A discrete time population genetic model for X-linked recessive diseases

Carmen Del Vecchio[1], Francesca Verrilli[1], Luigi Glielmo[1] and Martin Corless[2]

***Abstract*—**The epidemiology of X-linked recessive diseases, a class of genetic disorders, is modeled with a discrete-time, structured, mathematical model. The model accounts for both *de novo* mutations and different reproduction rates of procreating couples depending on their health conditions. Relying on Lyapunov theory, asymptotic stability properties of equilibrium points of the model are demonstrated. The model describes the spread over time in the population of any recessive genetic disorder transmitted through the X-chromosome.

***Keywords*—**Population genetic dynamics, nonlinear dynamic analysis, stability, epidemiology.

I. Motivation

We study a specific class of genetic disorders named X-linked recessive diseases; these conditions include the serious diseases hemophilia A, Duchenne-Becker muscular dystrophy, and Lesch-Nyhan syndrome as well as common and less serious conditions such as male pattern baldness and red-green color blindness. A major reason for devoting attention to this topic is the inadequacy of the currently used mathematical instruments to describe the transmission of a genetic disease within a predefined population.

Related studies analyzed the inheritance mechanism of any gene —not necessarily responsible for a genetic disease— placed on the X-chromosome ([1], [2], [3]); they belong to the field of population genetics. In these works genotypes frequencies —i.e., the frequency or proportion of genotypes in a population— are frequently chosen as model's variables. Under the hypothesis of infinite population and starting from a genotypes' distribution, the genotypes' proportions in the next generation are evaluated according to the inheritance mechanism and to the effects of selection or mutation. The average fitness (see [4] pag. 385-387) is frequently studied as a suitable Lyapunov function candidate to analyze stability properties of model's equilibrium points.

Even in this generic scenario seldom contributions examine the combined effects of selection and mutation on population's dynamics and equilibrium (see [5], [6], [7] and reference therein). Moreover, results of these researches cannot be applied to epidemiological studies. In fact the ultimate goal of genetic epidemiology is to predict the number of individuals carrying the disease responsible gene, this number can not be inferred from genotypes' frequency distribution when assuming infinite population size.

The results in this paper are an extension of some previous preliminary works on the same topic ([8], [9]). The major original result of our work is to model the peculiar inheritance mechanism of X-linked diseases while taking into account the reduced reproduction capacity of affected individuals as well as the occurrence of the diseases in healthy couple progeny due to *de novo* genetic mutations. This represents an advancement over current epidemiological models, and could be exploited to better understand the epidemiology of X-linked genetic diseases; moreover, it could be generalized to allow application to other classes of genetic disorders. Although the mathematical model developed to describe the epidemiology of these diseases within a population is nonlinear, it is suitable for analyses using classical nonlinear instruments to gain information about system behavior and equilibrium properties.

The paper is structured as follows: a brief description of X-linked genetic diseases and their peculiar inheritance pattern is provided in Section II-A. We pose our model and we derive general system properties and solutions characteristic in Sections II-B and III.Finally, we discuss the physiological implications corresponding to the mathematical properties derived from our model and some special model cases.

II. Mathematical model of X-linked recessive diseases

A. The transmission mechanism of X-linked recessive diseases

An X-linked recessive disease may be inherited as per the following rules (see [9] for details):

- Affected males never spread the disease to their sons, as no male-to-male transmission of the X chromosome occurs.
- Affected males pass the abnormal X chromosome to all of their daughters, who are described as obligate carriers.
- On average, female carriers pass a defective X chromosome to half of their sons (who will born affected) and half of their daughters (who become carriers). The remaining half of their siblings inherit a normal copy of the chromosome.
- Affected females are the rare result of an affected male and a carrier female mating.

Figures 1 and 2 depict the inheritance patterns described above.

X-linked recessive conditions include the serious diseases Duchenne/Becker muscular dystrophy, hemophilia A, and

[1] Dipartimento di Ingegneria, Università degli Studi del Sannio, Benevento (Italy) {fverrilli,glielmo,c.delvecchio}@unisannio.it
[2] School of Aeronautics and Astronautics, Purdue University, West Lafayette, Indiana (USA) corless@purdue.edu.

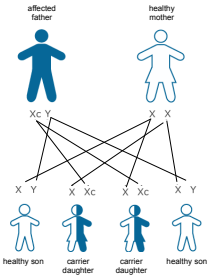


Fig. 1. Inheritance pattern of affected father and healthy mother.

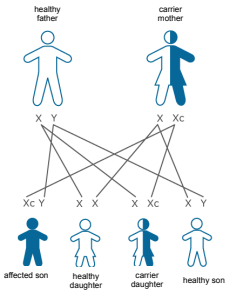


Fig. 2. Inheritance pattern of healthy father and carrier mother.

Lesch-Nyhan syndrome as well as common and less serious conditions such as male pattern baldness and red-green color blindness. X-linked dominant diseases are very uncommon, although some inherited forms of rickets are transmitted in this manner. Unlike the recessive diseases discussed above, the prevalence of X-linked dominant diseases is similar in males and females, even though the absence of male-to-male transmission distinguishes them from autosomal dominant diseases.

Other than the result of the described, well characterized patterns of inheritance, the spread of genetic disorders within a given population is influenced by additional factors, including sporadic mutations and prenatal diagnosis.

A genetic disease that occurs when neither parent is affected or a carrier of any genetic defect is called *sporadic mutation* or *de novo gene mutation*. These cases arise via random genetic mutations within the DNA sequence; the mutation can occur in the the germ-line cell population — i.e. in eggs and sperm cells— in subjects without any prior genetic defect and can be transmitted down to one of the offspring. The genetic mutation can also occur in the zygote cell, i.e. the initial cell formed when two gametes cell are joined. A sporadic mutation can be the cause of an X-linked recessive disease (whereas it is unlikely for an autosomal recessive disorder) as a single mutation is enough in males to cause the disease. Males can be born affected due to a spontaneous gene mutation as a single abnormal gene copy is enough for the disease to become symptomatic; females can also be born carriers owing to random mutations.

The rate of *de novo* mutations varies widely among different genetic regions, and depends on a number of factors, including environmental exposure to mutagenic agents, the length of the gene sequence and ability of the cell machinery to actually repair or correct the mutations. For instance, the dystrophin gene, whose mutations may give rise to Duchenne/Becker muscular dystrophy, is particularly prone to *de novo* mutations due to its massive length ([10]). It is estimated that up to a third of all cases of this disorder are due to *de novo* mutations, a rate considerably higher than any other X-linked disorder. Not all mutations in X-chromosome genes confer sufficient survival fitness to give rise to a viable embryo. Therefore, an individual born affected due to *de novo* mutations may not pass on the affected gene to progeny due to premature, spontaneous intrauterine death.

