So many researchers had done some previous studies on the modelling of disease progression and transmission dynamics using Markov model see (Sweeting et al 2010, Commenges et al 2004, Jackson et al 2002) and references therein. However, with the application of Markov model the study considers the question of determining the probability that given the chain is in state *i* in the 1st generation, it will be in state *j* in the generation to come.

Modeling background for renewal equation.

A population is often regarded as a group of individuals that can cross (mate) and give birth to offspring (neonates). Human population differs with respect to their genotype, age, Rhesus factor, physiology etc. Hence they can be classified according to their distinguishable differences. The following assumptions are taken into consideration:

i The entire population has an age structure that spans through (O, L] where L is the life span.

ii A genotype structure composed of three distinct genetic subgroups is duly imposed on the population i.e. normal **AA**, carriers **AS**, and sickles **SS**.

iii Monogamy form of family setting is also assumed.

iv The spread of the population is spatially homogeneous.

v Variables x, y ε $\mathbb{R}_+$ are the independent variables age and time respectively.

Consider below:

Genotype	Sex		Phenotype
	Females	Males	
AA	f_1	m_1	Normal
AS	f_2	m_2	Carriers
SS	f_3	m_3	Sicklers

Suffixes $\boldsymbol{i}$ = 1, j = 2, k = 3 corresponding to:

AA: = 1

AS: = 2

SS: = 3 respectively

$f_{\boldsymbol{i}}$ (x,y) ≥ 0 represents the population density of females in group $\boldsymbol{i}$ with age x at time y.

$m_{\boldsymbol{i}}(x, y) \geq 0$ represents the population density of males in group $\boldsymbol{i}$ with age x at time y.

$F_{\boldsymbol{i}j}$ {[$m_{\boldsymbol{i}}$ (x,y); f_j (x^1,y)], x,x^1y} – function governing the interaction between males of class '$\boldsymbol{i}$' males and females of class 'j'.

δ_{ij}^{k}- Is the probability of getting a neomate of class 'k' from mating between class '$\boldsymbol{i}$' males and class 'j' females.

Γ – The probability of acquiring the S gene from any of the parents.

f_i^j – Number of females of class'$\boldsymbol{i}$' interacting with class 'j'males (Referred to as number of couples) in a monogamous setting.

(a) $F_{\boldsymbol{i}j}$ [($f_{\boldsymbol{i}}$; m_1, m_2, m_3,) α, α^1, y]. In this case, every females at reproductive age has just a life partner and vice-versa.

(b) $F_{\boldsymbol{i}j}$ [($m_{\boldsymbol{i}}$; f_1, f_2, f_3,) α^1, α, y]. This means some males of class '$\boldsymbol{i}$'interact with normal females f_1, others with females carriers f_2, and the remaining with sicklers f_3. In this case polygamy may arise. In the sequel we shall refer only to the former.

We cannot affirm with certainty that a woman aged x will marry a man aged x^1. This is a problem that takes into cognizance numerous factors that cannot be controlled mathematically. We have therefore modified the interaction function $F_{\boldsymbol{i}j}$ to suit our case where the population is subdivided into three groups; namely, **AA, AS** and **SS.** Our two-sex mixing function $F_{\boldsymbol{i}j}$ can be defined as:

$$F_{\boldsymbol{i}j} = F_{\boldsymbol{i}j} \{[f_{\boldsymbol{i}} (\alpha, y); m_j ((\alpha^1, y)]; \alpha, \alpha^1, y\}.$$

Where $F_{\boldsymbol{i}j}$ is a positive real-valued function and has compact support with respect to (α, α^1). The reproductive age of females spans the interval $[w_{f_r}, w_{f_s}]$ and $[w_{m_r}, w_{m_s}]$ for males. Hence;

