

Epidemic Models and Protective Mutations

Claudio Parmeggiani
claudiparm@gmail.com

March 30, 2021

Abstract

The paper proposes and discusses an epidemic model which assumes that a population, affected to a disease, is (partially) protected from the disease. These protections may be generated by some inherited genetic mutations; mutations favoured by the presence, in the past, of a deadly endemic disease (the Haldane hypothesis). The protections can also have a non-genetic origin: the cause can be the age or the sex of the subjects; in this case the model can only be applied for short periods, typically in the early stages of the epidemic. Anyway we shall see that the presence of these protections, if wisely managed, not only does it benefit the whole population, but it also makes the measures taken to contain the epidemic more effective.

1 Introduction

Here it is proposed and discussed an *epidemic model* which assumes that part of a population, affected to a disease, is, in some way and at different levels, *protected* from the disease. These protections may be generated by some inherited *genetic mutations*; mutations eventually favoured by the presence, in the past, of some deadly endemic disease (the Haldane hypothesis, see the Section 2). But these protections may very well have a *non-genetic* origin: the cause can be, say, the age or the sex of the subjects; in this case the epidemic model can, of course, be applied only for short periods, typically in the early stages of the epidemic. And you should take into account the circumstance that these protections can be rapidly lost, due to (random) counter-mutations of the disease agents.

Accordingly, in Sections 3 and 4, it is described a mathematical model which attempts to incorporate the above mentioned fact, that part of the affected population is, to varying degrees, protected; so, for example, while some individuals are fully susceptible, others are only mildly infected and weakly contagious. The Section 5 contains examples and numerical solutions of the differential equations of the model; you shall see that the protections, if wisely managed, not only does it benefit the whole population, but also makes the measures taken to contain the epidemic more effective. In the Appendices are outlined some mathematical techniques here used and are sketched some variant of the model.

The model can be regarded as a sort of extension of to the “classic” Kermack and McKendrick one [5] (see also [7]); here, as in [5], the disease is assumed to be spread only by *contact* and not, say, by the action of some vector. And, for certain matters, see also the still relevant Volterra “Leçons” [11]: the spread of the disease can be seen as a fight between an evolved DNA and a primitive RNA/DNA for the control of the “body”, body which both need to perpetuate themselves.

2 Protective Mutations

It is possible that at the onset of an epidemic, of a completely *new* type, a substantial part of the affected population is, partially or completely, protected? An answer (one of the most plausible, but not the only one, evidently) is that this part of the population has inherited some *protective mutations*, that is some *random* mutations, eventually slightly harmful, but strongly protective against an *old*, perhaps now eradicated, endemic disease. And also protective against the new one. This is one aspect of the so called “Haldane hypothesis” [3]:

1932 – The Causes of Evolution – What is Fitness? A study of the causes of death in man, animals and plants leaves no doubt that one of the principal characters possessing survival value is immunity to disease. Unfortunately, this is not a very permanent acquisition, because the agents of disease also evolve, and on the whole more rapidly than their victims. ... It seems likely that when a species is subjected to a series of attacks by an evolving parasite it may be forced along a path of structural change by its temporarily successful acquisition of immunity. But in the end it may driven, so to say, into a corner, where further immunity involves structural changes which are disastrous to it in its everyday life.

More explicitly, always Haldane, in [4]:

1949 – Disease and Evolution. I want to suggest that the struggle against disease, and particularly infectious diseases, has been a very important evolutionary agent, and that some of its results have been rather unlike those of the struggle against natural forces, hunger, and predators, or with members of the same species. Under the heading infectious disease I shall include, when considering animals, all attacks by smaller organism, including bacteria, viruses, fungi, protozoa, and metazoan parasites.