In modern population genetics, prenatal diagnosis and birth control measures can play a major role and significantly influence the rate of affected cases in a given generation. Couples with a family history of X-linked genetic disorders often access prenatal or pre-implantation embryonic genetic screening, with the consequent negative selection of affected ones or selective therapeutic abortion.

B. Model formulation

In this section we present a mathematical model we have developed to describe the epidemiology of genetic diseases linked to the X chromosome. Our model fits in the category of structured models. In these models, a population is divided into homogeneous groups according to some major parameters, such as subject's age, sex or health conditions with respect to a disease. The model dynamics describes the distribution of the population over time according to the chosen parameters ([11], [12]).

Population dynamic models differ depending on assumptions regarding population size and mating rules among groups. The population is often assumed to be isolated (i.e. migration and selection are not modeled) and of constant finite size. This allows mathematical tractability. However, models with variable population size or including selection factors (such as the early death of affected individuals, or selection due to prenatal diagnosis) are definitely more realistic.

The most frequently adopted rule for mating is that individuals in the population mix randomly, i.e. individuals mate according to the product rule of probability; this is more realistic in large populations and assumes that the studied trait does not influence reproduction ([13]).

We have developed a discrete-time dynamic model to describe the inheritance mechanism of X-linked recessive diseases in a finite size population grouped by sex and health condition with respect to the disease.

We divide the population into four classes, namely healthy and affected males, healthy and carrier females. We do not consider affected females as they very rarely occur in nature. Thus our model has four variables:

- $x_1(k)$: the number of healthy males at time k
- $x_2(k)$: the number of affected males at time k
- $x_3(k)$: the number of healthy females at time k
- $x_4(k)$: the number of carrier females at time k.

We make the following assumptions.

- In each generation there is an equal number of males and females; thus there is the the same number of males and females in newborn children.
- Each person breeds with a person of the opposite sex from his/her own generation.
- The number of sons (which is equal the number of daughters) of each couple varies according to the parents' health conditions and is modeled through the *fertility factors* w_{ij}.
- Spontaneous genetic mutations are modeled; they are assumed to occur in the zygote cells; a child of a healthy couple can be born affected or a carrier due to mutation.

The number of males (and females in view of the first assumption above) born to a person of class $i \in \{1,2\}$ breeding with a person of class $j \in \{3,4\}$ is

$$\frac{1}{2} w_{ij} \frac{x_i}{(x_1+x_2)} \frac{x_j}{(x_3+x_4)} (x_1+x_2+x_3+x_4), \quad (1)$$

where

$$w_{ij} \geq 0 \quad (2)$$

is the fertility factor, procreation rate or reproduction rate of couples of type (i,j). The parameters w_{ij} are bounded by clinical considerations; the more severe the disease the smaller its value in couples formed by affected males and/or carrier females. A son (daughter) is a healthy or affected male (healthy or carrier female) depending on the health conditions of his (her) parents. As an example consider sons born from couples formed by a healthy father (an individual of class 1) and a carrier mother (an individual of class 4). According to the inheritance patterns of X-linked recessive diseases (see Figure 2), half of these sons (on average) will be affected and half will be healthy. Hence the number of affected males in the next generation due to such couples is

$$\frac{1}{4} w_{14} \frac{x_1}{(x_1+x_2)} \frac{x_4}{(x_3+x_4)} (x_1+x_2+x_3+x_4)$$

which is the same as the number of healthy males in the next generation due to these couples.

For ease in describing our model we introduce the state vector

$$x := [x_1 \; x_2 \; x_3 \; x_4]^T.$$

According to the previous assumptions and to the inheritance pattern of X-linked recessive disease described in Section II-A, the population dynamics is described by the following non-linear discrete-time system:

$$x(k+1) = f(x(k)) \quad (3)$$

where the vector function $f = [f_1 \; f_2 \; f_3 \; f_4]^T$ is given by

$$\begin{aligned}
f_1(x) &= P(x)\Big[(1-\alpha) w_{13} x_1 x_3 + w_{23} x_2 x_3 + \\
&\quad \frac{1}{2} w_{14} x_1 x_4 + \frac{1}{2} w_{24} x_2 x_4\Big] \quad (4a)\\
f_2(x) &= P(x)\Big[\alpha w_{13} x_1 x_3 + \frac{1}{2} w_{14} x_1 x_4 + \\
&\quad \frac{1}{2} w_{24} x_2 x_4\Big] \quad (4b)\\
f_3(x) &= P(x)\Big[(1-\beta) w_{13} x_1 x_3 + \frac{1}{2} w_{14} x_1 x_4\Big] \quad (4c)\\
f_4(x) &= P(x)\Big[\beta w_{13} x_1 x_3 + w_{23} x_2 x_3 + \frac{1}{2} w_{14} x_1 x_4 + \\
&\quad w_{24} x_2 x_4\Big] \quad (4d)
\end{aligned}$$

with

$$P(x) := \frac{x_1+x_2+x_3+x_4}{2(x_1+x_2)(x_3+x_4)}. \quad (5)$$

The term αw_{13} (βw_{13}) is the fraction of affected sons (carrier daughters) born from healthy parents due to *de novo* gene mutation; thus α and β model the spontaneous mutation rate of the disease in males and females respectively. Their values strictly depend on the genetic disease and they range between 10^{-3} and 10^{-8} ([14]). We will consider α and β to be strictly less then $\frac{1}{2}$; otherwise spontaneous genetic mutation would be more relevant than the ordinary disease transmission mechanism; in contrast setting $\beta = 0$ and $\alpha = 0$ implies that gene mutations do not apply to the disease. Thus in what follows α and β will range in $[0, \frac{1}{2})$.

Finally we explicitly note that once system (3)-(4) is initialized with all state variables non-negative (that is, non-negative initial populations) —the state variables remain nonnegative for all times $k \geq 0$:

$$x_i(k) \geq 0 \quad \text{for} \quad i = 1,2,3,4 \quad \text{when} \quad x_{i0} := x_i(0) \geq 0 \quad (6)$$

for $i = 1, \ldots, 4$. Hence we are dealing with a *positive system*.

