Monte Carlo simulation on the effect of different approaches to thalassaemia on gene frequency

F. Habibzadeh[1] and M. Yadollahie[1]

نموذج مونت كارلو الحاسوبي لمحاكاة تأثير أساليب مختلفة
للوقاية من الثلاسيمية على تواتر الجين المسبّب لها

فرّخ حبيب زاده، محبوبه يد اللهي

الخلاصـة: قامت الباحثتان في إطار هذه الدراسة باستخدام طريقة المحاكاة الحاسوبية لتحديد الاختلاف في الجين، وتكرار الزيجوت المتغاير الألائل، وتكـرار الزيجوت المتماثل الألائل، الناتجـة عـن 4 أسـاليب مختلفـة للوقايـة مـن الثلاسيمية. وتشمل هذه الأساليب: العلاج الداعم فقط؛ معالجة المرضى الحاملين للزيجوت المتماثل الألائل بطريقة النمط الظاهري فقط؛ إجهاض جميع الأجنّة الحاملة للزيجوت المتماثل الألائل؛ ومنع الـزواج بـين حـاملي الجـين. وبيّنت الدراسة ثبات تواتر الجين عند اتّباع الاستراتيجية الثانية أو الرابعة، وانخفاض تـواتر الجـين مـع الاسـتراتيجية الأولى أو الثالثة. ولوحظ أيضاً أن تكرار الزيجوت المتغاير الألائل يختلف بشكل متواز مع تواتر الجين. وعند تطبيـق الاستراتيجية الأولى، لوحظ انخفاض تواتر الزيجوت المتماثل الألائل بمرور الوقت، في حـين يثبت هـذا التـواتر مـع الاستراتيجية الثانية، وينخفض عند تطبيق الاستراتيجية الثالثة أو الرابعةِ إلى صـفر بعـد الجيـل الأول. وأيّـاً كانـت الاستراتيجية المتّبعة، فإنه في أسوأ الافتراضات، سيظل جين الجمهرة ثابتاً بمرور الوقت.

ABSTRACT We used computer simulation to determine variation in gene, heterozygous and homozygous frequencies induced by 4 different approaches to thalassaemia. These were: supportive therapy only; treat homozygous patients with a hypothetical modality phenotypically only; abort all homozygous fetuses; and prevent marriage between gene carriers. Gene frequency becomes constant with the second or the fourth strategy, and falls over time with the first or the third strategy. Heterozygous frequency varies in parallel with gene frequency. Using the first strategy, homozygous frequency falls over time; with the second strategy it becomes constant; and with the third and fourth strategies it falls to zero after the first generation. No matter which strategy is used, the population gene frequency, in the worst case, will remain constant over time.

Simulation de Monte Carlo pour l'effet sur la fréquence génique des différentes approches face à la thalassémie
RÉSUMÉ Nous avons recouru à la simulation par ordinateur pour déterminer la variation du gène, les fréquences homozygotes et hétérozygotes induites par 4 approches différentes face à la thalassémie, à savoir : traitement de soutien seulement ; traitement des patients homozygotes sur le mode hypothétique au niveau phénotypique seulement ; avortement de tous les fœtus homozygotes ; et prévention du mariage entre les porteurs du gène. La fréquence génique devient constante avec la deuxième ou la quatrième stratégie, et diminue dans le temps avec la première ou la troisième stratégie. La fréquence hétérozygote varie parallèlement à la fréquence génique. En utilisant la première stratégie, la fréquence homozygote diminue dans le temps ; avec la deuxième stratégie, elle devient constante et avec les troisième et quatrième stratégies, elle tombe à zéro après la première génération. Indépendamment de la stratégie utilisée, la fréquence génique dans la population restera constante avec le temps, dans le pire des cas.