Evidently the old disease, to have given rise to effective mutations, must have been endemic, long lasting (hundreds of generations) and, if now eradicated, its disappearance must not have been too far in time. For the human species, on Earth, a truly plausible candidate for this role is *malaria* (“aria mala”, evil air): for thousands of years it has spread throughout the planet and has marked the history of peoples. And even when, after a long struggle, it was eradicated, all this occurred in recent times, in the last century; consider, for example, the case of Italy [10] or of Taiwan [6].

Unfortunately the informations about the *genetic legacy* of malaria are very far to be systematized and complete; but see, for example, [12] (WHO 2019) or Carter & Mendis [2]. In this regard it is interesting to note that, apparently, the genetic legacy linked to the presence of *P. falciparum* (predominant in Sub-Saharan Africa) is quite different from the one related to the presence of *P. vivax* (mainly widespread in East Asia, Europe and Middle East, in the past).

3 The rules

3.1 Infectible, infected, recovered or deceased

Let $N(t)$ be the size of some population \mathcal{N} (a *set of individuals*) living, at time t , in a fixed region. Part of the population is affected by some *new* contagious disease and can spread it; taking into account the fact that the disease affects individuals in different ways, the set \mathcal{N} is partitioned into several disjointed classes, $(\mathcal{N}_k)_{k \in K}$; the members of each class are assumed to react differently to infection. K is here a *finite set* (but see the Appendix A); $|K|$ will denote the number of elements of K .

At the onset of the epidemic each \mathcal{N}_k can be split into *two* subclasses, the *uninfected* \mathcal{X}_k , which are, in varying degrees, capable to be infected; and the *infected*, \mathcal{Y}_k , which can eventually transmit the disease. Then, say, different \mathcal{X}_k have different degrees of *susceptibility*: at one extreme there are totally susceptible individuals, at the other extreme individuals who can never catch the disease, they are fully *protected*. Correspondingly, members of different \mathcal{Y}_k present different severity of the infection and variable *infectiousness* and *evidence of symptoms*. We shall denote by \mathcal{X} (\mathcal{Y}) the *union* of all the \mathcal{X}_k (\mathcal{Y}_k); \mathcal{N} is the union of \mathcal{X} and \mathcal{Y} .

Besides, during the epidemic, some individual can *acquire*, after recovering, an antibody driven (typically temporary) *immunity*; these immunities can sometime accumulate in complicate ways, but here we are only marginally interested to them; we shall simply introduce a new subset of \mathcal{N} , \mathcal{Z} (the *immune*) and its subclasses, $(\mathcal{Z}_k)_{k \in K}$, whose members are recovered individuals, from \mathcal{Y}_k , now assumed to be *totally* but *temporarily* immune. The \mathcal{Z}_k are uninfected, as the \mathcal{X}_k , which are also non-immune and can now be more properly referred to as *infectible*. These immunities can also be *induced*, by external intervention; induced immunities are, generally, longer lasting; those ones will not considered here.

If $n_k(t)$, $x_k(t)$, $y_k(t)$ and $z_k(t)$ are the sizes (number of individuals) of \mathcal{N}_k , \mathcal{X}_k , \mathcal{Y}_k and \mathcal{Z}_k ; and if $X(t)$, $Y(t)$ and $Z(t)$ are the sizes of \mathcal{X} , \mathcal{Y} and \mathcal{Z} (everything at time t):

$$X(t) = \sum_k x_k(t), \quad Y(t) = \sum_k y_k(t), \quad Z(t) = \sum_k z_k(t) \quad (1a)$$

$$n_k(t) = \sum_k (x_k(t) + y_k(t) + z_k(t)) \quad (1b)$$

$$N(t) = \sum_k n_k(t) = X(t) + Y(t) + Z(t). \quad (1c)$$

3.2 The spread of the disease

The decrease of the number of the uninfected, x_h , $h \in K$ (and the corresponding increase

of the number of the infected, y_h), per unit of time, due to *infectious contacts*, is assumed to be proportional to x_h ; that is to be equal to $c_h(t) x_h(t)$ where $c_h(t)$ is the (infection) *spread rate* – always ≥ 0 . c_h generally depends on the likelihood of meeting contagious individuals; if only the infected (not the recovered) can spread the disease, we can assert that, in a first, linear approximation ($r_{h,k}(t) \geq 0$, $h, k \in K$),

$$c_h(t) = \sum_k r_{h,k}(t) p_k(t) \quad (2)$$

where the $p_k(t)$ are the *probabilities* that a \mathcal{X}_h (an uninfected member of \mathcal{X}_h) meets an (infected and infectious) \mathcal{Y}_k .

