

Breaking the Transmission Chain of Thalassemia: A Predictive Approach Using Controlled Marriages

Farhani^{1*}, Saiful Amri², Muhammad Ikhwan³, Dara Irsalina⁴, Reza Wafdan⁵

^{1,2,3,4,5} Program Studi Matematika, Universitas Syiah Kuala, Aceh, Indonesia

*Corresponding author: farhani.math@usk.ac.id

*Submission date: 25 September 2025, Revision: 01 October 2025, Accepted: 24 January 2026

ABSTRACT

This study aims to analyze the probabilities of thalassemia gene types in the n -th generation under controlled marriage conditions. Marriage is controlled by assume at least one of the parent is normal (neither has thalassemia minor nor thalassemia major). To analyze the probability, we begin by constructing a probability matrix of thalassemia inheritance, then we determine the eigenvalues and eigenvectors for matrix diagonalization, and then we use those eigen values and those eigen vectors to compute the probability matrix for n -th generation. This approach is employed to predict the distribution of thalassemia probability in the n -th generation. The findings indicate that under controlled marriage conditions, the probability thalassemia matrix shows that thalassemia major will be completely eliminated by the first generation, thereby preventing its occurrence in subsequent generations. Furthermore, over successive generations, the proportion of thalassemia minor probability will gradually decline and eventually approach zero and probability for normal type getting bigger that eventually almost certain to happen. It means, for far generations, it almost impossible to have thalassemia minor.

KEYWORDS

Thalassemia, Eigen Vector, Matrix Diagonalization.

1. INTRODUCTION

Thalassemia is a hereditary disorder that affects the production of normal hemoglobin [1]. It is inherited in an autosomal recessive manner from parents to their children [2, 3, 4]. Clinically, Thalassemia ranges from thalassemia minor (carrier) to thalassemia major [5]. Thalassemia major is the most severe clinical form of thalassemia. Similar to thalassemia major, individuals with thalassemia intermedia also experience abnormalities in two chromosomes inherited from both parents. The difference lies in the type of mutant gene that is passed down. Thalassemia minor can also be referred to as a carrier, trait, mutant carrier, or thalassemia carrier [6].

Autosomal recessive single-gene disorders arise only in individuals carrying two mutant alleles, one inherited from each parent [7]. If both parents are affected by thalassemia, all offspring will inherit the disease. If only one parent is affected, all children will become carriers [8]. Consequently, preventing thalassemia major requires strategies that limit marriages between two carriers, as such unions carry a high risk of producing affected offspring [9, 10].

Because marriages involving thalassemia carriers could transmit thalassemia genes [11], controlling marriages so that at least one partner is genetically normal may reduce disease propagation. This study aims to determine the probability distribution of thalassemia genotypes across the n -th generation under controlled marriage conditions.

2. LITERATURE REVIEW

We will conduct how thalassemia will be transmitted by observing how genes inherited in general and then we will predict the transmission using matrix diagonalization. Since we will conduct prediction for long time generation, we will use information

about limit at infinity.

2.1 Inheritance of genes

A gene is a unit of information that translates genetic characteristics. Genes are located within chromosomes. Chromosomes consist of autosomes and gonosomes. Suppose gene A produces trait A, and if it mutates—denoted as a—the resulting trait is also denoted as a. A and a are called alleles. A person's genetic condition can be described using a genotype, which usually consists of a pair of alleles, for example, denoted as AA. The genetic information a person possesses is called the genotype, while the observable characteristics resulting from a certain genotype are called the phenotype. The principle of gene segregation states that the two alleles from each parent will separate, and each allele will be passed on to each gamete [12]. For example, if the parents have genotypes AA and Aa, then their offspring may inherit the genotypes AA (with a probability of 50%) and Aa (with a probability of 50%).

2.2 Matrix Diagonalization

2.2.1 Matrix

A matrix is a collection of numbers arranged in a rectangular form and enclosed by parentheses or square brackets. The numbers inside the matrix are called entries. The size of a matrix is determined by the number of rows and columns, and the size is also referred to as the order of the matrix. For example,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

The matrix above can be written as $A(3 \times 2)$ because matrix A has three rows and two columns. A matrix consisting of only one row or one column is also called a vector [13].

2.2.2 Eigen Value and Eigen Vector

If A is an $n \times n$ matrix, then a nonzero vector x in R^n is called an eigen vector of A if Ax is a scalar multiple of x , which can be written as

$$Ax = \lambda x \quad (1)$$

for some scalar λ . Scalar λ is called an eigen value A, and x is called eigen vector of A corresponding to λ . Eigen value can be found by solving the equation $\det(\lambda I - A) = 0$, then each of the values is substituted to equation $Ax = \lambda x$, to get a corresponding eigen value [14].

