

A Mathematical Model to Control the Transmission of Thalassemia Disease using Pure Fractions

S. Thakur^{1*}, S. N. Raw¹ and R. Sharma²

¹Department of Mathematics, National Institute of Technology Raipur, Raipur - 492010, Chhattisgarh, India; sapnarajput85@gmail.com, shardaraw@gmail.com

²Department of Electrical and Information Engineering, School of Informatics, Engineering and Technology, Regent University College of Science and Technology, Accra, Ghana, West Africa; dravindrasharma@yahoo.com

Abstract

Objectives: In this paper, we apply the concept of pure fractions to create a mathematical model for the control of Thalassemia disease. **Methods:** The theory of pure fractions has generated various properties that make it suitable for formalizing the uncertain information upon which medical diagnosis and treatment is usually based. In this study we use pure fraction to generate a mathematical model for Thalassemia disease diagnosis. **Findings:** Thalassemia disease is one of the medical problems which could be controlled by premarital screening. This disease has major sign and symptoms in the first year of life. The objective of the work is to diagnose Thalassemia using pure fractions. The proposed model would help the health center to automate Thalassemia risk in future generation and to improve the medical care. **Application:** The proposed method is apply to various other genetic diseases such as G6PD deficiency etc.

Keywords: Pure Fraction, Reducibility of Pure Fractions, Thalassemia Disease, Unit- Interval

1. Introduction

Thalassemia is also known as Cooley's Anemia and Mediterranean Anemia, see the following references¹⁻⁴ for more on studies of the Thalassemia-gene, and references⁵⁻¹¹ for an introduction and complication of Thalassemia disease in children. In this paper, we use partial fractions to create a mathematical model for Thalassemia disease diagnosis. The rest of the paper is organized as follows: in Section 2, definitions and propositions of pure fractions are described. In Section 3, define intervals for Thalassemia genes and Section 4, we create a mathematical model for Thalassemia disease. Also results, discussion and conclusion of the mathematical model are given in Sections 5 and 6 respectively.

2. Preliminaries

The theory of pure fraction originates from the work on the application of fuzzy sets¹². In this section there are a number of definitions and propositions are defining which will be used for the control of the transmission of Thalassemia disease. In this section, we will give some known and useful definitions and notations regarding pure fractions. See¹³ for the definitions and notions of pure fractions.

Definition 2.1 If s is called a pure fraction define in the unit interval $[0,1]$ then s can be represented in the form:

$$s = \frac{r + \alpha - 1}{\alpha}$$

Where $r \in [0,1]$ and α is a positive real number.

*Author for correspondence

Definition 2.2 Reducibility of the pure fractions is generated by the two subintervals of $[0,1]$, i.e., irreducible pure fractions are defined in the subinterval $[0, \frac{1}{2}]$ and reducible pure fractions are defined in the subinterval $[\frac{1}{2}, 1]$. Let R^+ is the set of non-negative real numbers, then for a given pure fraction and a positive real number α , we define $r^{[\alpha]}$ by:

$$r^{[\alpha]} := \frac{r + \alpha - 1}{\alpha} \tag{1}$$

Every pure fraction can be represented in the form (1). If $r^{[\alpha]} \notin [0,1]$ then it is called indeterminate.

Definition 2.3 A given pure fraction has more than one representation of the form given by form (1). Then we call α the index of s relative to r , denoted by $i(s,r)$. Thus, from (1),

$$\alpha = i(s,r) = \frac{1-r}{1-s}$$

Proposition 2.1 Let $s,r,t \in [0,1]$ and $\alpha,\beta \in R^+$, then we have the following properties:

1. $t = r^{[\alpha]} \Leftrightarrow \alpha = i(t,r)$
2. $r^1 = r$ and $1^{[\alpha]} = 1$.
3. $r^{[\alpha]} < r^{[\beta]}$ whenever $\alpha < \beta$ and $i(t,r) < i(u,r)$ whenever
4. $(r^{[\alpha]})^{[\beta]} = r^{[\alpha\beta]} = r^{[\beta\alpha]}$.
5. $i(t,r) = i(r,t)^{-1}$.