[1]*National Iranian Oil Company Medical Education and Research Centre, Shiraz, Islamic Republic of Iran (Correspondence to F. Habibzadeh: habibzaf@sums.ac.ir).*
Received: 12/11/03; accepted: 25/07/04

Introduction

Thalassaemia is a dreadful inherited disease affecting tens of thousands of people worldwide. It is the commonest single-gene, autosomal, hereditary disease [1]. Thalassaemia is more prevalent among people of the Mediterranean area, parts of North and West Africa, the Middle East, the Indian subcontinent, southern Far East and south-eastern Asia, the so-called "thalassaemia belt". The thalassaemia gene frequency is higher in some regions; it is around 10% in Sicily, 5%–15% in Greece and 4%–10% in the Islamic Republic of Iran [1,2]. This haemoglobinopathy, even though it confers some protection against fatal falciparum malaria, has now become a socioeconomic burden for many countries [2,3].

With the exception of the extremely rare mutations, only those individuals born to 2 gene carriers (heterozygotes or homozygotes) may develop thalassaemia major. So far, several strategies have been proposed to deal with the burden this disease places on the community. However, only a few approaches are applicable and currently in use. These include treatment of affected children by bone marrow transplantation, prenatal diagnosis of the disease and termination of affected pregnancies, and eliminating the chance of a child with thalassaemia major being born at all, i.e. to prevent marriage between 2 gene carriers. Employing this last strategy has been approved for several years in the Islamic Republic of Iran.

Some health care professionals, however, worry that this approach might result in accumulation of the disease gene, which is normally removed from the population through the death of homozygous patients. Therefore, they believe that if we prevent the marriage of 2 gene carriers, frequencies of the thalassaemia gene and individuals heterozygous for the gene might increase over time, and ultimately rise to a level where most people are gene carriers. They believe that unless gene therapy is available for treatment of thalassaemia in the near future, preventing the marriage of 2 gene carriers is, in fact, merely postponing our current problem to the future, and might even worsen the situation in the long term.

This study was conducted to determine the effect of 4 different approaches to thalassaemia on the disease gene frequency.

Methods

Computer simulation has helped science in general, and medicine in particular, in numerous ways. In this study, using a computer simulation, we investigated the effect on the population disease gene frequency and the frequencies of heterozygous and homozygous individuals in 4 different hypothetical approaches to thalassaemia. The 4 strategies were:

- to do nothing for thalassaemia patients but supportive therapy (e.g. regular blood transfusion and iron chelation therapy);
- to treat all homozygous patients using a hypothetical therapeutic modality that treats patients only phenotypically but not genetically (although people receiving this treatment would need no special medical care related to thalassaemia during their lifetimes, they still carry the disease gene and will pass it to their offspring);
- to carry out prenatal diagnosis of the disease and to abort all homozygotes;
- to prevent marriage between 2 gene carriers.

Simulation algorithm

To run a simulation, it is important to develop a computer model based on a real life problem. Different simulation techniques are available. Monte Carlo simulation is a type of stochastic simulation that incorporates random variability into the model. Monte Carlo simulation differs from traditional simulation in that the model parameters are treated as stochastic or random variables, rather than as fixed values. Monte Carlo simulation repeatedly simulates the model, each time drawing a different random set of input values to determine the resultant set of possible outcomes [4].

A computer simulation program was developed based on the Monte Carlo method. The pseudocode of the program is shown in Table 1. In this simulation, the elements of a 1 000 000-element 1-dimensional array of integer numbers were corresponded to each person of a 1 000 000-person population (500 000 males and 500 000 females). The number assigned to each array element reflected the genotype of the corresponding person (0 = normal, 1 = heterozygote, 2 = homozygote). Assuming a binomial distribution, with a disease gene frequency (f_{gene}) of p, the frequencies of homozygous (f_{homo}), and heterozygous (f_{hetero}) individuals in the population were calculated to be p^2 and $2 \times p \times (1-p)$ respectively. Therefore, with an f_{gene} of 5% (the β-thalassaemia f_{gene} in Fars province, southern Iran [2]) the f_{hetero}, and f_{homo} were calculated as 9.5%, and 0.25%, respectively. Based on these frequencies, the numbers of heterozygous and homozygous persons in our 1 000 000-individual imaginary population were 95 000 and 2500, respectively. Using these frequencies, in the first step of the simulation the parent population was categorized into heterozygous, homozygous and normal individuals (initialization step) (Figure 1).