Due to the model hypotheses discussed above —each couple procreates an equal number of males and females— one can easily see that

$$x_1(k) + x_2(k) = x_3(k) + x_4(k) \quad (7)$$

for all $k > 0$. Hence $P(x)$ in equation (5) simplifies to

$$P(x) = \frac{1}{x_1+x_2} = \frac{1}{x_3+x_4}$$

that is, $P(x)$ is the inverse of half the total population. Since $w_{ij} \geq 0$ for $i = 1,2$ and $j = 3,4$ and assuming $w_{ij} > 0$ for at least one couple (i,j), it is not difficult to show that

$$x_1(k) + x_2(k) > 0 \qquad \text{for all } k \text{ when} \qquad x_{10} + x_{20} > 0\,. \quad (8)$$

This guarantees that $P(x(k))$ is always well defined. Let

$$\mathcal{X} := \{x : x_i \geq 0 \text{ for } i = 1, \ldots, 4 \text{ and } x_1 + x_2 = x_3 + x_4 > 0\}.$$

Then this set is invariant for system (3)-(4), that is, if the state starts in this set, it never leaves it.

The model presented in this paper is a generalization of previous models in ([8], [9]). It significantly improves the modeling of *de novo* mutations (i.e. affected or carrier children born to healthy parents) and reproduction rates consistent with the health conditions of reproducing couples. These model features, enabling the modeling of any X-linked disease, were not included in the first version ([8]). Moreover the results we gain on system equilibrium, stability and convergence properties are more general than those in [9] where only the special case of negligible *de novo* mutations and few combinations of reproduction rates values could be analyzed. Finally in this paper the sporadic genetic mutation can affect sons and daughters born from healthy couples with different rates α and β; in [9] it was assumed that only males could be born affected due to a spontaneous mutation.

III. Some System Properties

A. A Lyapunov function

Noting that

$$\begin{aligned}
x_1 &= P(x)(x_1x_3 + x_1x_4), \qquad x_2 &&= P(x)(x_2x_3 + x_2x_4)\\
x_3 &= P(x)(x_1x_3 + x_2x_3), \qquad x_4 &&= P(x)(x_1x_4 + x_2x_4)
\end{aligned} \quad (9)$$

system (3)-(4) can be described by

$$x(k+1) = x(k) + g(x(k)) \tag{10}$$

where $g = [g_1\ g_2\ g_3\ g_4]^T$ and

$$\begin{aligned}
g_1(x) &= P(x)\Big[((1-\alpha)w_{13}-1)x_1x_3 + \\
&\quad \Big(\frac{w_{14}}{2}-1\Big)x_1x_4 + w_{23}x_2x_3 + \frac{w_{24}}{2}x_2x_4\Big] \\
g_2(x) &= P(x)\Big[\alpha w_{13}x_1x_3 + \frac{1}{2}w_{14}x_1x_4 - x_2x_3 + \\
&\quad \Big(\frac{w_{24}}{2}-1\Big)x_2x_4\Big] \\
g_3(x) &= P(x)\Big[((1-\beta)w_{13}-1)x_1x_3 + \\
&\quad \frac{1}{2}w_{14}x_1x_4 - x_2x_3\Big] \\
g_4(x) &= P(x)\Big[\beta w_{13}x_1x_3 + \Big(\frac{w_{14}}{2}-1\Big)x_1x_4 + \\
&\quad w_{23}x_2x_3 + (w_{24}-1)x_2x_4\Big].
\end{aligned}$$

In investigating the behavior of system (3)-(4) the function V_1 defined by

$$V_1(x) = x_1 + x_2 \tag{12}$$

is very useful. This is simply the total male population which is the same as the total female population, that is,

$$V_1(x) = x_3 + x_4. \tag{13}$$

This function will be called a Lyapunov function. The change in this population from one stage k to the next stage $k+1$ is given by

$$V_1(x(k+1)) - V_1(x(k)) = \Delta V_1(x(k)) \tag{14}$$

where

$$\begin{aligned}
\Delta V_1(x) &:= V_1(f(x)) - V_1(x) \\
&= g_1(x) + g_2(x) \\
&= -P(x)[(1-w_{13})x_1x_3 + (1-w_{14})x_1x_4 + \\
&+ (1-w_{23})x_2x_3 + (1-w_{24})x_2x_4].
\end{aligned} \tag{15}$$

If $w_{ij} \leq 1$ for all i,j then, for all $k \geq 0$, we have $\Delta V_1(x(k)) \leq 0$ and

$$V_1(x(k+1)) \leq V_1(x(k))\,. \tag{16}$$

Therefore

$$V_1(x(k)) \leq V_1(0) \tag{17}$$

for all $k \geq 0$. That is the total population is bounded by the initial population. We can now readily obtain the following boundedness result. Noting that

$$x_i(k) \leqslant V_1(k) \qquad \text{for } i = 1, \ldots, 4.$$

we obtain our first result.

Proposition 3.1: Consider system (3)-(4) with $0 \leq w_{ij} \leq 1$ for $i = 1, 2$ and $j = 3, 4$. If the initial state x_0 lies in $\mathcal{X}$, then, for all $k \geq 0$, $x(k) \in \mathcal{X}$ and

$$x_1(k) + x_2(k) \leqslant x_{10} + x_{20} \tag{18}$$

$$x_3(k) + x_4(k) \leqslant x_{30} + x_{40} \tag{19}$$

$$x_i(k) \leqslant x_{10} + x_{20} \quad \text{for } i = 1, \ldots, 4. \tag{20}$$

a) **Special case: $w_{ij} = 1$ for all i and j:** In this case, $\Delta V_1(x) = 0$ for all x; hence $V(x(k+1) = V(x(k))$ for all k which implies that $V_1(x(k)) = V_1(x(0))$, that is,

$$x_1(k) + x_2(k) = x_{10} + x_{20} \tag{21}$$

for all k. This means that the total male (hence female) population remains constant. We examine this special case in further detail in Section IV Now we consider what happens when $w_{ij} < 1$ for at least one ij.

B. Some convergence properties

Our first result tells us that if one of the fertility rates w_{ij} is strictly less than one, then the state converges to a limit with the number of the affected males and carrier males equal to zero. If in addition $w_{13} < 1$ or there is a non-zero mutation rate then the whole population goes to zero.

Proposition 3.2: Consider system (3)-(4) with initial stat x_0 in $\mathcal{X}$ and $0 < w_{ij} \leq 1$ for $i = 1, 2$ and $j = 3, 4$.

