



ISSN 2278 – 0211 (Online)

## The Study of Population Dynamics of Sickle Cell Anaemia by Renewal Equation and Markov Chain Process

Akanbi, Olumuyiwa O.

Lecturer, Department of Mathematics & Statistics, Federal Polytechnic Ilaro, Ogun-State, Nigeria

### Abstract:

*Sickle cell Anaemia (SCA) is the most common form of sickle cell Diseases. Sickle cell anaemia is a potentially devastating condition that is caused by a defective allele (mutant form) of the gene coding for a sub unit of the haemoglobin protein. This paper investigates the dynamics of the disease by renewal equation and the probability of acquiring the S gene by Markov chain process. The renewal equation reflects the abundance of AS group (carrier) in the population at the long run. In the same vein the markov process also reflects that AS genotypic group has greater potential to dominate the population at the long run. (i.e. at 20th generation). Three absorbing states were noticed: the mating between three pairs AA&AA, AA&SS and SS&SS. Therefore, the study concludes that, with or without genotype screening before marriage, the population of AS genotypic group is expected to grow abundantly in the population. The abundance of the population of AS indicates that the relative activities of this heterogeneous group cannot be permanently zero. Hence, the study recommends that, a cheaper curative measure remains the only solution to the problem of sickle cell anaemia in our society.*

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

### 1. 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 possess 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 are 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 Anaemia 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 Sickle Cell Disease (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 author to carryout mathematical transmission dynamics and its probability of acquiring the alleles.

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). A new set of functional differential equations modeling heterosexual population dynamics where  $K$  was called the polygamy factor was developed by (Sowunmi,1993). Sowunmi proved that population renewal is the product of an interaction between the reproductive males and non-gestating reproductive females. Sowunmi and Tchuenche were the first to use the general interaction function  $F_{ij}$ . Where  $F_{ij}$  contributes the major mathematical background for the formulation of our birth renewal equation. 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 (Tchuenche,2003). He 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. 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 (Tchuenche,2005).

Chelsea 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 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. 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 have 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, then it will be in state  $j$  in the generation to come.

### 1.1. 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  $(0, L]$  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  $x, y \in \mathbb{R}_+$  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

Table 1

Suffixes  $i = 1, j = 2, k = 3$  corresponding to:

**AA**: = 1

**AS**: = 2

**SS**: = 3 respectively

$f_i(x, y) \geq 0$  represents the population density of females in group  $i$  with age  $x$  at time  $y$ .

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

$F_{ij} \{ [m_i(x, y); f_j(x', y)], x, x' y \}$  – function governing the interaction between males of class ‘ $i$ ’ males and females of class ‘ $j$ ’.

$\delta_{ij}^k$  – Is the probability of getting a neomate of class ‘ $k$ ’ from mating between class ‘ $i$ ’ males and class ‘ $j$ ’ females.

$\Gamma$  – The probability of acquiring the S gene from any of the parents.

$f_i^j$  – Number of females of class ‘ $i$ ’ interacting with class ‘ $j$ ’ males (Referred to as number of couples) in a monogamous setting.

(a)  $F_{ij} [(f_i; m_1, m_2, m_3), \alpha, \alpha', y]$ . In this case, every female at reproductive age has just a life partner and vice-versa.

(b)  $F_{ij} [(m_i; f_1, f_2, f_3), \alpha', \alpha, y]$ . This means some males of class ‘ $i$ ’ interact with normal females  $f_1$ , others with female 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  $x'$ . This is a problem that takes into cognizance numerous factors that cannot be controlled mathematically. We have therefore modified the interaction function  $F_{ij}$  to suit our case where the population is subdivided into three groups; namely, **AA**, **AS** and **SS**. Our two-sex mixing function  $F_{ij}$  can be defined as:

$$F_{ij} := F_{ij} \{ [f_i(\alpha, y); m_j(\alpha', y)]; \alpha, \alpha', y \}.$$

Where  $F_{ij}$  is a positive real-valued function and has compact support with respect to  $(\alpha, \alpha')$ . The reproductive age of females spans the interval  $[w_{f_r}, w_{f_s}]$  and  $[w_{m_r}, w_{m_s}]$  for males.