Normally, many of those homozygous for thalassaemia would die of the disease complications before having offspring, however, a proportion of them, say m (fertility scalar), who had either a mild form of

Figure 1 **Pseudo-code of the simulation program**

Initialize parent population

loop for each generation from 0 to 20

   if (prevent marriage of gene carriers)

        Shuffle parent population so that carriers only occupy even positions of the array; place normal individuals elsewhere

   else

        Shuffle parent population array

   endif

        After random mating of parents (Mendelian inheritance), generate the offspring population

        Calculate the gene frequency and the percentage of heterozygotes or homozygous individuals in offspring's population (to be the next parent population)

        Reinitialize the new parent population

        Record the calculated frequencies

   endloop

the disease (thalassaemia intermedia) or who had received better care, would live long enough to pass their genes on. With an f_{homo} of p^2 at birth, the frequency of those homozygous for thalassaemia who could have offspring was therefore mp^2. This fertility scalar was considered in our simulation program.

In step 2, to produce the same chance for each individual to marry another, using a pseudo-random generator algorithm [5], the parent population array elements were shuffled at random. In this parent population, those corresponding to even positions of the array (0, 2, 4, etc.) mated with those corresponding to their next element (positions 1, 3, 5, etc.) in the array to produce 5 children. In determining the genotypes of their children, Mendelian inheritance was assumed. Based on the genotypes of the offspring, f_{gene}, f_{hetero}, and f_{homo} were calculated. Based on these newly derived frequencies, our 1 000 000-individual hypothetical parent population was reinitialized. The process was repeated for 20 generations and f_{gene}, f_{hetero}, and f_{homo} of the offspring population were recorded for each generation (Figure 1).

If the second approach, i.e. treating all homozygous patients phenotypically only with an imaginary therapeutic modality, the fertility scalar m was set to 1. This meant that all homozygous people survived to pass their genes to their offspring.

If the third approach was employed, when a homozygous child was generated, the program ignored it (prenatal diagnosis and abortion) and continued to generate another child until the number of non-homozygous children was 5.

If marriage of 2 gene carriers was forbidden, the parent population was initialized in such a way that after shuffling of the population, gene carriers could only occupy the parent population array elements with even positions (0, 2, 4, etc.); normal individuals could occupy other positions (odd and empty even positions). As before, those corresponding to even number positions of the array mated with those corresponding to their next odd number positions. In this way, gene carriers were forced to mate with normal individuals.

To eliminate the chaotic effects caused by inherent randomness of the Monte Carlo method [4], the arithmetic mean of the results obtained from execution of the program 100 times for each strategy was taken as the final, refined result.

To determine the effect of mean number of children for each couple (i.e. population growth rate), the processes were repeated for 3 children. The results were compared with those obtained for 5 children.

Results

If we do nothing for either thalassaemia prevention or treatment, f_{gene}, f_{hetero} and f_{homo} fall over time (Figure 2).

On the other hand, if all homozygous patients are treated phenotypically, not genetically, f_{gene} will be 5%, f_{hetero} 9.5% and f_{homo} 0.25%. These frequencies remain constant over time (Figure 2).

If prenatal diagnosis of thalassaemia plus abortion for all homozygous children is carried out, f_{gene} and f_{hetero} decrease with time (Figures 2A and B).

When marriage between 2 gene carriers is prevented, f_{gene}, after a drop to 4.79%, and f_{hetero}, after a rise to 9.58%, remain constant after the first generation (Figures 2A and B).

No homozygous individual is born if either of the 2 latter approaches is employed (Figure 2C).

Mean number of children (population growth rate) had no effect on the results.