If $w_{ij} < 1$ for at least one ij, then

$$\lim_{k\to\infty} x_2(k) = \lim_{k\to\infty} x_4(k) = 0 \tag{22}$$

and

$$\lim_{k\to\infty} x_1(k) = \lim_{k\to\infty} x_3(k) = \underline{x}_1 \qquad \text{for some } \underline{x}_1 \geq 0. \tag{23}$$

If either $w_{13} < 1$ or $\alpha > 0$ or $\beta > 0$ then $\underline{x}_1 = 0$, that is,

$$\lim_{k\to\infty} x(k) = 0.$$

Proof: Since $w_{ij} \leq 1$ for all i and j, it follows that (16) holds, that is, $\{V_1(x(k))\}$ is a non-increasing sequence. Since this sequence is bounded below by zero, it converges that is

$$\lim_{k\to\infty} V_1(k) = \underline{V_1} \tag{24}$$

for some $\underline{V_1} \geq 0$. Hence

$$\begin{aligned}
\lim_{k\to\infty} \Delta V_1(x(k)) &= \lim_{k\to\infty} V_1(x(k+1)) - \lim_{k\to\infty} V_1(x(k)) \\
&= \underline{V_1} - \underline{V_1} \\
&= 0\,.
\end{aligned} \tag{25}$$

Suppose that $w_{ij} < 1$ for some ij. Since $w_{ij} \leq 1$ for all ij it follows from (15) that

$$\Delta V_1(x(k)) \leqslant -P(x(k))(1-w_{ij})x_i(k)x_j(k) \leq 0$$

Since $\Delta V_1(x(k)) \to 0$, it now follows that $P(x(k))x_i(k)x_j(k)$ tends to 0. Noting that $P(x(k)) \geq P(x(0))$ we obtain that $x_i(k)x_j(k)$ goes to zero. This implies that $f_i(x(k))f_j(x(k)) = x_i(k+1)x_j(k+1)$ also converges to zero.

First suppose that $w_{14} < 1$. Then

$$P(x(k))x_1(k)x_4(k) \to 0\,. \tag{26}$$

Also $f_1(x(k))f_4(x(k)) \to 0$. Since all terms in (4a) and (4d) are non-negative we must have

$$P(x(k))x_2(k)x_3(k) \to 0 \quad \text{and} \quad P(x(k))x_2(k)x_4(k) \to 0\,. \tag{27}$$

Using the second relationship in (9) it follows from (27) that $x_2(k) \to 0$ and, recalling (24), we also have $x_1(k) \to \underline{x}_1 := \underline{V_1}$. The fourth relationship in (9) along with (26) and the second condition in (27) imply that $x_4(k) \to 0$ and, recalling (24), we also have $x_3(k) \to \underline{x}_1$.

Now suppose that $\beta > 0$. Since $f_1(x(k))f_4(x(k)) \to 0$ equations (4a) and (4d) imply that

$$P(x(k))x_1(k)x_3(k) \to 0. \tag{28}$$

When $\alpha > 0$, the fact that $f_2(x(k)) = x_2(k+1) \to 0$ and (4b) also implies (28). Using (28) and (26) along with the first relationship in (9) we obtain that $x_1(k) \to 0$, that is, $\underline{x}_1 = 0$.

Now suppose that $w_{24} < 1$. Then $f_2(x(k))f_4(x(k)) \to 0$ and it follows from (4b) and (4d) that (26) holds and as we have just shown this results in (22) and (23) with $\underline{x}_1 = 0$ when either $\alpha > 0$ or $\beta > 0$.

In a similar fashion one can show that if $w_{23} < 1$ then, (22) and (23) hold with $\underline{x}_1 = 0$ when either $\alpha > 0$ or $\beta > 0$.

Finally suppose $w_{13} < 1$. Then $f_1(x(k))f_3(x(k)) \to 0$ and it follows from (4a) and (4c) that (26) and (28) hold. From this we can conclude as before that (22) and (23) hold with $\underline{x}_1 = 0$ ■

The next result tells us that even if $w_{ij} = 1$ for all ij, properties (22) and (23) still hold. To prove this we need to introduce a new function V_2.

Proposition 3.3: Consider system (3)-(4) with initial state x_0 in $\mathcal{X}$ and $0 < w_{ij} \leq 1$ for $i = 1, 2$ and $j = 3, 4$. If $\alpha = \beta = 0$ then (22) and (23) hold.

Proof: To prove this result we consider the behavior of the following function:

$$V_2(x) := x_2 + x_4 \tag{29}$$

Along any solution $x(\cdot)$ we have

$$V_2(x(k+1)) = V_2(x(k)) + \Delta V_2(x(k)) \tag{30}$$

and with $\alpha = \beta = 0$ we have

$$\begin{aligned}
\Delta V_2(x) &= g_2(x) + g_4(x) \\
&= -P(x)\Big[(1-w_{14})x_1x_4 + \\
&\quad (1-w_{23})x_2x_3 + (2 - \frac{3w_{24}}{2})x_2x_4\Big] \\
&\leqslant -P(x)(2 - \frac{3w_{24}}{2})x_2x_4 \leqslant 0.
\end{aligned}$$

Proceeding as in the proof of Propositions 3.2, we can show that we must have

$$P(x(k))x_2(k)x_4(k) \to 0, \tag{31}$$

and from this we can deduce that

$$f_2(x(k))f_4(x(k)) \to 0. \tag{32}$$

It follows from (32) that

$$P(x(k))x_1(k)x_4(k) \to 0. \tag{33}$$

Using the fourth relationship in (9) it follows from (31) and (33) that $x_4(k) \to 0$. It now follows that $f_4(x(k)) = x_4(k+1)$ goes to zero and, hence

$$P(x(k))x_2(k)x_3(k) \to 0 \tag{34}$$

Using the second relationship in (9) it follows from (31) and (34) that $x_2(k) \to 0$. The proof of (23) is the same as that in Proposition 3.2 ■

b) **The special case:** $w_{13} = 1$ **and** $\alpha, \beta = 0$**:** In this special $x(k)$ does not always converge to zero; only $x_2(k)$ and $x_4(k)$ always converge to zero. This can be seen by observing that, in this case, any state of the form $[x_1\ 0\ x_3\ 0]^T$ is an equilibrium state of system (3)-(4). If the system starts in one of these equilibrium states it remains there. Such a state corresponds to all the population being healthy.

When all w_{ij} are strictly less than one, the next result claims that the population exponentially decays to zero.