Supp $F_{\boldsymbol{i}j} \leq [w_{f_r}, w_{f_s}]$ x $[w_{m_r}, w_{m_s}]$

<u>Derivation of birth equation:</u>

The probabilities of acquiring the gene from either parents are the same, this point is worth noting, because this is what implies $F_{ij} = F_{ji}$,

For x$\in$ (O, L] $0 \leq \delta_{ij}^{k} \leq 1$ and the compactness of the support of F_{ij}; the birth equations are given below:

$$\boldsymbol{B}_{1(y)} = \int_0^{\infty} \int_0^{\infty} (\delta_{11}^{(1)} F_{11} + \delta_{12}^{(1)} F_{12} + \delta_{22}^{(1)} F_{22})\, d\propto d\propto^{\iota} \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(i)}$$

$$\boldsymbol{B}_{2(y)} = \int_0^{\infty} \int_0^{\infty} (\delta_{12}^{(2)} F_{12} + \delta_{13}^{(2)} F_{13} + \delta_{22}^{(2)} F_{22} + \delta_{23}^{(2)} F_{23})\, d\propto d\propto^{\iota} \ldots\ldots\ldots\ldots\ldots\text{(ii)}$$

$$\boldsymbol{B}_{3(y)} = \int_0^{\infty} \int_0^{\infty} (\delta_{22}^{(3)} F_{22} + \delta_{23}^{(3)} F_{23} + \delta_{33}^{(3)} F_{33})\, d\propto d\propto^{\iota} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(iii)}$$

And

$$\beta = \sum_{n=1}^{3} \mathcal{B}_n(y)$$

Properties of $B_n(y)$

1. $B_n(y)$ is well defined and non negative $F_{ij} > 0$. Human beings are a good example of birth flow population in which birth occur continuously over the time interval.
2. $B_{n(y)}$ is a first degree homogeneous function of its arguments i.e $B_n K F_{ij(y)} = K B_n F_{ij(y)}$.
3. $B_{n(y)} = 0$ for any population in which either male or female is absent.

Now if genotype screening is duly imposed on the model (i.e the relative activities between AS x AS, AS x SS, and SS x SS become zero). Hence, we have the birth equation as given below:

$$\hat{B}_{1(y)} = \int_0^{\propto} \int_0^{\propto} (\delta_{11}^{(1)} F_{11} + \delta_{12}^{(1)} F_{12})\, d\propto d\propto^{1} < B_{1(y)} \ldots\ldots\ldots\ldots\ldots\text{(iv)}$$

$$\hat{B}_{2(y)} = \int_0^{\propto} \int_0^{\propto} (\delta_{12}^{(2)} F_{12} + \delta_{13}^{(2)} F_{13})\, d\propto d\propto^{1} < B_{2(y)} \ldots\ldots\ldots\ldots\ldots\text{(v)}$$

$$\hat{B}_3 = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(iv)}$$

From (iv), (v) and (vi) there will be no birth of sicklers, there is drastically decline in the birth equation and there is abundant of **AS** in the population.

Markov Chain Representation

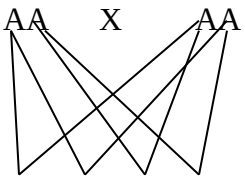
We describe the Markov chain as: Let f = {$f_1, f_2, \ldots\ldots, f_n$} be a set of states. The process starts in one of these states and moves successively from one state to another. If the chain is currently in

state f_i, then it moves to state f_j at the next step, with a probability denoted by F_{ij}. F_{ij} are referred to as transition probabilities. The process can remain in the same state, this occurs with probability of F_{ii}.

In this study, since each of the offspring has two parents and hence the genotype of an individual depends on those of both parents. Therefore to build a Markov model for the three genotypic group, we consider the evolution of the genotypic groups.