2.2.3 Matrix diagonalization

A square $n \times n$ matrix A is said to be diagonalizable if there exists an invertible matrix P and diagonal matrix D such that

$$A = PDP^{-1} \quad (2)$$

In fact, $A = PDP^{-1}$, if only if the columns of P are n linearly independent eigenvectors of A [15].

2.2.4 Limit at infinity

$\lim_{x \rightarrow c} f(x) = L$ means that when x is near but different from c , then $f(x)$ is near to L . For limit $f(x)$ with x to infinity, is denoted by $\lim_{x \rightarrow \infty} f(x)$, symbol $x \rightarrow \infty$ is used to show that x gets larger and larger without bound [16].

If $r > 0$ is a rational number, then

$$\lim_{n \rightarrow \infty} \frac{1}{x^r} = 0 \quad (3)$$

By taking x large enough, we can make $\frac{1}{x}$ close to 0 as we please [17]

3. METHODOLOGY

This type of research is mathematics research using the literature study method, in which the material used consists of books and journals discussing eigenvalues and eigenvectors, matrix diagonalization, as well as limits and the application of matrix diagonalization in the field of genetics. In this paper, we assume that the transmissions of thalassemia only occurred from inherited genes, omitting the genetic mutation case. The steps in conducting this research are explained in detail as follows.

1. Constructing the probability matrix
2. Finding the eigen values and eigen vectors.
3. Diagonalizing the probability matrix.
4. Determining the probability for n-th generation.

4. RESULT & DISCUSSION

In this study, the thalassemia gene type is denoted by T for the dominant gene and t for the recessive gene, with possible allele pairs being TT, Tt, and tt. TT represents the dominant gene type that does not suffer from thalassemia (normal), Tt represents the carrier/thalassemia minor gene type that may pass on thalassemia to offspring, and tt represents the thalassemia gene type, also known as thalassemia major.

4.1 Constructing the Equation in a Matrix

If one of the parent is normal, we could have pairing type such as TT-TT, TT-Tt, Tt-TT, TT-tt and tt-TT. In this paper, we assume that TT-Tt is same as Tt-TT and TT-tt and tt-TT are also same, because we only focus on the probability of the next generation and such pairings will inherit same type and we assume they will inherit the same probability. Therefore, we only consider TT-TT, TT-Tt and TT-tt marriage.

According to the principle of gene segregation, TT-TT will inherit TT only, TT-Tt marriage will inherit TT and Tt, and TT-tt will produce Tt only. To see how probability for each cases, we will make punnet square of each marriage types. A Punnett square illustrates how alleles combine in offspring [18]. These are the probability of pairing combination.

1. $TT - TT$

Table 1. Punnet square for $TT - TT$

	T	T
T	TT	TT
T	TT	TT

According to **Table 1**, every pairing of alleles will inherit TT type only, therefore we will write probability for TT is 1, and probability will be zero for Tt and tt .

2. $TT - Tt$

Table 2. Punnet square for $TT - Tt$

	T	T
T	TT	TT
t	Tt	Tt

According to **Table 2**, two of four possible genotype will be TT , and another two will be Tt genotype, therefore in this marriage, probability for TT is $\frac{1}{2}$, probability for Tt is $\frac{1}{2}$, probability for tt will be 0.

3. $TT - tt$

Table 3. Punnet square for $TT - tt$

	T	T
t	Tt	Tt
t	Tt	Tt

According to **Table 3**, from this marriage, the possible genotype is only Tt . So, the probability for Tt in this marriage case is 1, and the probability for TT and tt will be 0.

All of probability distribution could be shown as following table. Suppose,

Table 4. Distribution probability of genotype in each marriage case

Genotype	$TT \times TT$	$TT \times Tt$	$TT \times tt$
TT	1	$\frac{1}{2}$	0
Tt	0	$\frac{1}{2}$	1
tt	0	0	0

a_n = The probabily nth generation of type TT ,

b_n = The probabily nth generation of type Tt ,

c_n = The probabily nth generation of type tt

Based on **Table 4**, we can say that in the first generation, because TT is obtained from the genotype TT and Tt of the previous offspring, the probability of genotype occuring is

$$a_1 = a_0 + \frac{1}{2} b_0$$

For the same reason, it is obtained $b_1 = \frac{1}{2} b_0 + c_0$.

We can see that, none of the marriage type will produce tt . Hence, there are no offspring with the tt genotype (no thalassemia mayor patients). For second generation, it only has TT and Tt genotype, therefore we will get the probability as **Table 5**. The

Table 5. Distribution probability for second generation

Genotype	$TT \times TT$	$TT \times Tt$
TT	1	$\frac{1}{2}$
Tt	0	$\frac{1}{2}$

mariage still has been controlled which require at least one parent is normal (not a carrier).