Proposition 3.4: Consider system (3)-(4) with $w_{ij} < 1$ for $i = 1, 2$ and $j = 3, 4$ and let

$$\bar{w} := \max\{w_{ij}\} < 1.$$

If the initial state x_0 lies in $\mathcal{X}$ then, for all $k \geq 0$,

$$x_1(k) + x_2(k) \leqslant \bar{w}^k(x_{10} + x_{20}) \tag{35}$$

$$x_3(k) + x_4(k) \leqslant \bar{w}^k(x_{30} + x_{40}) \tag{36}$$

$$x_i(k) \leqslant \bar{w}^k(x_{10} + x_{20}) \text{ for } i = 1, \ldots, 4 \tag{37}$$

Proof: It follows from (15) that

$$\begin{aligned}
\Delta V_1(x) &\leqslant -\frac{1}{V_1}(1-\bar{w})(x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4) \\
&= -\frac{(1-\bar{w})}{V_1}[(x_1+x_2)(x_3+x_4)] \\
&= (-1+\bar{w})V_1(x).
\end{aligned}$$

It now follows from (14) that $V_1(x(k+1)) \leqslant \bar{w}V_1(x(k))$ and, consequently, $V_1(x(k)) \leqslant \bar{w}^kV_1(x(0))$ which yields the desired results. ■

Remark 1: Note that the results in propositions 3.1 and 3.4 are independent of the mutation rates α and β. Hence, these results hold for any α and β in $[0, \frac{1}{2})$.

Now we derive system' solutions upper and lower bounds and provide the exact solutions in some specific cases. Comments on the medical implications of the mathematical results can be found in the Discussion Section.

C. Lower bound on convergence values of $\mathbf{x_1}$ and $\mathbf{x_3}$

To obtain estimates of the values to which x_1 and x_3 converge, consider

$$V_3(x) = x_1 + x_3 - x_2 - x_4.$$

Along any solution $x(\cdot)$ we have $V_3(x(k+1)) = V_3(x(k)) + \Delta V_3(x(k))$ where

$$\begin{aligned}\Delta V_3 &= g_1(x) + g_3(x) - g_2(x) - g_4(x)\\ &= P(x)(2 - w_{24})x_2x_4\\ &\geqslant 0.\end{aligned}$$

This guarantees that V_3 does not decrease. If $\underline{x}_1$ and $\underline{x}_3$ are the values to which x_1 and x_3 converge, then

$$\underline{x}_1 = \underline{x}_3 \geqslant \frac{1}{2}V_3(x_0) = \frac{1}{2}[x_{10} + x_{30} - x_{20} - x_{40}].$$

For the above bound to be positive, it must hold

$$x_{10} + x_{30} \geqslant x_{20} + x_{40},$$

that is in the initial population distribution the number of healthy people must be greater or equal to the number of affected and carrier one. This hypothesis is reasonable and consistent with the epidemiological observation of the diseases.

D. Upper bound on convergence values of $\mathbf{x_1}$ and $\mathbf{x_3}$

To obtain an upper bound on $\underline{x}_1$ and $\underline{x}_3$ consider

$$\begin{aligned}V_4(x) &= V_3(x) + \Big(\frac{4 - 2w_{24}}{4 - 3w_{24}}\Big)V_2(x)\\ &= x_1 + x_3 + \Big(\frac{w_{24}}{4 - 3w_{24}}\Big)(x_2 + x_4).\end{aligned}$$

Along any solution $x(\cdot)$ we have $V_4(x(k+1)) = V_4(x(k)) + \Delta V_4(x(k))$ where

$$\begin{aligned}\Delta V_4 &= \Delta V_3(x) + \Big(\frac{4 - 2w_{24}}{4 - 3w_{24}}\Big)\Delta V_2(x)\\ &= P(x)(2 - w_{24})x_2x_4 -\\ &\quad \Big(\frac{4 - 2w_{24}}{4 - 3w_{24}}\Big)P(x)[(1 - w_{14})x_1x_4 + (1 - w_{23})x_2x_3] -\\ &\quad \Big(\frac{4 - 2w_{24}}{4 - 3w_{24}}\Big)P(x)\Big(2 - \frac{3w_{24}}{2}\Big)x_2x_4\\ &= -\Big(\frac{4 - 2w_{24}}{4 - 3w_{24}}\Big)P(x)[(1 - w_{14})x_1x_4 +\\ &\quad (1 - w_{23})x_2x_3] \qquad (38)\\ &\leqslant 0\end{aligned}$$

This implies that V_4 does not increase and yields the upper bounds:

$$\underline{x}_1 = \underline{x}_3 \leqslant \frac{1}{2}\Big[x_{10} + x_{30} + \Big(\frac{w_{24}}{4 - 3w_{24}}\Big)(x_{20} + x_{40})\Big]$$

E. A special case

Consider the case in which $w_{13} = w_{14}$=w_{23}=1 that correspond to diseases where the fertility is altered only in couples formed by affected males and carrier females, that is only w_{24} ranges in $[0, 1]$. In this case from (38) we can deduce that V_4 is constant; hence

$$\underline{x}_1 = \underline{x}_3 = \frac{1}{2}\Big[x_{10} + x_{30} + \Big(\frac{w_{24}}{4 - 3w_{24}}\Big)(x_{20} + x_{40})\Big],$$

which is positive for non-zero initial conditions.

Another interesting case is when $w_{13} = 1$, w_{14} in [0,1] and $w_{23} = w_{24} = 0$. This models the reproduction scenario for severe X-linked recessive diseases such as the aforementioned hemophilia A and ectodermal dysplasia where healthy males and females always contribute to next generation, while healthy males and carrier women have a variable reproduction capacity (w_{14}) depending on the disease gravity. Affected males do not contribute to next generation, independently from the females health condition as they rarely reach the reproduction age.

For this combination of reproduction rates the exact solution of system (3)-(4) is:

$$\begin{aligned}x_1(k) &= A_1\frac{\prod_{i=1}^{k}\Big(2^i x_{30} + \sum_{j=1}^{i} 2^{(i-j)} w_{14}^j x_{40}\Big)}{\prod_{i=1}^{k-1}\Big(2^{i-1}x_{30} + \sum_{p=1}^{i}\Big(\prod_{j=1}^{i-p-1} 2\Big) w_{14}^p x_{40}\Big)}\\ x_2(k) &= A_2\frac{\prod_{i=1}^{k-1}\Big(2^i x_{30} + \sum_{j=1}^{i} 2^{i-j} w_{14}^j x_{40}\Big)}{\prod_{i=1}^{k-1}\Big(2^{i-1}x_{30} + \sum_{p=1}^{i}\Big(\prod_{j=1}^{i-p-1} 2\Big) w_{14}^p x_{40}\Big)}\\ x_3(k) &= A_1\frac{\prod_{i=1}^{k}\Big(2^i x_{30} + \sum_{j=1}^{i} 2^{(i-j)} w_{14}^j x_{40}\Big)}{\prod_{i=1}^{k-1}\Big(2^{i-1}x_{30} + \sum_{p=1}^{i}\Big(\prod_{j=1}^{i-p-1} 2\Big) w_{14}^p x_{40}\Big)}\\ x_4(k) &= A_2\frac{\prod_{i=1}^{k-1}\Big(2^i x_{30} + \sum_{j=1}^{i} 2^{i-j} w_{14}^j x_{40}\Big)}{\prod_{i=1}^{k-1}\Big(2^{i-1}x_{30} + \sum_{p=1}^{i}\Big(\prod_{j=1}^{i-p-1} 2\Big) w_{14}^p x_{40}\Big)}\end{aligned}$$