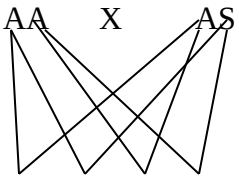
Possible Outcome

Case I



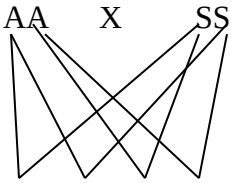
100% of AA

Case II



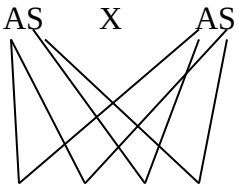
50% of AA & 50% of AS

Case III



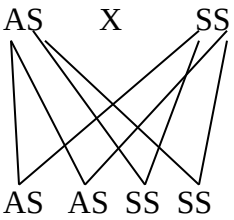
100% of AS

Case IV



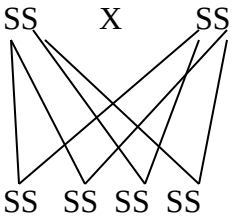
50% of AS, 25% of AA, 25% of SS

Case V



50% of AS & 50% of SS

Case VI



100% of SS

Representing the different cases of mating above as an array we have the following probabilities

Parents			AA	AS	SS
AA	X	AA	1	0	0
AA	X	AS	$^1/_2$	$^1/_2$	0
AA	X	SS	0	1	0
AS	X	AS	$^1/_4$	$^1/_2$	$^1/_4$
AS	X	SS	0	$^1/_2$	$^1/_2$
SS	X	SS	0	0	1

We need to note that the inclusion of SS marrying SS is very rear. We included it for completion sake. Hence to make our model a realistic model SS mating SS will not appear in our matrix representation.

Therefore we have;

Parents			AA	AS	SS
AA	X	AA	1	0	0
AA	X	AS	$^1/_2$	$^1/_2$	0
AA	X	SS	0	1	0
AS	X	AS	$^1/_4$	$^1/_2$	$^1/_4$
AS	X	SS	0	$^1/_2$	$^1/_2$

Formulation of Transition Matrix

Fij = Fji. No importance is attached to the interaction which one of the pairs is a male or female.

Assuming the states are 1,2,…….,n then the state transition matrix is given by;

$$F = \begin{pmatrix} F_{11} & F_{12} \ldots\ldots F_{1n} \\ F_{21} & F_{22} \ldots\ldots F_{2n} \\ . & . & . \\ . & . & . \\ F_{n1} & F_{n2} & F_{nn} \end{pmatrix}$$

Fij ≥ 0 $\forall i$ we have

$$\sum_{k=1}^{n} F_{ik} = \sum_{k=1}^{n} F(X_{m+1} = k \quad X_m = i) = 1$$

Thus, the model considers two cases

Case I:

If an offspring in chosen at random and is mated with AS and this process is repeated through a number of generations. The states obtained are AA, AS and SS. Hence the transition probabilities are

F_{ij} =		AA	AS	SS
	AA	0.5	0.5	0

$$\begin{array}{c} \\ AS \\ SS \end{array} \begin{pmatrix} 0.25 & 0.5 & 0.25 \\ 0 & 0.5 & 0.5 \end{pmatrix}$$

Case II:

If an offspring chosen at random and is mated with AA. Hence the transition probabilities are

$$F_{ij} = \begin{array}{c} \\ AA \\ AS \\ SS \end{array} \begin{array}{c} \begin{array}{ccc} AA & AS & SS \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 1 & 0 \end{pmatrix} \end{array}$$

Analysis & Results:

Case I

To obtain the stationary matrix

SF = S , Where $S = S_1, S_2, S_3$

$$[S_1\ S_2\ S_3] \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 0.5 & 0 \\ 0 & 1 & 0 \end{pmatrix} = [S_1\ S_2\ S_3]$$

With $S_1 + S_2 + S_3 = 1$

By TORA mathematical package;

$[S_1\ S_2\ S_3] = [0.25\ \ 0.50\ \ 0.25]$

This implies that 25% of the population will be AA, 50% will be in AS and 25% will belong to SS genotypic group.