Assuming that the distributions of infected is quite uniform – but this is not always the case – the $p_k(t)$ are simply equal to $\hat{y}_k(t) = y_k(t)/N(t)$, the *density of infected*; $\sum_k (\hat{x}_k(t) + \hat{y}_k(t) + \hat{z}_k(t)) = 1$. Thus we can conclude that the *decrease* of x_h (*increase* of y_h), per unit of time (per day, per week ...), due to infectious contacts, is equal to:

$$x_h(t) \sum_k r_{h,k}(t) \hat{y}_k(t) \quad (3)$$

and that the analogous decrease of X (increase of Y) is:

$$\sum_{h,k} r_{h,k}(t) x_h(t) \hat{y}_k(t). \quad (4)$$

If $|K| = 1$ the above formulae reduce to the well know one, $r(t) X(t) Y(t)/N(t)$.

The *contact matrix* $(r_{h,k})_{h,k \in K}$ may significantly vary over time due to *external* interventions aimed, for example, at preventing uninfected-infected contacts; in most case $r_{h,k}(t)$ can be simply factored as $r(t) \gamma_{h,k}$ (or as $r_h(t) \gamma_{h,k}$) where the $\gamma_{h,k}$ depend only on the disease characteristics and are *constant*, for a stable disease agent. The *contact rate* $r(t)$ is here normalized by setting the *maximum* of the $\gamma_{h,k}$ equal to 1.

Not all the infected can contact others individuals: some of them are *isolated*. This fact can be handled introducing, for every $k \in K$, the *isolation weights*, $q_k(t)$, ≥ 0 and ≤ 1 ; $q_k(t) = 1$ would mean that all the individuals of \mathcal{Y}_k are isolated; we shall always assume that all the $q_k(t)$ are < 1 ; $\sum_k q_k(t) y_k(t) < Y(t)$. Therefore, in the above formulae, $\hat{y}_k(t)$ must be replaced with:

$$\frac{(1 - q_k(t)) y_k(t)}{N_q(t)} \quad (5)$$

where $N_q(t) := N(t) - \sum_k q_k(t) y_k(t)$.

Even the q_k may vary over time due to *external* interventions, aimed, in this case, to detect the presence of *asymptomatic* individuals. But sometimes it is not so easy to recognize the presence of the infection, it is then necessary to do some time (and resources) consuming tests and to organize and regulate the screening activities, all not very popular things. So these activities are often carried out when it is too late, when there are too many infected.

Moreover the number of infected decrease because some of them will recover or die (or they are suppressed or extirpated, being “expendable”); hence, if $a_k(t)$ and $d_k(t)$ are, respectively, the *recovery* and the (infection related) *deaths rates* (dead-plus-suppressed),

$$- (a_k(t) + d_k(t)) y_k(t) \quad (6)$$

is the corresponding change of $y_k(t)$, per unit of time. Then, if $b_k(t)$ is the rate of (acquired) *immunity loss*, the change of $z_k(t)$, per unit of time, is:

$$+a_k(t) y_k(t) - b_k(t) z_k(t) \quad (7)$$

while the consequent changes of $x_k(t)$ is:

$$+b_k(t) z_k(t). \quad (8)$$

Generally $a_k + d_k > 0$, but $a_k = d_k = 0$ for *disease tolerant* individuals: infectible individuals, eventually infectious when infected, but always “healthy”: they *coexist* with the disease agents.