Note that in the second generation, TT probability will depend on both $TT - TT$ and $TT - Tt$ marriage. It means TT will depend on previous TT and Tt probability, with probability

$$a_2 = a_1 + \frac{1}{2} b_1 \quad (4)$$

Tt probability for second generation only depend $TT - Tt$ marriage, so it depends only on Tt probability that will be this equation.

$$b_2 = \frac{1}{2} b_1 \quad (5)$$

From equation (4) and (5), we could construct this matrix

$$\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 + \frac{1}{2}b_1 \\ \frac{1}{2}b_1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \quad (6)$$

It means that TT type for n -th generation (a_n) depend on a_{n-1} and b_{n-1} because TT will obtained from the marriage of a_{n-1} and b_{n-1} . Meanwhile, b_n only will be obtained b_{n-1} . Thus, the transition of genotype distribution from generation to generation can be written in the following equation

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = M \begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix} \quad (7)$$

with $M = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}$

By mathematical induction, the vector can be written as

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = M^n \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} \quad (8)$$

4.2 Eigen Values and Eigen Vectors

For compute the matrix M^n in equation (8), we need to find eigen values and eigen vectors from matrix M . Computing with phyton program, Eigen values of matrix M are $\lambda_1 = 1$ and $\lambda_2 = \frac{1}{2}$. And then, we use equation (1) or $(\lambda I - A)x = 0$ to find eigen vectors, we found that the corresponding eigen vectors are $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

4.3 Matrix M Diagonalization and Finding the Values of a_n and b_n

For diagonalizing matrix $M = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{pmatrix}$, it will need eigen values for diagonal matrix entries and eigen vector for matrix P and then finding the inverse matrix of P . Eigen values and eigen vectors of the matrix M are $\lambda_1 = 1$ and $\lambda_2 = \frac{1}{2}$ with corresponding eigen vectors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

So, it obtain

$$\begin{aligned} M &= PDP^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

Because

$$\begin{aligned} M^2 &= PDP^{-1}PDP^{-1} \\ &= P \cdot D \cdot I \cdot D \cdot P^{-1} \\ &= P \cdot D^2 P^{-1} \end{aligned}$$

then by induction, it will be $M^n = PD^nP^{-1}$. Therefore,

$$\begin{aligned} \begin{pmatrix} a_n \\ b_n \end{pmatrix} &= M^n \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} \\ \begin{pmatrix} a_n \\ b_n \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}^n \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (\frac{1}{2})^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 - (\frac{1}{2})^n \\ 0 & (\frac{1}{2})^n \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \end{aligned}$$

So,

$$a_n = a_1 + b_1 - \left(\frac{1}{2}\right)^n b_1 \quad (9)$$

$$b_n = \left(\frac{1}{2}\right)^n b_1 \quad (10)$$

For the first generation,

$$a_1 = a_0 + \frac{1}{2} b_0$$

$$b_1 = \frac{1}{2} b_0 + c_0$$

At the initial generation, before controlled marriage happen, all the types of genotype could be happen, so

$$a_0 + b_0 + c_0 = 1$$

as a result at the first generation,

$$\begin{aligned} a_1 + b_1 &= a_0 + \frac{1}{2} b_0 + \frac{1}{2} b_0 + c_0 \\ &= a_0 + b_0 + c_0 \\ &= 1 \end{aligned}$$

Therefore,

$$\begin{aligned} a_n &= 1 - \left(\frac{1}{2}\right)^n b_1 \\ b_n &= \left(\frac{1}{2}\right)^n b_1 \end{aligned}$$

and for n to infinity, using equation (3), we will get

$$\begin{aligned} \lim_{n \rightarrow \infty} a_n &= \lim_{n \rightarrow \infty} \left(1 - \left(\frac{1}{2}\right)^n b_1\right) \\ &= \lim_{n \rightarrow \infty} 1 - \lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^n \cdot \lim_{n \rightarrow \infty} b_1 \\ &= 1 - 0 \cdot b_1 \\ &= 1 \end{aligned}$$

Hence, $a_n \rightarrow 1$.

$$\begin{aligned} \lim_{n \rightarrow \infty} b_n &= \lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^n b_1 \\ &= \lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^n \cdot \lim_{n \rightarrow \infty} b_1 \\ &= 0 \cdot b_1 \\ &= 0 \end{aligned}$$

Hence, $b_n \rightarrow 0$. To sum up, for n to infinity, $a_n \rightarrow 1$ and $b_n \rightarrow 0$. We can say that probability for TT getting close to 1 and probability for Tt getting close to zero as n becomes large. **So, for far generations, TT almost certain to happen and Tt becomes almost impossible to happen. On another words, probability for Thalassemia minor will be smaller that event almost impossible to happen and normal type almost certain to happen.**