Hence;

$$\text{Supp } F_{ij} \subseteq [w_{f_r}, w_{f_s}] \times [w_{m_r}, w_{m_s}]$$

**1.1.1. Derivation of Birth Equation**

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

$$B_1(\mathcal{Y}) = \int_0^\infty \int_0^\infty (\delta_{11}^{(1)} F_{11} + \delta_{12}^{(1)} F_{12} + \delta_{22}^{(1)} F_{22}) d\alpha d\alpha' \dots\dots\dots (i)$$

$$B_2(\mathcal{Y}) = \int_0^\infty \int_0^\infty (\delta_{12}^{(2)} F_{12} + \delta_{13}^{(2)} F_{13} + \delta_{22}^{(2)} F_{22} + \delta_{23}^{(2)} F_{23}) d\alpha d\alpha' \dots\dots\dots (ii)$$

$$B_3(\mathcal{Y}) = \int_0^\infty \int_0^\infty (\delta_{22}^{(3)} F_{22} + \delta_{23}^{(3)} F_{23} + \delta_{33}^{(3)} F_{33}) d\alpha d\alpha' \dots\dots\dots (iii)$$

And

$$\beta = \sum_{n=1}^3 B_n(\mathcal{Y})$$

**1.1.2. Properties of  $B_n(\mathcal{Y})$**

1.  $B_n(\mathcal{Y})$  is well defined and non negative  $F_{ij} \geq 0$ . Human beings are a good example of birth flow population in which birth occur continuously over the time interval.
2.  $B_n(\mathcal{Y})$  is a first degree homogeneous function of its arguments i.e  $B_n K F_{ij}(\mathcal{Y}) = K B_n F_{ij}(\mathcal{Y})$ .
3.  $B_n(\mathcal{Y}) = 0$  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:

$$\hat{B}_1(\mathcal{Y}) = \int_0^\infty \int_0^\infty (\delta_{11}^{(1)} F_{11} + \delta_{12}^{(1)} F_{12}) d\alpha d\alpha' < B_1(\mathcal{Y}) \dots\dots\dots (iv)$$

$$\hat{B}_2(\mathcal{Y}) = \int_0^\infty \int_0^\infty (\delta_{12}^{(2)} F_{12} + \delta_{13}^{(2)} F_{13}) d\alpha d\alpha' < B_2(\mathcal{Y}) \dots\dots\dots (v)$$

$$\hat{B}_3 = 0 \dots\dots\dots (vi)$$

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.

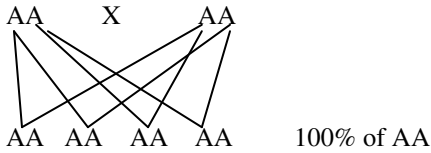
**1.2. Markov Chain Representation**

We describe the Markov chain as: Let  $S = \{S_1, S_2, \dots, S_n\}$  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 (Akanbi, 2006).

1.2.1. Possible Outcome

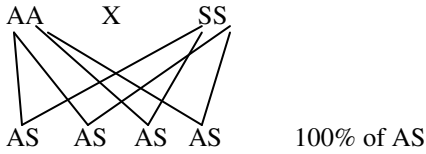
Case I



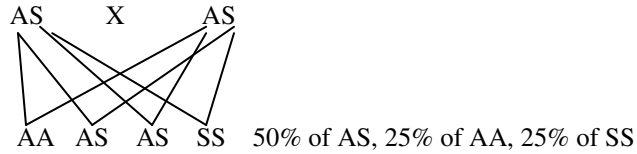
Case II



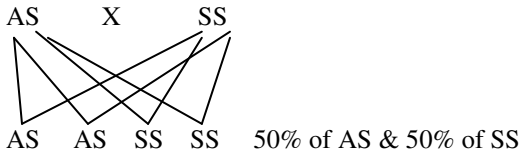
Case III



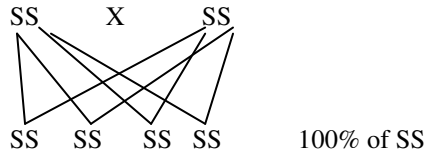
Case IV



Case V



Case VI



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

Parents		AA	AS	SS
AA X AA	$\begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 1 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}$	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	$\begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 1 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/2 & 1/2 \end{pmatrix}$	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

1.2.2. 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;

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

Fij ≥ 0 ∀ 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 is 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

$$F_{ij} = \begin{matrix} & \begin{matrix} AA & AS & SS \end{matrix} \\ \begin{matrix} AA \\ AS \\ SS \end{matrix} & \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.25 & 0.5 & 0.25 \\ 0 & 0.5 & 0.5 \end{pmatrix} \end{matrix}$$

Case II:

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

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

## 2. Analysis & Results

### 2.1. Various Calculation done in this Section was Executed by Matlab Mathematical Software

Case I

To obtain the stationary matrix

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

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

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.