Proof

For the proof of all properties see¹³.

Definition 2.4 When $s = r^{[n]}$ and n is a positive integer, we call s a power of r , and r a root of s .

Definition 2.5 Two sequences $\{x_n\}$ and $\{y_n\}$ of pure fractions are said to constitute a Genetic Remediation Scheme (GRS) if:

- x_1 is reducible.
- $y_m = x_m^{[\frac{1}{2}]}, m = 1, 2, \dots$
- $x_{m+1} = \frac{1}{2}(x_m + y_m)$.
- The iteration stops when x_N is reducible for some positive integer N

Theorem 2.1 The sequences $\{x_m\}$ and $\{y_m\}$ of pure fractions that constitute a GRS are decreasing and finite. In addition, for each k such that $\{y_k\}$ is reducible we have $y_{k+1} \geq y_k^{[\frac{1}{2}]}$.

Prove. It is given that there are two sequences with a GRS property, then we have y_k such that:

$$y_k = x_k^{[\frac{1}{2}]} = 2x_k - 1 < 2x_k - x_k = x_k \quad \therefore x_k^{[\frac{1}{2}]} = 2x_k - 1$$

$$\Rightarrow x_{k+1} = \frac{1}{2}(x_k - y_k) < x_k.$$

Therefore, from Proposition 2.1, we have

$$y_{k+1} = x_{k+1}^{[\frac{1}{2}]} < x_k^{[\frac{1}{2}]} = y_k$$

$$\Rightarrow y_{k+1} < y_k.$$

Therefore, the given sequences are decreasing.

To prove the given sequences are reducible we have

$$y_k < x_k$$

$$\Rightarrow y_k < \frac{1}{2}(x_k + y_k) = x_{k+1}.$$

$$\Rightarrow y_k^{[\frac{1}{2}]} < x_{k+1}^{[\frac{1}{2}]}$$

Whenever y_k is reducible.

Therefore, both the sequences are decreasing and reducible.

Now to prove the given sequences are finite let us suppose that there exist no positive integer N such that x_k is irreducible. Then $x_k \geq \frac{1}{2}, \forall n$.

Previously we have shown that both the sequences are decreasing and reducible or bounded, hence both the sequences are converging to some pure fractions.

Hence x_k converges to some pure fraction s and y_k converges to some pure fractions t .

Consequently,

$$s = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} \left[\frac{1}{2(x)_n} + y_n \right] = \frac{1}{2(s+t)}$$

Thus $s = t$... (2)

$$t = \lim_{n \rightarrow \infty} y_n = \lim_{n \rightarrow \infty} [2(x)_n - 1] = 2s - 1$$

EMBED Equation. 3 $s = 1$

since $s \leq x_1 < 1$

Therefore, we have a contradiction. It completes the proof.

3. Thalassemia and Pure Fractions

The present work introduces a simple and effective methodology for Thalassemia diagnosis based on the theory of the pure fractions. The developed methodology is suitable for application in a much wider range of Thalassemia disease. We define the concept of pure

fractions relevant to the considered genetic diseases T (Thalassemia). In this article a mathematical model is created to handle the prevalence of the thalassemia gene in which the concept of reducibility of pure fractions is used.

Since Thalassemia gene is inherited by the children in a recessive manner. Then from Principles of Inheritance¹⁴ (Mendel's Laws) Genetic mutations can occur in the gene that codes for normal Hemoglobin A (**HbA**). Thalassemia is an example of such a mutation. The disease is mainly due to the production of abnormal Hemoglobin a (**Hba**) instead of Hemoglobin A, and it occurs in people who are homozygous for the recessive Thalassemia major gene (**aa**) pair. Therefore, to understand the occurrence of Thalassemia disease on an individual having a gene known as the Thalassemia gene and a genotype known as the Thalassemia genotype (**aa**). There are three types of genotypes encountered in studies of Thalassemia disease. They are called **AA**, **Aa**, and **aa**. The effect of genotype **AA** is that the carrier has no **a**-gene, no fear of ever developing Thalassemia disease, and no fear of ever transmitting the **a**-gene to offspring. The carrier of the genotype **Aa** is a heterozygote who cannot suffer from Thalassemia disease, but can transmit the **a**-gene to offspring. The carrier of the genotype **aa** is called a Thalassemia major. Such a person will surely develop Thalassemia disease over time and will surely transmit the **a**-gene to offspring. In this article pure fractions representing genotypes are only probabilities of possession and transmission of the gene. The genotypes of Thalassemia disease are represented by the pure fractions $0, \frac{1}{2}$ and 1 corresponding to **AA**, **Aa**, and **aa** respectively.