The a_k and d_k may vary over time, sometimes dramatically, due to external causes (improvements in therapies or, for example, collapse of healthcare structures); the b_k , on the contrary, are generally more constant.

If $a_k + d_k$ is > 0 , the $\tau_k = 1/(a_k + d_k)$ can be understood as the *average durations* of the infection: in that period of time an average infected (\mathcal{Y}_k) heals or dies. The dimensionless “indices” $\tau_k a_k$ and $\tau_k d_k$ are, respectively, the *probabilities* that someone, after contracting the disease, recover or die, by the disease; $\tau_k a_k + \tau_k d_k = 1$, how it should be. Likewise the $\sigma_k = 1/b_k$ (if b_k is > 0) are the average durations of the acquired immunity; if σ_k is of the order of magnitude of the average life span of the population, the immunity can be regarded as permanent.

Note that if $|K|$ is > 1 the increase of y_k is *not proportional* to y_k and the dimensionless parameters $r\tau_k$ (or $r_k\tau_k$) lose, as they are, their meaning. See the Appendix B, The infected.

Finally, for long-lasting epidemics (when the protections are of genetic origin), we have to consider the changes of number of individuals due to *births* and *natural deaths*. It can happen that the newborns are infected or immune; assuming that they are not, we can say that these changes – births, ν and deaths, μ – are, per unit of time and for each one of the subclasses \mathcal{X}_k , \mathcal{Y}_k and \mathcal{Z}_k , simply equal to:

$$\begin{array}{ccc} \nu_k n_k(t), & 0, & 0 \\ \mu_k x_k(t), & \mu_k y_k(t), & \mu_k z_k(t) \end{array} \quad (9)$$

The deaths rates are here assumed to be the same for the \mathcal{X} , \mathcal{Y} and \mathcal{Z} ; ν_k and μ_k are constant, $\nu_k \geq 0$, $\mu_k > 0$; it is generally not necessary to introduce a carrying capacity for the ν . In some case the time period examined is relatively short – if, for example, we are only interested to the initial stages of the epidemic – and the ν and μ can be ignored.

4 The equations

Here I is the real interval $[0, +\infty[$; the *linear space* $E = \mathbb{R}^K$ is normed by, say, $\|u\| = \sum_k |u_k|$;

$$E_+ := \{u \in E : u_k \geq 0\};$$

a, b, d, q are *continuous* and *bounded* maps $I \rightarrow E_+$, $q_k < 1$; the $r_{h,k}$ are *continuous* and *bounded* functions $I \rightarrow \mathbb{R}_+$, $\sum_{h,k} r_{h,k} > 0$.

Translating the “changes per unit of time” with derivatives with respect to t , we arrive at a system of *ordinary differential equations*, in \mathbf{x} , \mathbf{y} and \mathbf{z} ($\mathbf{n} = \mathbf{x} + \mathbf{y} + \mathbf{z}$):

$$\mathbf{x}' = -c(t, \hat{\mathbf{y}}) \mathbf{x} + b(t) \mathbf{z} + \nu \mathbf{n} - \mu \mathbf{x} \quad (10a)$$

$$\mathbf{y}' = +c(t, \hat{\mathbf{y}}) \mathbf{x} - (a(t) + d(t)) \mathbf{y} - \mu \mathbf{y} \quad (10b)$$

$$\mathbf{z}' = +a(t) \mathbf{y} - b(t) \mathbf{z} - \mu \mathbf{z} \quad (10c)$$

where \mathbf{x} , \mathbf{y} and \mathbf{z} are *continuously differentiable* maps $I \rightarrow E_+$ assuming some fixed value when $t = 0$. Moreover ($\lambda_k := \nu_k - \mu_k$):

$$\mathbf{n}' = \lambda \mathbf{n} - d(t) \mathbf{y}, \quad N' = \sum_k \lambda_k n_k - D(t) Y \quad (11)$$

and

$$Y' = +C(t, \mathbf{x}, \hat{\mathbf{y}}) - (A(t) + D(t)) Y - \sum_k \mu_k y_k. \quad (12)$$