Long run behavior

For case I:

The limit of transition matrix at the 20^{th} generation is obtained as;

$$F^{20} = \begin{array}{c} \\ AA \\ AS \\ SS \end{array} \begin{array}{c} \begin{array}{ccc} AA & AS & SS \end{array} \\ \begin{pmatrix} 0.25000048 & 0.50000000 & 0.24999952 \\ 0.25000000 & 0.50000000 & 0.25000000 \\ 0.24999952 & 0.50000000 & 0.25000048 \end{pmatrix} \end{array}$$

In the next 20 generations the population will stabilize at 25%, 50% and 25% for AA, AS & SS genotypic group respectively.

For Case II

The limit of transition matrix of case II is obtained as;

$$F^{20} = \begin{array}{c} \\ AA \\ AS \\ SS \end{array} \begin{array}{c} \begin{array}{ccc} AA & AS & SS \end{array} \\ \begin{pmatrix} 1.00000000 & 0.00000000 & 0.00000000 \\ 0.99999905 & 0.00000095 & 0.00000000 \\ 0.99999809 & 0.00000191 & 0.00000000 \end{pmatrix} \end{array}$$

That is, if other genotypic group is focused on marrying AA only. This implies that in the next 20 generations the whole population will have AA genotype. But how realistic is this?

Conclusion

The renewal equation being a continuous process reflects that the birth of sicklers is an absolutely continuous process. Since the relative activities between AS & AS and AS & SS are not zero. Even if the relative activities of AS & AS and AS & SS are zero, (equation iv & v) there will be decrease in the population with aboundant AS class in the population. Hence there is a great possibility for the population of sicklers to increase over long period of time. The renewal equation shows that AS genotypic group will be in abundant in the long run. From the two cases considered under a Markov process, for the 1st case we realized that 50% of the population will be AS in the long run. For the 2nd case which is not likely to occur in real life the whole population will assume AA genotypic group over a long period of time. However, the remaining two genotypic group will be without life partner which is not realistic any way.

Considering facts from the model, it is trivial that AS genotypic group will be abundant in the population and might eventually engage in mating since an average African will not observe celibacy as a way of life. Engagement of AS &AS brings about the emergence of SS class. Though, bone marrow transplants as a curative measure has seen the limelight. But it is too expensive in the developing countries.

Hence the study recommends that the vacuum in the curative measure still stands a space to be filled. Therefore, establishment of a body to scout and nature ideas that will lead to curative drug is supreme.

References

Akanbi O.O 2006: Discrete and continuous-time models of the genetic structure of Sickle-Cell Anaemia. M.Sc Thesis University of Ibadan.

Das B, Sengupta. S (2003): A note on some Morphogenetics variables among the Sonowal Kacharis of Assam. *Antropologist 5(3): 211-212.*

Fifty –ninth World Health Assembly Provisional Agenda item 11.4, 24th April 2006.

Herrick J.B. (1910): Peculiar Elongated and sickle shaped red corpuscles in a case of severe anaemia. *Arch. Int med 6: 516-518.*

Kathleen A. Neville, Julie A panepinto 21-25 march 2011: pharmacotherapy of SCD: *18th expert committee on the selection and use of essential medicines.*

Lamba J. K, (2002): Common allelic variants of cytochrome P4503A4 and their prevalence in different populations. *Phamacogenetics 12:121-132*

Seeley RR, Stephens T.D, Tate P (1998): Anatomy and physiology 4th edition. The McGraw Hill Companies, Inc. USA Pp 1098.

Sergeant GR 1985: Sickle Cell Disease Oxford University Press, London pp 1-25.

Sowunmi C.O.A 1993: "A model of Heterosexual population Dynamics with Age structure and gestation period, ''Academy Press, Inc.

Tchuenche J.M. 2002: Mathematical population Dynamics of Sickle- Cell Anaemia, Ph.D thesis University of Ibadan.

U.S. Census Bureau Jan 2013.