Consider a Thalassemia disease T and a gene G responsible for the disease T . As done in fuzzy logic for truth values Nguyen and Walker¹⁵, we imagine every number in the unit interval $[0,1]$ as representing a possible genotype for G . That is, each pure fraction $s \in [0,1]$ measures a corresponding degree of possession and transmission of the gene G by an individual denoted by X . We note here that the pure fraction s can also be thought of as the probability of possession and transmission of the gene G by the individual denoted by X . A person with a "high incidence" of the gene G is defined as someone who has a genotype falling in the subinterval $I_2 = [\frac{1}{2}, 1)$. A person with such a genotype may suffer from discrimination when selecting a spouse. However, we make the assumption that such an individual would like to have a spouse so that their descendant of some future generation would be more fortunate and would have a genotype of

G that does not fall in subinterval I_2 . Such an individual is a descendant with a "low incidence" of G and has a genotype that falls in the subinterval $I_1 = [0, \frac{1}{2})$. Based on what is stated above, we will devise a mathematical model, where an individual with a high incidence of the gene G can be matched with the appropriate spouse so that a future descendant can have a low incidence of the gene G . Since our model is intended for a general genetic disease, we have assumed, as done in fuzzy logic for truth values. Thus, every pure fraction s is a possible genotype for the Thalassemia disease T .

4. Mathematical Model for Thalassemia Disease

In the research now a day's several kind of model are used for disease diagnosis such as¹⁶ used weight based k-means algorithm is developed for identifying the leukemia, inflammatory, bacterial or viral infection, HIV infection and pernicious anaemia diseases from the hemogram blood test samples data set. Also¹⁷ proposed a work uses the benefits of genetic algorithms and fuzzy inference system for effective prediction of heart disease in patients. In this section, the theory and properties of pure fractions are used which are defined by Eke in 2008. Let us suppose that G denote a gene which is responsible for the Thalassemia disease T . Also, there is a function $T : H \rightarrow [0,1]$, such that the following assumptions are satisfied

- For a "genotype" $t = T(X)$ the followings are holds:
 - i. $t = 0 \Leftrightarrow X$ is completely free of the gene G .
 - ii. $t \in [0, \frac{1}{2}) \Leftrightarrow X$ has a "low incidence" of gene G .
 - iii. $t \in [0, \frac{1}{2}) \Leftrightarrow X$ has a "high incidence" of gene G .
- If $t = r^n$, then r is also in the range of T .
- The genetic structure of an individual is proportional to that of his/her parents X^0 and Y^0 , hence

$$T(X) = \frac{T(X^0) + T(Y^0)}{2}$$

- If both parents $T(X^0)$ and $T(Y^0)$ are in danger of the Thalassemia disease T then an individual Y will not accept X for a marital partner if

$$T(Y) < \min \left\{ T(X^0)^{|\frac{1}{2}|} + T(Y^0)^{|\frac{1}{2}|} \right\}$$

Since we have the function $T : H \rightarrow [0,1]$, then we can set the condition "marital connections" such that an individual X with genotype $T(X) \in [0,1]$ can be "tolerated

for marriage” by an individual Y only if they can have a descendant Z such that $T(Z) \in [0, \frac{1}{2}]$.