### 2.2. Long Run Behavior

For case I:

The limit of transition matrix at the 20<sup>th</sup> generation is obtained as;

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

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;

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

That is, if another 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? Every individual cannot marry AA since love is dynamic

## 3. 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 abundant 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 1<sup>st</sup> case we realized that 50% of the population will be AS in the long run. For the 2<sup>nd</sup> 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.

#### 4. References

- i. Akanbi O.O 2006: Discrete and continuous-time models of the genetic structure of Sickle-Cell Anaemia. M.Sc. Thesis University of Ibadan.
- ii. Chelsea Liddell, Nicci Owusu-Brackett, Dorothy Wallace September 22, 2011: "A mathematical model of Sickle cell genome frequency in response to selective pressure from malaria. Darmouth College, Hanover, NH, U.S.A.
- iii. Commenses D, soly' P, Letenneur L, Dartifues J. (2004). Incidence and Mentality of Alzheim's Diseases or demand using an illness-death model. Stat med. (23) 199-210.
- iv. Das B, Sengupta. S (2003): A note on some Morphogenetics variables among the Sonowal Kacharis of Assam. Antropologist 5(3): 211-212.
- v. Dehane S.M, Bielefeld R.A, cauthan G.M, Daniel T.M, Rowland D.Y. (2000). Multivariate Markovian Modeling of Tuberculosis. Forecast for the United States. Energy infect dis (6) 148-457
- vi. Eunha Shim, Zhilan Feng and Carlos Castillo-Chavez Dec 12, 2013: "Differential impact of Sickle Cell Trait on Symptomatic and Asymptomatic malaria". Math Biosci Eng; 9(4);877-898.
- vii. Fifty-ninth World Health Assembly Provisional Agenda item 11.4, 24<sup>th</sup> April 2006.
- viii. Herrick J.B. (1910): Peculiar Elongated and sickle shaped red corpuscles in a case of severe anaemia. Arch. Int med 6: 516-518.
- ix. Jackson CH, Sharpless L.D. Hidden Markov models for the onset and Profession of Bronchiolitis Oblite and synonyms in ling transplant recipients. Stat. med (21) 113-128.
- x. Kathleen A. Neville, Julie A panepinto 21-25 march 2011: pharmacotherapy of SCD: 18<sup>th</sup> expert committee on the selection and use of essential medicines.
- xi. Kay R.A (1981). A Markov model for Analysis Cancer Markovs and Disease state in Survival Studies. Biometrics (H2) 855-865
- xii. Kuo H.S, Chang H.J, chou P, Tengh, Chan T.A (1999) "A Markov chain Model to assess the efficiency of screening for non-insulin dependent diabetes meccitus (N1ddm) hut. J Epidemiol. (28) 233-340
- xiii. Lamba J. K, (2002): Common allelic variants of cytochrome P4503A4 and their prevalence in different populations. Phamacogenetics 12:121-132
- xiv. Lamba J. K, (2002): Common allelic variants of cytochrome P4503A4 and their prevalence in different populations. Phamacogenetics 12:121-132
- xv. Seeley RR, Stephens T.D, Tate P (1998): Anatomy and physiology 4<sup>th</sup> edition. The McGraw Hill Companies, Inc. USA Pp 1098.
- xvi. Sergeant GR 1985: Sickle Cell Disease Oxford University Press, London pp 1-25.
- xvii. Sowunmi C.O.A 1993: "A model of Heterosexual population Dynamics with Age structure and gestation period, "Academy Press, Inc.
- xviii. Sweeting M.J, Farewell V.T, De Angelis.D. (2010). Multi state marker models for disease profession in the presence of informative examinations: an application to hepatitis C. state med. (29) 1161-11674
- xix. Tchuene J.M 2003: "A note on the Mackendrick Von Foerster type model in a population with genetic structure". Int.J. math. Educ.Sci Technol. Vol. 34, No. 3 pp 463-469.
- xx. Tchuene J.M 2005: "Realistic Patterns of inheritance of Sickle-Cell Anaemia Gene: A theoretical Approach.
- xxi. Tchuene J.M. 2002: Mathematical population Dynamics of Sickle- Cell Anaemia, Ph.D. thesis University of Ibadan.
- xxii. U.S. Census Bureau Jan 2013.