Above, if a and $\mathbf{x} \in E$, $a\mathbf{x}$ stands for $(a_k x_k)_{k \in K}$, $\hat{\mathbf{y}} = (\mathbf{y} - q(t) \mathbf{y}) / N_q(t)$,

$$c_h(t, \hat{\mathbf{y}}) = \sum_k r_{h,k}(t) \hat{y}_k, \quad C(t, \mathbf{x}, \hat{\mathbf{y}}) = \sum_{h,k} r_{h,k}(t) x_h \hat{y}_k \quad (13)$$

and A, D are defined by $A(t) Y = \sum_k a_k(t) y_k$, $D(t) Y = \sum_k d_k(t) y_k$.

We can eventually assume that the maps a, b, d, q and r are only *measurable* and *bounded* and that the functions \mathbf{x} , \mathbf{y} and \mathbf{z} are only *absolutely continuous* – hence almost everywhere differentiable. For this and all the other mathematical aspects of the problem, see Birkhoff & Rota [1] and Sansone & Conti [8].

When $t = 0$ (the *epidemic onset*) all the n_k and x_k are assumed to be > 0 (typically $x_k(0) \cong n_k(0)$); it follows that all the $n_k(t)$ and $x_k(t)$ are *always* > 0 , for $t \geq 0$. See the Appendix B for this and the following remarks.

$Y(0)$ also has to be > 0 (otherwise $Y(t)$ would be identically $= 0$), $Y(t)$ is then always > 0 . But, for some $k \in K$, $y_k(0)$ can be $= 0$; generally, in this case, $y_k(t) > 0$, for $t > 0$. However it may be that instead $y_k(t)$ remains $= 0$ (so $x_k(t) = n_k(t)$): these individuals are not really taking part in the epidemic, they are fully protected.

$Z(0)$ is obviously $= 0$, but, apart from special cases ($a_k = 0$ or $y_k = 0$), $z_k(t)$ is > 0 , for $t > 0$.

Note that the $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ differential system is very far from being *autonomous*: the r and the q are frequently varied along the epidemic in a planned and controlled way, having decided what to maximize and what to minimize. However, it can happen – mainly at the epidemic onset – that r and q are instead capriciously modified, following the results of opinion polls or the decrees of some autocrat.

In any case, in Section 5, will be examined some simple case; the equations are *numerically integrate* choosing two different, somehow opposite, planning and control strategies for the contact rate and the isolation weights. It will also be examined the case of *deadly diseases* ($a_k = 0$ and $d_k > 0$ for some $k \in K$) and the *endemic phase* of the disease – when the density of infected, Y/N , is, so to speak, *sensibly constant*.

5 Numerical integrations

In all examples of Sections 5.1 and 5.2, $K = \{S, P\}$: the x_S are totally susceptible, unprotected individuals; the x_P are less susceptible and partially protected (they cannot dye by the disease); $r_{h,k}(t) = r(t) \gamma_{h,k}$;

$$\gamma = \begin{pmatrix} 1.0 & 0.7 \\ 0.8 & 0.5 \end{pmatrix} \quad (14)$$

and (D is some unit of measure of time, say *days*)

$$\begin{array}{lll} a_S = 0.04 & d_S = 0.00004 & \tau_S = 25 \text{ D} \\ a_P = 0.15 & d_P = 0 & \tau_P = 6.7 \text{ D} \end{array} \quad (15)$$

$$\begin{array}{ll} b_S = 0.005 & \sigma_S = 200 \text{ D} \\ b_P = 0.008 & \sigma_P = 125 \text{ D} \end{array} \quad (16)$$

The probability that an unprotected individual dies, after contracting the disease, is equal to $d_S \tau_S = 0.001$. The a , b and d are assumed to be constant; the contact rate r and the isolation weights q_S and q_P are instead *variable*:

- when $\mathbf{0} \leq t < t_1$ the presence of the disease is *not perceived* (or, for “some reason”, underestimated) and:

$$r = 0.20 \quad q_S = 0.0 \quad q_P = 0.0 \quad (17)$$

- when $t_1 \leq t < t_2$ two different strategies are adopted. In the first one r is slightly decreased and *kept constant*; the tracking and screening activities are *enhanced*, so that, say:

$$r = 0.18 \quad q_S = 0.40 \quad q_P = 0.20 \quad (18)$$

In the second strategy r is *continuously adapted* to Y/N (via the *function* F) while the isolation weights are less markedly increased:

$$r = F(Y/N) \quad q_S = 0.25 \quad q_P = 0.10 \quad (19)$$

- when $t_2 \leq t \leq t_M$ it is assumed, *hopefully* but perhaps unwisely and surely recklessly, that the epidemic is about to end. Consequently the pre-disease value of r is restored while q_S and q_P are kept at quite low levels:

$$r = 0.20 \quad q_S = 0.10 \quad q_P = 0.05 \quad (20)$$

5.1 Few protected individuals

Now we have to fix the *initial conditions*; it is here assumed that only *one third* of the population is protected; and that, when $t = 0$, there are $10^{-6} N_0$ ($N_0 := N(0)$) infected individuals (protected):

$$\begin{array}{ll} n_S(0) = x_S(0) = 0.65 & y_S(0) = 0.0 \\ n_P(0) = x_P(0) + y_P(0) = 0.35 & y_P(0) = 0.000001 \end{array} \times N_0 \quad (21)$$

As always in this Section, $\nu_{S/P} = 0$, $\mu_{S/P} = 0$ and $z_{S/P}(0) = 0$.

Keeping r constant ($= 0.18$) for $90 \text{ D} \leq t \leq 900 \text{ D}$ (the *first strategy*), we arrive at the graph of **Fig. 1a**: after the initial peak, the infected curve flattens out, through some *damped oscillations* (sometimes improperly called “waves”).)