Now to solve the problem “marital connections” suppose that for the any positive integer n , and $x=T(X)$, $\{X_n\}$ and $\{Y_n\}$ are the sequences of individuals with their genotype sequences $\{X_n\}$ and $\{Y_n\}$.

$$T(Y) < \min \left(T(X_n)^{\lfloor \frac{1}{2} \rfloor} + T(Y_n)^{\lfloor \frac{1}{2} \rfloor} \right)$$

Case-Study

Suppose that X_1 be an individual such that $x_1=T(X_1)$, where x_1 is reducible. Then x_1 generates a GRS (Genetic Remediation Scheme), it has two finite sequences $\{x_1, x_1, \dots, x_N\}$ and $\{y_1, y_1, \dots, y_{N-1}\}$, \exists two corresponding sequences of humans $\{X_1, X_1, \dots, X_N\}$ and $\{Y_1, Y_1, \dots, Y_{N-1}\}$. Also for $m = 1, 2, 3, \dots, N - 1$, Y_m is assumed to be the spouse of X_m and X_{m+1} is assumed to be an offspring of X_m and Y_m satisfying,

$$T(X_{m+1}) = x_{m+1} = \frac{1}{2}(x_m + y_m)$$

x_m and y_m are reducible therefore, we have

$$x_{m+1} \geq y_m^{\lfloor \frac{1}{2} \rfloor} = \min(x_m^{\lfloor \frac{1}{2} \rfloor}, y_m^{\lfloor \frac{1}{2} \rfloor})$$

Example 4.1 Let $x_1 = \frac{15}{17}; x_1 \in I_2$ Can be a pure fraction associated with an individual X . Then the following steps can be taken to find out the descendant:

It is given that

$$\begin{aligned} x_1 &= \frac{15}{17} \\ \therefore y_m &= x_m^{\lfloor \frac{1}{2} \rfloor} = 2x_m - 1 \\ \Rightarrow y_1 &= \frac{13}{17} \end{aligned}$$

Here, Iteration is given in the following Table 1.

5. Results and Discussion

We constructed the Pure Fractions for the control of Thalassemia disease. Example 4.1 has been taken to exhibit the simplicity of the technique. Through the computed iteration see Table 1, we have x_m such that for $N' = 5, x_5 < \frac{1}{2}$, hence $x_5 \in I_1$. Therefore, after the V -th generation X_1 and Y_1 will have the descendant X_5 who has a “low incidence” of the Thalassemia gene G with genotype. The motivation for the development of the mathematical model in Section 4 is the study of the genotypes for Thalassemia disease which is a genetic disease. For the general genetic disease D , the people in danger are those with the gene-types $x \in [\frac{1}{2}, 1)$. The genotypes in $[\frac{1}{2}, 1)$ are all reducible pure fractions. The mathematical model created in this paper is for the purpose of controlling the transmission of genes with “high incidence” genotypes to arrive at genotypes $x \in [0, \frac{1}{2})$ with “low incidence” in order to produce heterozygous offspring over future generations. The mathematics of pure fractions was very convenient for this purpose. The theory of pure fractions is quite new in the sense that the authors believe

Table 1. Iterations for computing genotype frequencies

| Generation (N') | x_m | Subinterval I_1 / I_2 | y_m |
|---------------------|---|-------------------------|----------------------------------|
| i. | $x_1 = \frac{15}{17}$ | $x_1 \in I_2$ | $y_1 = 2x_1 - 1 = \frac{13}{17}$ |
| ii. | $x_2 = \frac{1}{2}(x_1 + y_1) = \frac{14}{17}$ | $x_2 \in I_2$ | $y_2 = 2x_2 - 1 = \frac{11}{17}$ |
| iii. | $x_3 = \frac{1}{2}(x_2 + y_2) = \frac{25}{34}$ | $x_3 \in I_2$ | $y_3 = 2x_3 - 1 = \frac{8}{17}$ |
| iv. | $x_4 = \frac{1}{2}(x_3 + y_3) = \frac{41}{68}$ | $x_4 \in I_2$ | $y_4 = 2x_4 - 1 = \frac{7}{34}$ |
| v. | $x_5 = \frac{1}{2}(x_4 + y_4) = \frac{55}{136}$ | $x_5 \in I_1$ | Stops |

this work is the first study to formalize and standardize special structural properties of pure fractions. The theory shows potential for further research and applications. In particular the type of Remediation Scheme presented by Definition 2.5 and Theorem 2.1 can be modified to apply to various other genetic diseases such as Sickle Cell Disease (SCD), G6PD deficiency etc.