5. CONCLUSION

Based on the results of the study, it can be concluded that:

1. The inheritance of the thalassemia gene through controlled marriage (at least one parent has the TT gene type) will prevent thalassemia patients from appearing in the first generation.
2. In the n -th generation, it will satisfy the following equation

$$a_n = 1 - \left(\frac{1}{2}\right)^n b_1$$

$$b_n = \left(\frac{1}{2}\right)^n b_1$$

$$\text{with } b_1 = \frac{1}{2}b_0 + c_0$$

3. For n to infinity,

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \left(1 - \left(\frac{1}{2}\right)^n b_1\right) = 1, a_n \rightarrow 1$$

$$\lim_{n \rightarrow \infty} b_n = \lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^n b_1 = 0, b_n \rightarrow 0$$

Hence, for far generations, TT almost certain to happen and Tt becomes almost impossible to happen. On another words, probability for Thalassemia minor will be smaller that event almost impossible to happen and normal type almost certain to happen.

REFERENCES

- [1] D. Ahirrao and N. Desale, *A Textbook of Genetics*. Harshwardhan Publication, Aug. 2019.
- [2] D. Rediyanto, "Deteksi dini thalassemia," *Majalah Pratista Patologi*, vol. 8, no. 1, Jan. 2023.
- [3] Pratama and I. Kurniati, "Pendekatan Diagnosis Berbasis Molekuler pada Pasien Talasemia," *Medula*, vol. 9, no. 2, pp. 339–345, Jul. 2019.
- [4] R. Galanello and R. Origa, "Beta-thalassemia," *Orphanet J Rare Dis*, vol. 5, no. 1, p. 11, Dec. 2010.
- [5] J. Regar, "Aspek Genetik Talasemia," *JBM*, vol. 1, no. 3, Feb. 2013.
- [6] L. Rujito, *Buku Referensi Talasemia : Genetik Dasar dan Pengelolaan Terkini*, 11 2019.
- [7] H. Mustami and C. Muthiadin, *Konsep Dasar Pewarisan Gen Pada Manusia Disertai Hasil Riset Dan Analisis Resiko Mendel*. Makassar: Alauddin University Press 2021.
- [8] R. Panigoro, D. Prihatni, Y. Sribudiani, M. Ghazali, A. M. Maskoen, S. Susanah, D. K. Turbawati, R. Purwanti, and E. Sahiratmadja, "Upaya Menurunkan Angka Kejadian Talasemia Mayor Melalui Edukasi Dan Pemeriksaan Darah Pada Acara Donor Darah: Menuju Zero Growth Talasemia Mayor di Indonesia," *Dharmakarya: Jurnal Aplikasi Ipteks untuk Masyarakat*, vol. 12, no. 2, pp. 182–188, Jun. 2023.
- [9] A. d'Arqom, *Talasemia Ditinjau dari Segi Medis dan Sosial*. Airlangga University Press, Aug. 2023.
- [10] N. Ain, "Kajian Keluarga Dalam Upaya Pencegahan Penyakit Thalassemia Mayor Di Indonesia : Family Study In The Effort To Prevent Thalassemia Major In Indonesia," *Jurnal Kedokteran STM (Sains Dan Teknologi Medik)*, vol. 8, no. 2, pp. 185–193, 2025.

- [11] M. S. Hossain, F. Islam, S. Akhter, and A. Al Mossabbir, “Thalassemia in Bangladesh: progress, challenges, and a strategic blueprint for prevention,” *Orphanet J Rare Dis*, vol. 20, no. 1, p. 358, Jul. 2025.
- [12] B. A. Pierce, *Genetics: a conceptual approach*, 4th ed. Basingstoke: Palgrave, 2012.
- [13] H. Anton and C. Rorres, *Elementary linear algebra: applications version*, 11th ed. Hoboken, NJ: John Wiley & Sons Inc, 2014.
- [14] H. Anton and A. Kaul, *Elementary Linear Algebra*, 12th ed. Wiley, 2019.
- [15] D. C. Lay, S. R. Lay, and J. J. McDonald, *Linear Algebra and Its Applications*, 5th Edition, 2016.
- [16] D. E. Varberg, E. J. Purcell, and S. E. Rigdon, *Calculus: early transcendentals*. Upper Saddle River, N.J: Pearson/Prentice Hall, 2007.
- [17] J. Stewart, *Single variable calculus: early transcendentals*, 6th ed. Australia ; Belmont, CA: Thomson Brooks/Cole, 2008.
- [18] R. Lewis, *Human genetics: concepts and applications*, 9th ed. New York: McGraw-Hill, 2010.