Discussion

The method used to approach thalassaemia influences f_{gene}, f_{hetero}, and f_{homo} substantially. Considering our findings, it seems that under none of the studied conditions would f_{gene} rise, and in the worse case it would remain constant over generations.

Although f_{gene}, f_{hetero}, and f_{homo} decline with time, when we do nothing in the way of either prevention or treatment of
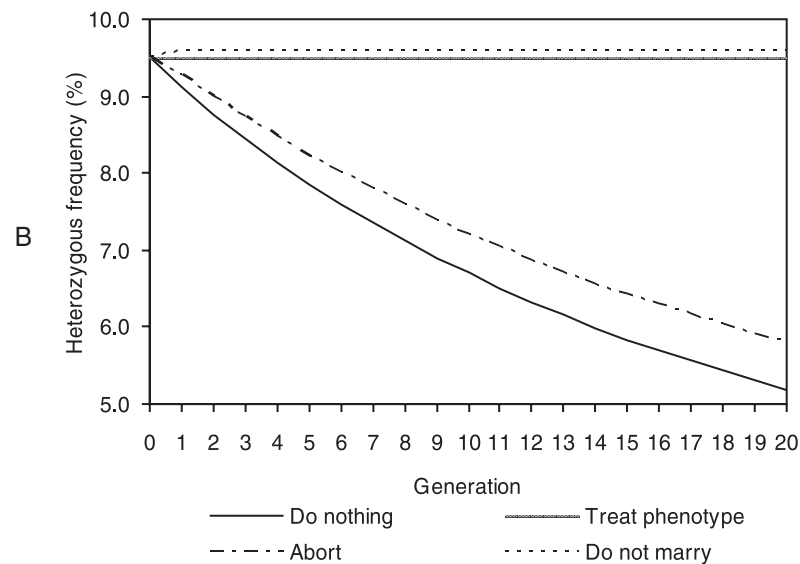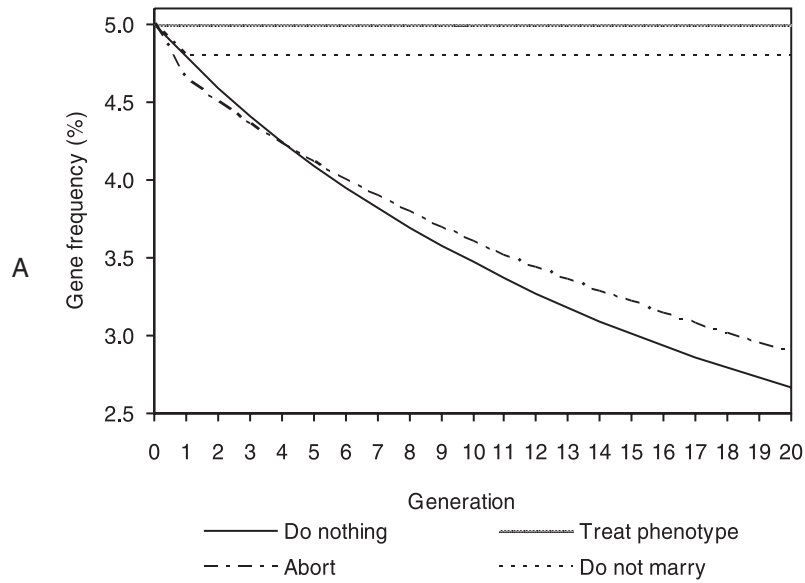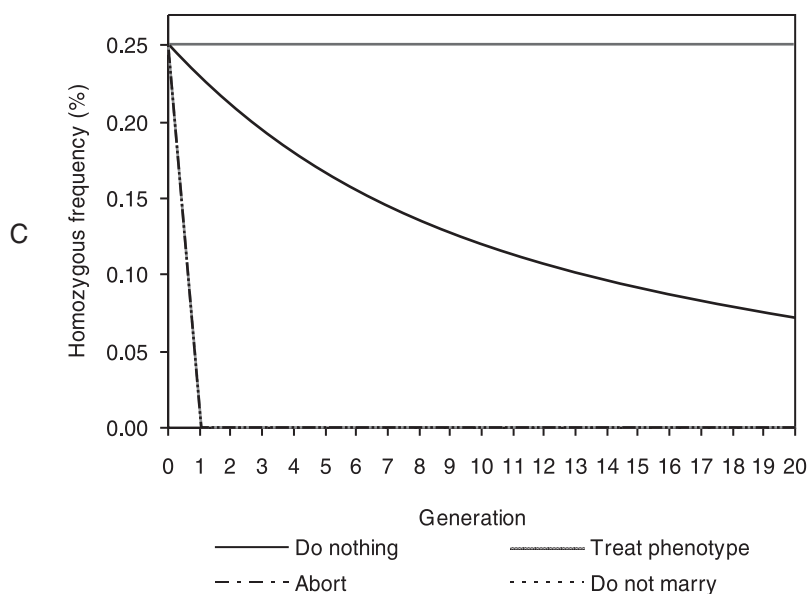
Figure 2 **Variation of A gene, B heterozygote and C homozygote frequencies with time (generation) in four different hypothetical approaches to thalassaemia patients (mean values from 100 executions of the simulation program)**