6. Conclusion

Thalassemia is silently prevalent in the tribes of Chhattisgarh. The prevalence of Thalassemia (in heterozygous or homozygous state) in tribal population in this region is probably due to selection pressure of endemic malaria in this part of India. In these research article pure fractions plays a very useful role to control the transmission of Thalassemia gene and to create a new generation without Thalassemia Major or severe forms of Thalassemia gene.

7. Acknowledgment

I would like to place on record my deep sense of gratitude to Mr. Pramod Puri, General Secretary of Thalassemia Welfare Society, Bhilai, Chhattisgarh, for his stimulating cooperation, unfailing inspiration and continuous encouragement throughout the completion of present work.

8. References

1. Angastiniotis M, Kyriakidou S, Hadijiminas M. How Thalassaemia was controlled in Cyprus. *World Health Forum*. 1986; 7:291–7.
2. De M, Das SK, Bhattacharya DK, et al. The occurrence of beta thalassaemia mutation and its interaction with Hb E in the eastern India. *Int J Hematol*. 1997; 66(1):31–4.
3. Atkin K, Ahmad WIU. Genetic screening and haemoglobinopathies: Ethics, politics and practice. *Social Science and Medicine*. 1998; 46(3):445–58.
4. Cao A, Galanello R. Beta-Thalassemia. *Genetics in Medicine*. 2010; 12(2):61–76.
5. Cooley TB, Witwer ER, Lee P. Anemia in children with splenomegaly and peculiar changes in the bones. *Am J Dis Child*. 1927; 34:347–63.
6. Weatherall JD, Clegg JB. *The Thalassemia syndromes*. 3rd ed. Oxford: Blackwell Scientific; 1981. p. 132–74.
7. Ghodekar SR. Thalassemia: A review. *International Journal of Pharma Research and Development*. 2010; 2(10):101–8.
8. Trent RJA. Diagnosis of hemoglobinopathies. *Clin Biochem*. 2006; 27(1):27–38.
9. Nienhuis AW, Nathan DG. Pathophysiology and clinical manifestations of the β -Thalassemias. *Cold Spring Harb Perspect Med*. 2012; 2:1–13.
10. Adeyemo TA, Adeyemo WL, Adediran A, et al. Orofacial manifestations of hematological disorders: Anemia and hemostatic disorders. *Indian Journal of Dental Research*. 2011; 22(3):454–61.
11. Madhok S, Madhok S. Dental considerations in Thalassemic patients. *IOSR Journal of Dental and Medical Sciences*. 2014; 13(6):57–62.
12. Eke B. A mathematical simulation of a social norm using fuzzy sets. In: Paul P. Wang editor. *Proceedings of the 7th Joint Conference on Information Sciences*; 2003. p. 126–7.
13. Eke B, Nkwanta A. Mathematical model for the control of the transmission of genetic diseases using pure fractions. *International Journal of Pure and Applied Mathematics*. 2008; 45:243–54.
14. Laird NM, Lange C. *The fundamentals of modern statistical genetics*. Statistics for Biology and Health. Springer Science+Business Media, LLC; 2011. DOI: 10.1007/978-1-4419-7338-2_2.
15. Nguyen HT, Walker EA. *A first course in fuzzy logic*. 3rd ed. Boca Raton: Chapman and Hall – CRC; 2005.
16. Vijayarani S, Sudha S. An efficient clustering algorithm for predicting diseases from hemogram blood test samples. *Indian Journal of Science and Technology*. 2015 Aug; 8(17):521–23.
17. Santhanam T, Ephzibah EP. Heart disease prediction using hybrid genetic fuzzy model. *Indian Journal of Science and Technology*. 2015; 8(9):797–803.