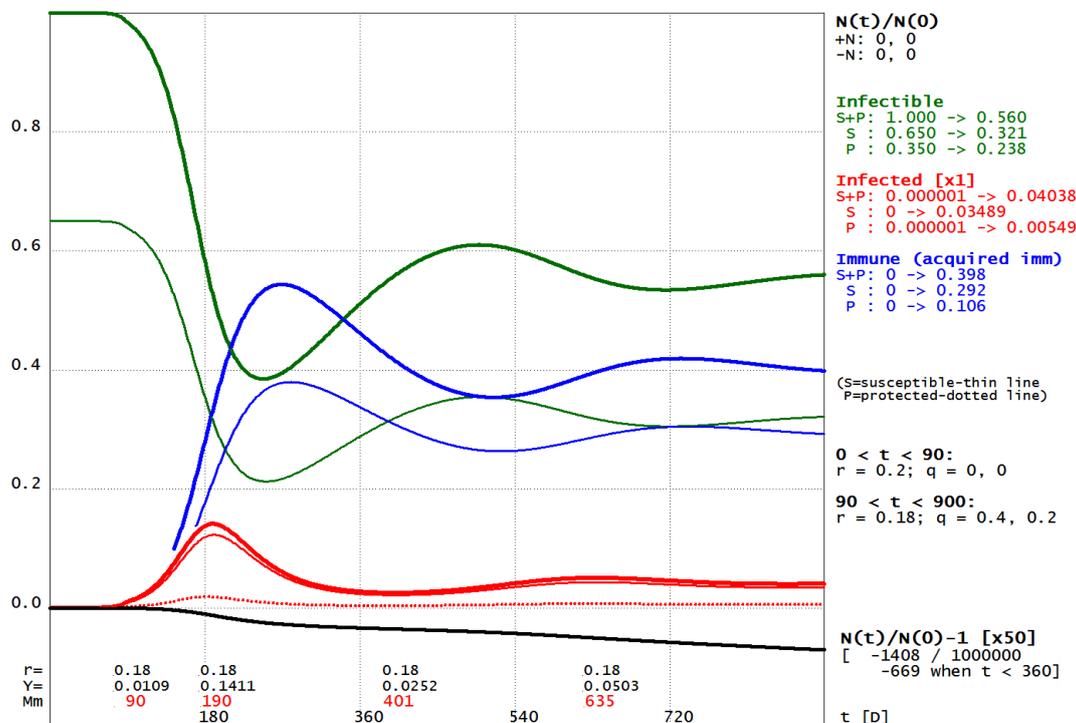


Fig. 1a - $x_S(0) = 0.65$, $x_P(0) = 0.35$. r is decreased from 0.20 to 0.18, $q = 0.4, 0.2$.

But if, following the *second strategy*, r is continuously adapted to Y/N – the total number of infected – by, for example,

$$r(t) = F(Y(t)/N(t)) = \frac{r(0)}{(1 + \kappa Y(t)/N(t))^\lambda} \quad (22)$$

with $\kappa = 3$ and $\lambda = 2$ (always for $90 \text{ D} \leq t \leq 900 \text{ D}$) you will get the graph of **Fig. 1b**.

You can see, as it was easily predictable, that the infected curve is now more smooth and flat (at the curve peak, $r = 0.12$), but the “results” are, more or less, the same of **Fig. 1a**. Obviously a flatter curve places less stress on healthcare facilities; but the continuous adaptation, more or less forcibly, of r to Y/N has a considerable social and economic cost.

What happens if, from a certain point in time ($t \geq 345 \text{ D}$), the pre-disease parameters are restored? You will inevitably return to the starting point ($t = 0$, “back to square one”), now clearly with more infected and (fortunately) with more immune too; moreover this second, “artificial” peak, is quite similar to the first.

See **Fig. 1c**, where it is taken into consideration the second strategy; for the first one the situation is similar, but slightly better: there are *more immune* individuals.