where

$$\begin{aligned}A_1 &= \frac{x_{10}(x_{10} + x_{20} + x_{30} + x_{40})}{2^{2k}(x_{10} + x_{20})(x_{30} + x_{40})}\\ A_2 &= \frac{w_{14}^k x_{10}x_{40}(x_{10} + x_{20} + x_{30} + x_{40})}{2^{2k}(x_{10} + x_{20})(x_{30} + x_{40})}.\end{aligned}$$

The demonstration is straightforward and can be easily obtained substituting the previous solution in system (3)-(4).

IV. Special case: unitary fertility factors

In this section we consider a special case of model (3)-(4). This case is obtained setting all reproduction rates equal to one ($w_{ij} \equiv 1$), thus modeling non-disabling diseases which do not affect the fertility of affected/carrier couples.

When $w_{ij} = 1$ for $i = 1, 2$ and $j = 3, 4$, the functions f_i

in (4) simplify to

$$\begin{aligned} f_1(x) &= P(x)\Big[(1-\alpha)x_1x_3 + x_2x_3 + \frac{1}{2}x_1x_4 + \\ &\quad \frac{1}{2}x_2x_4\Big] && \text{(40a)} \\ f_2(x) &= P(x)\Big[\alpha x_1x_3 + \frac{1}{2}x_1x_4 + \frac{1}{2}x_2x_4\Big] && \text{(40b)} \\ f_3(x) &= P(x)\Big[(1-\beta)x_1x_3 + \frac{1}{2}x_1x_4\Big] && \text{(40c)} \\ f_4(x) &= P(x)\Big[\beta x_1x_3 + x_2x_3 + \frac{1}{2}x_1x_4 + x_2x_4\Big] && \text{(40d)} \end{aligned}$$

It follows from Section III-A that $x_1(k+1) + x_2(k+1) = x_1(k) + x_2(k)$ for all k, that is, that the number of males, equal to the number of females at each generation is constant. Consider a population of $2N$ individuals; the following constraints apply:

$$\begin{aligned} x_1(k) + x_2(k) &= N \text{ for all } k, && \text{(41a)} \\ x_3(k) + x_4(k) &= N \text{ for all } k; && \text{(41b)} \end{aligned}$$

thus system dynamics can be rewritten using one state variable of the male population (i.e. x_1 or x_2) and one variable of the female population (i.e. x_3 or x_4); moreover, $P(x)$ as given (5) simplifies to $P = \frac{1}{N}$.

We normalize the system states by introducing $z = [z_1\ z_2\ z_3\ z_4]^T$ where

$$z_i := \frac{x_i}{N} \qquad \text{for} \qquad i = 1, \ldots, 4$$

The evolution of z is governed by

$$z(k+1) = h(z(k)) \tag{42}$$

where $h = [h_1\ h_2\ h_3\ h_4]^T$ with

$$\begin{aligned} h_1(z) &= (1-\alpha)z_1z_3 + z_2z_3 + \frac{1}{2}z_1z_4 + \frac{1}{2}z_2z_4 && \text{(43a)} \\ h_2(z) &= \alpha z_1z_3 + \frac{1}{2}z_1z_2 + \frac{1}{2}z_2z_4 && \text{(43b)} \\ h_3(z) &= (1-\beta)\, z_1z_3 + \frac{1}{2}z_1z_4 && \text{(43c)} \\ h_4(z) &= \beta z_1z_3 + z_2z_3 + \frac{1}{2}z_1z_4 + z_2z_4\,. && \text{(43d)} \end{aligned}$$

Exploiting constraints (41), formulas (43) can be rewritten as

$$\begin{aligned} h_2(z) &= q_2(z_2, z_4) = -\alpha z_2 + \left(\frac{1}{2} - \alpha\right) z_4 + \alpha z_2 z_4 + \alpha \\ h_4(z) &= q_4(z_2, z_4) \\ &= (1-\beta)\, z_2 + \left(\frac{1}{2} - \beta\right) z_4 - \left(\frac{1}{2} - \beta\right) z_2 z_4 + \beta \\ h_1(z) &= 1 - q_2(z_2, z_4) \\ h_3(z) &= 1 - q_4(z_2, z_4) \end{aligned}$$

The stability analysis of system (3)-(40) can be achieved by studying the stability properties of system (42)-(44).

Proposition 4.1: The set $\{z | z_1 + z_2 = z_3 + z_4\}$ is invariant for system $z(k+1) = h(z(k))$ and contains a unique exponentially globally stable equilibrium point.

Proof: The first part of the proposition can be straightforwardly derived from the previous discussions on system properties.

Consider now any two pairs (z_2, z_4) and $(\bar{z}_2, \bar{z}_4)$ with $z_2, z_4, \bar{z}_2, \bar{z}_4$ in $[0, 1]$ and let

$$\tilde{z}_2 \triangleq z_2 - \bar{z}_2\,, \qquad \tilde{z}_4 \triangleq z_4 - \bar{z}_4$$

and

$$\tilde{q}_2 \triangleq q_2(z_2, z_4) - q_2(\bar{z}_2, \bar{z}_4)\,, \qquad \tilde{q}_4 \triangleq q_4(z_2, z_4) - q_4(\bar{z}_2, \bar{z}_4)$$

Using equation (44) one can easily show that

$$\begin{aligned} \tilde{q}_2 &= \left(\frac{1}{2} - \alpha\right) \tilde{z}_4 - \alpha \tilde{z}_2 + \alpha(z_2 z_4 - \bar{z}_2 \bar{z}_4) && \text{(45a)} \\ \tilde{q}_4 &= \left(\frac{1}{2} - \beta\right) \tilde{z}_4 + (1-\beta)\tilde{z}_2 - \\ &\quad \left(\frac{1}{2} - \beta\right) (z_2 z_4 - \bar{z}_2 \bar{z}_4)\,. && \text{(45b)} \end{aligned}$$