thalassaemia, the process is very slow and takes place over centuries. About 15% of homozygous patients develop thalassaemia intermedia ($m = 0.15$) [6]. Although they may experience conditions such as extramedullary haematopoiesis which need special medical attention, these non-transfusion-dependant patients, in general, have normal life expectancy, marry and have children. The remaining 85% of homozygotes, however, usually do not become sexually active and die of the disease complications (e.g. haemochromatosis) in their second or third decades of life [6]. Although better care of homozygous patients (e.g. better blood transfusion and iron chelation therapy protocols) increases their life expectancy, the outcome for patients is still grave, particularly in developing countries, where the disease is

prevalent and resources are scarce. This is why in our simulation we assumed that only 15% (i.e. $m = 0.15$) of homozygotes would have the chance of producing children if no treatment was instituted.

If we treat all homozygous patients phenotypically, f_{gene}, f_{hetero}, and f_{homo} become steady and do not change over time. This steady state is, in fact, the well-known Hardy–Weinberg equilibrium state [7].

Control of β-thalassaemia by carrier screening, genetic counselling and prenatal diagnosis has been successful in many countries [8–13]. Obviously, no one homozygous for the disease will be born after institution of prenatal diagnosis and termination of all affected pregnancies.

Although no homozygous individual is born if marriage of 2 gene carriers is prevented, f_{gene} (after a slight drop) and f_{hetero}

(after a slight rise) become constant from the first generation offspring onward. The slight drop in f_{gene} is due to the fact that treatment is started from the first generation onward; about 85% of the homozygous offspring of the parent population die of the disease complications and do not pass the disease gene to their (2nd generation) offspring. These genes are, in effect, removed from the population. For the same reason, a slight rise in f_{hetero} is observed. The remaining 15% of the homozygotes who marry (because of the restriction on marriage of 2 gene carriers) had to marry a normal individual. This would eventually result in the birth of heterozygous offspring. This mechanism at first produces a slight increase in f_{hetero}, but with the absence of homozygote individuals after the first generation, f_{hetero} becomes constant thereafter.

In reality, thalassaemia, with more than 150 different mutations, is too complex to be simulated accurately, and the relatively unsophisticated model presented in this study suffers from several limitations. The first approach that we explored, doing nothing, i.e. "no treatment, no prevention", is certainly not a practical method to be discussed for management of thalassaemia. It is, nevertheless, a situation occurring in certain areas of the world which, it is to be hoped, will disappear in the near future. Treatment of all homozygotes phenotypically is still not available. Bone marrow transplantation, though promising, is associated with several complications such as infertility, graft rejection, graft versus host disease and even death in a noticeable proportion of patients. Besides, transplantation for all patients affected with thalassaemia major is obviously an impossible target. Abortion of all affected fetuses is also not practical.