The term $(z_2 z_4 - \bar{z}_2 \bar{z}_4)$ can be factorized as $\tilde{z}_2 z_4 + \bar{z}_2 \tilde{z}_4$; substitution in (45) gives

$$\begin{aligned} \tilde{q}_2 &= -\alpha(1 - z_4)\tilde{z}_2 + \left[\frac{1}{2} - \alpha(1 - \bar{z}_2)\right] \tilde{z}_4 \\ \tilde{q}_4 &= \left[1 - \beta - \left(\frac{1}{2} - \beta\right) z_4\right] \tilde{z}_2 + \left(\frac{1}{2} - \beta\right)(1 - \bar{z}_2)\tilde{z}_4. \end{aligned}$$

Recalling that $0 \leqslant z_2 \leqslant 1$,

$$\begin{aligned} |\tilde{q}_2| &\leqslant \alpha(1 - z_4)|\tilde{z}_2| + \left(\frac{1}{2} - \alpha\right) |\tilde{z}_4| \\ &\leqslant \alpha|\tilde{z}_2| + \left(\frac{1}{2} - \alpha\right) |\tilde{z}_4|. \end{aligned}$$

Last inequality holds because of z_2 and α bounds. Similarly we have that

$$|\tilde{q}_4| \leqslant (1 - \beta)|\tilde{z}_2| + \left(\frac{1}{2} - \beta\right) |\tilde{z}_4|$$

when $\beta < \frac{1}{2}$. Note that different factorizations of $(z_2 z_4 - \bar{z}_2 \bar{z}_4)$ would give the same bounds for $|\tilde{q}_2|$ and $|\tilde{q}_4|$. Using Lemma 7.1 in the appendix one can deduce there exists a $\kappa < 1$ such that

$$\lambda|\tilde{q}_2| + |\tilde{q}_4| \leqslant \kappa(\lambda|\tilde{z}_2| + |\tilde{z}_4|) \tag{46}$$

for $\lambda \in \left(\frac{1-\beta}{1-\alpha}, \frac{1+2\beta}{1-2\alpha}\right)$ and for any α and β in $[0, \frac{1}{2}]$. Hence $q = (q_2, q_4)$ is a contraction in $[0, 1] \times [0, 1]$ with respect to the norm

$$V_5(z_2, z_4) = \lambda|z_2| + |z_4|.$$

From this we conclude that q possesses an attractive fixed point $(\bar{z}_2, \bar{z}_4)$,that is

$$q_2(\bar{z}_2, \bar{z}_4) = \bar{z}_2, \qquad q_4(\bar{z}_2, \bar{z}_4) = \bar{z}_4.$$

It follows from (46) that

$$V_5(\tilde{z}_2(k), \tilde{z}_4(k)) \leq \kappa^k V(\tilde{z}_2(0), \tilde{z}_4(0))$$

for all $k \geq 0$. This yields the desired result. ■

V. Discussion

There are about 1,098 genes on the human X-chromosome. Most of them code for characteristics other than female anatomical traits. Many of the non-sex determining X-linked genes are responsible for abnormal conditions such as hemophilia A, Duchenne muscular dystrophy, fragile-X syndrome, some high blood pressure dysfunctionalities, congenital night blindness, G6PD deficiency, and the most common human genetic disorder, red-green color blindness. X-linked genes are also responsible for a common form of baldness referred to as "male pattern baldness".

Some of these diseases are severely disabling and affected people usually do not reach the reproduction age. Model (3)–(4) reproduces the severity of the X-linked recessive diseases through an appropriate choice of reproduction rates w_{ij}. Disabling diseases such as hemophilia A and Duchenne muscular dystrophy can be modeled by assigning small values or zero to reproduction rates w_{23} and w_{24}.

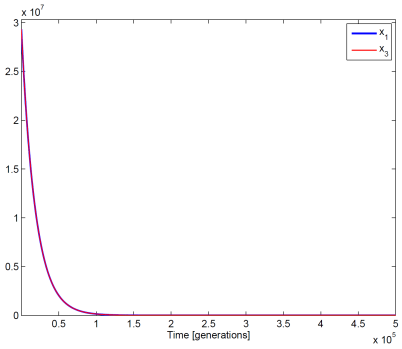
Less serious conditions (such as the red and green color blindness or male pattern baldness) do not cause premature death of affected males and these males usually reach reproduction age. System (3)–(4) could also be exploited to model these diseases by setting all w_{ij} to the same value. However in these cases the number of affected women in the population should also be considered; in fact affected daughters can be born from affected fathers and carrier or affected mothers; they can reach the reproduction age and contribute to the next generation. Model (3)–(4) does not consider affected woman thus it is more suitable to model the epidemiology of severe X-linked diseases. The development of a model with a five vector state comprising the class of affected woman is the object of current research.

We present some numerical results obtained simulating the distribution of hemophilia A disease, a hereditary bleeding disorder caused by a lack of the blood clotting factor VIII, a protein encoded by gene F8 on the X-chromosome. It is largely an inherited disorder; affected males show a reduced reproduction capacity related to the severity of the disease symptoms; carrier females do not usually show any sign of the disease ([15]). The spontaneous mutation rate of the disease is very small; it has been approximately evaluated to be $2.67 \cdot 10^{-5}$ for both males and females ([16]).

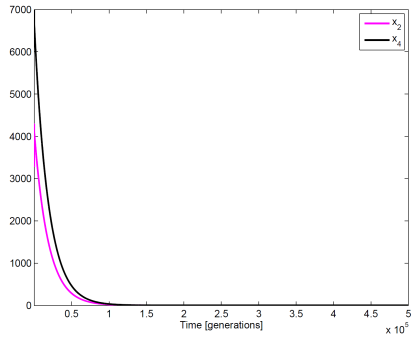
Relying on clinical observations we assigned a reproduction rate of $w_{13} = w_{14} = 1$ to couples formed by a healthy male and a carrier woman, while we choose $w_{23} = w_{24} = 0.62$. The following initial values have been assigned

$$x(0) = [29323162\ 3275\ 29320437\ 6000]^T$$

The initial distribution of males and females (i.e. $x_1(0)$ and $x_3(0)$) has been chosen according to the Italian population (58652874 in 2008 according to [17]). According to the census in [18], 3275 males were affected by hemophilia A. Simulation results are reported in Figure 3; note that, according to the theoretical results we demonstrated (Proposition 3.3), the simulation depicts the extinction of the population; particularly this should happen in a very long period (10^5 generations).



(a) *Trend of* x_1 *and* x_3.



(b) *Trend of* x_2 *and* x_4.

Fig. 3. An example of state evolutions for hemophilia A.

VI. Conclusions

We have presented a discrete-time nonlinear model for X-linked recessive diseases aiming at describing the spread of such diseases in a population; the model takes into account the role of sporadic or *de novo* mutations on the inheritance pattern and distinct reproduction rates according to the health conditions of breeding couples. We analyzed system properties and performed stability analysis of the equilibrium point using Lyapunov functions. Future studies will also need to take into account carrier females who could contribute to disease spread in less severe diseases, as well as the effects on the population of control actions such as prenatal diagnosis.

Acknowledgements

Authors are grateful to Emanuele Durante Mangoni, MD PhD, from the Internal Medicine Unit, Second University of Naples, Italy, for helpful discussions and literature suggestions on X-linked diseases and their inheritance pattern and for his critical review of the manuscript.

VII. Appendix

Lemma 7.1: Consider two real-valued functions g_1 and g_2 of two real variables y_1 and y_2, and suppose that there are scalars $a_{ij} > 0,\ i, j = 1, 2$ with

$$a_{ii} < 1 \quad \text{and} \quad a_{12}a_{21} < (1 - a_{11})(1 - a_{22})$$

such that for any (y_1, y_2), $(\bar{y}_1, \bar{y}_2)$,

$$|\tilde{g}_1| \leqslant a_{11}|\tilde{y}_1| + a_{12}|\tilde{y}_2| \quad (47a)$$
$$|\tilde{g}_2| \leqslant a_{21}|\tilde{y}_1| + a_{22}|\tilde{y}_2|. \quad (47b)$$

where

$$\tilde{y}_1 \triangleq y_1 - \bar{y}_1 \qquad \tilde{y}_2 \triangleq y_2 - \bar{y}_2$$
$$\tilde{g}_1 \triangleq g_1(y_1, y_2) - g_1(\bar{y}_1, \bar{y}_2) \qquad \tilde{g}_2 \triangleq g_2(y_1, y_2) - g_2(\bar{y}_1, \bar{y}_2)$$

Then for any

$$\lambda \in \left(\frac{a_{21}}{1 - a_{11}}, \frac{1 - a_{22}}{a_{12}} \right)$$

there exists a $\kappa < 1$ such that

$$\lambda|\tilde{g}_1| + |\tilde{g}_2| \leqslant \kappa(\lambda|\tilde{y}_1| + |\tilde{y}_2|).$$

Proof: If $\lambda > 0$, combining inequalities (47) yields:

$$\lambda|\tilde{g}_1| + |\tilde{g}_2| \leqslant \left(a_{11} + \frac{a_{21}}{\lambda}\right) \lambda|\tilde{y}_1| + (\lambda a_{12} + a_{22})|\tilde{y}_2|$$

Define $\kappa = \max(\kappa_1, \kappa_2)$ with $\kappa_1 = a_{11} + \frac{a_{21}}{\lambda}$ and $\kappa_2 = \lambda a_{12} + a_{22}$; clearly $\kappa > 0$ and the following inequality holds:

$$\lambda|\tilde{g}_1| + |\tilde{g}_2| \leqslant \kappa(\lambda|\tilde{y}_1| + |\tilde{y}_2|). \quad (48)$$

It is straightforward to verify that if

$$a_{12}a_{21} < (1 - a_{11})(1 - a_{22}) \quad \text{and} \quad \frac{a_{21}}{1 - a_{11}} < \lambda < \frac{1 - a_{22}}{a_{12}}$$

then $\kappa_1, \kappa_2 < 1$; hence $\kappa < 1$ and the desired result follows. ■

Notice that the function $g = [g_1\ g_2]^T$ satisfying the hypotheses of the above lemma is a contraction under a suitable norm.

REFERENCES

[1] J. H. Bennett and C. R. Oertel, "The approach to a random associations of genotypes with random mating," *Journal of Theoretical Biology*, vol. 9, pp. 67–76, 1965.

[2] C. Cannings, "Equilibrium, convergence and stability at a sex-linked locus under natural selection," *Genetics*, vol. 56, pp. 613–618, 1967.

[3] G. Palm, "On the selection model for a sex-linked locus," *Journal of Mathematical Biology*, vol. 1, pp. 47–50, 1974.

[4] D. Luenberger, *Introduction to Dynamic Systems*. John Wiley and Sons, 1979.

[5] T. Nagylaki, "Selection and mutation at an x-linked locus," *Annals of Human Genetics*, vol. 41, no. 2, pp. 241–248, 1977.

[6] J. M. Szucs, "Selection and mutation at a diallelic X-linked locus," *Journal of Mathematical Biology*, vol. 29, pp. 587–627, 1991.

[7] I. Heuch, "Partial and complete sex linkage in infinite populations," *Journal of Mathematical Biology*, vol. 1, no. 4, pp. 331–343, 1975.

[8] C. Del Vecchio, L. Glielmo, and M. Corless, "Equilibrium and stability analysis of x-chromosome linked recessive diseases model," in *Proc. IEEE 51st Annual Conference on Decision and Control (CDC)*, 2012, pp. 4936 – 4941.

[9] ——, "Non linear discrete time epidemiological model for x-linked recessive diseases," in *22nd Mediterranean Conference on Control and Automation, June, Palermo, Italy*, 2014.

[10] V. Dubowitz, *Muscle Disorders in Childhood, Second Edition*. Saunders, 1995.

[11] A. W. F. Edwards, *Foundations of Mathematical Genetics*. Cambridge University Press, 2000.

[12] W. J. Ewens, *Mathematical Population Genetics I: Theoretical Introduction*. Springer, 2004.

[13] M. Lachowicz and J. Miekisz, *From Genetics to Mathematics*, ser. Series on Advances in Mathematics for Applied Sciences. World Scientific Publications, 2009, vol. 79.

[14] J. W. Drake, B. Charlesworth, D. Charlesworth, and J. F. Crow, "Rates of spontaneous mutation," *Genetics*, vol. 148, pp. 1667–1686, 1998.

[15] D. J. Bowen, "Haemophilia a and haemophilia b: molecular insights," *Molecular Pathology*, vol. 55, pp. 1–18, 2002.

[16] J. Haldane, "The mutation rate of the gene for hemophilia, and its segregation ratios in males and females," *Annuals of Eugenics*, vol. 13, pp. 262–71, 1947.

[17] http://www.istat.it/it/charts/popolazioneresidente, 2014, authors' last visit on February 2014.

[18] M. R. Gualan, A. Sferrazza, C. Cadeddu, C. de Waure, F. D. Nardo, G. L. Torre, and W. Ricciardi, "Epidemiologia dell'emofilia A nel mondo e in Italia," *Italian Journal of public health*, vol. 8, no. 2, 2011, in italian.