In this study, the fertility scalar m was only applied to homozygotes. However, it is unlikely that heterozygotes or even normal individuals have $m = 1$. The simulation also did not take into consideration the fact that members of defined ethnic or racial groups are more likely to marry within the group than outside it, and therefore it could only be applied to a small, genetically homogenous population.

The simulation model presented only tests 4 hypothetical situations. In real life, however, we face a very complex situation that has resulted from a combination of all the different approaches to thalassaemia. One way to understand the real situation would be to explore it at its extremes, i.e. to analyse the situation when only 1 approach is used. This over-simplification in fact delineates the borders outlining the true situation and thus, will help in better understanding of the current situation, in improving short- and long-term planning and in more efficient utilization of our health care resources.

As an example, if we try to perform bone marrow transplantation (to treat phenotypically) for some homozygotes and do prenatal diagnosis and abortion of affected fetuses for some pregnancies, we expect the resultant f_{gene} curve would lie somewhere between the curves corresponding to the 2 methods used.

Since we did not use anything specific to the thalassaemia gene, the inferences of this simulation may apply to many other similar hereditary diseases. No matter which of the 4 strategies is used, with the assumptions made, the population disease gene frequency, in the worst case, will remain constant over time.

References

1. Lukens JN. The thalassaemia and related disorders, quantitative disorders of haemoglobin synthesis. In: Lee GR et al., eds. *Wintrobe's clinical haematology*, 9th ed. Philadelphia, Lea & Febiger, 1993: 1102–45.

2. Habibzadeh F et al. Thalassaemia in Iran: an overview. *Archives of Iranian medicine*, 1998, 1(1):27–33.

3. Ghani R, Manji MA, Ahmed N. Hemoglobinopathies among five major ethnic groups in Karachi, Pakistan. *Southeast Asian journal of tropical medicine and public health*, 2002, 33(4):855–61.

4. Bonate PL. A brief introduction to Monte Carlo simulation. *Clinical pharmacokinetics*, 2001, 40(1):15–22.

5. Matsumoto M, Nishimura T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM transactions on modeling and computer simulation*, 1998, 8(1): 3–30.

6. Pearson HA et al. The changing profile of homozygous beta-thalassaemia: demography, ethnicity, and age distribution of current North American patients and changes in two decades. *Pediatrics*, 1996, 97(3):352–6.

7. Thompson JS, Thompson MW. *Genetics in medicine*, 3rd ed. Philadelphia, WB Saunders, 1980:290–3.

8. Cao A, Rosatelli MC, Galanello R. Control of beta-thalassaemia by carrier screening, genetic counselling and prenatal diagnosis: the Sardinian experience. *Ciba Foundation symposium*, 1996, 197: 137–51.

9. Ghanei M et al. Pre-marriage prevention of thalassaemia: report of a 100 000 case experience in Isfahan. *Public health*, 1997, 111(3):153–6.

10. Keskin A et al. Premarital screening of beta-thalassaemia trait in the province of Denizli, Turkey. *Acta haematologica*, 2000, 104(1):31–3.

11. Perera DM et al. Thalassaemia control in developing countries—the Sri Lankan perspective. *Ceylon medical journal*, 2000, 45(1):12–6.

12. Oliva Berini E, Cladera Serra A, Torrent Quetglas M. Campaña para la detección de la betatalasemia minor y prevención de la major en la isla de Menorca. Experiencia de 10 anos [Campaign for the detection of minor beta-thalassaemia and prevention of major beta-thalassaemia in the isle of Menorca. 10-year experience]. *Medicina clínica*, 1998, 110(10):361–4.

13. Wanapirak C et al. Prenatal strategies for reducing severe thalassaemia in pregnancy. *International journal of gynaecology and obstetrics*, 1998, 60(3):239–44.