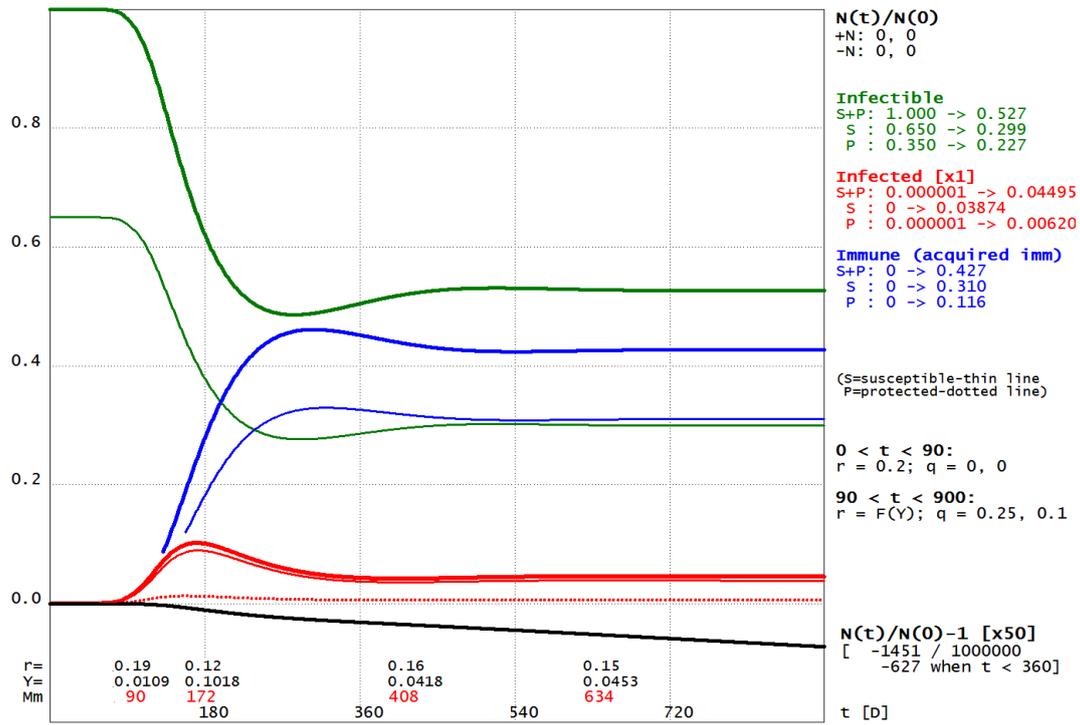


Fig. 1b - $x_S(0) = 0.65$, $x_P(0) = 0.35$. $r(0) = 0.20$, $r = r(0) / (1 + 3Y)^2$, $q = 0.25, 0.1$

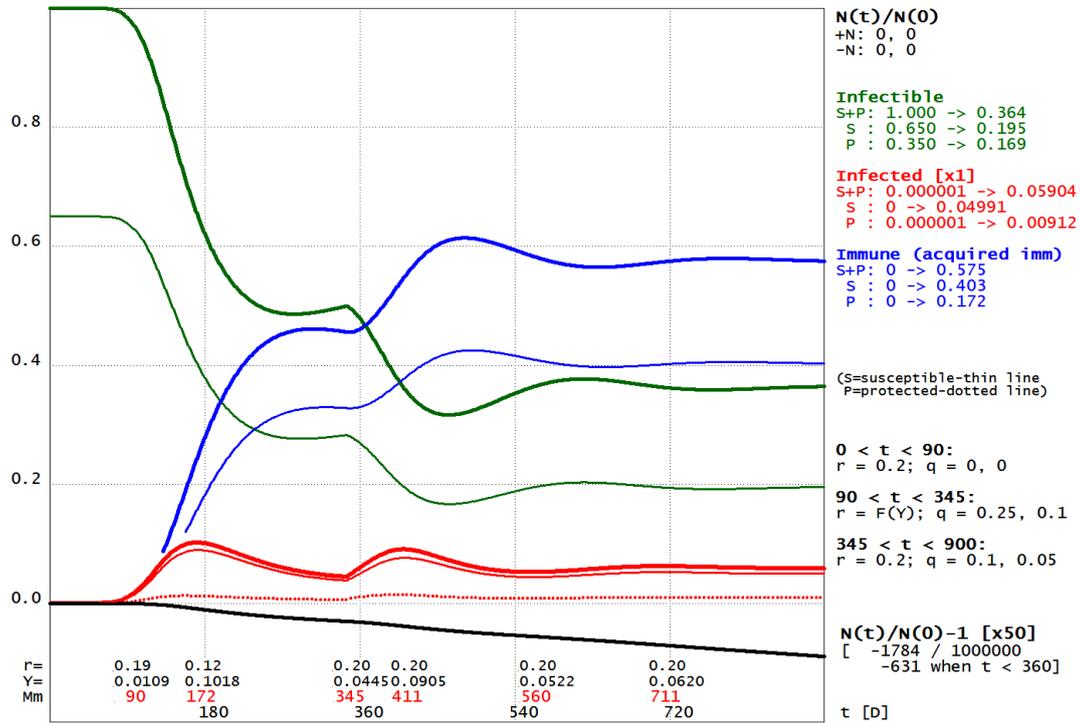


Fig. 1c - $x_S(0) = 0.65$, $x_P(0) = 0.35$; $r = 0.20 / (1 + 3Y/N)^2$. For $t \geq 345$, $r = 0.20$.

5.2 Many protected individuals

Now *two thirds* of the population are protected and, when $t = 0$, there are $2 \times 10^{-6} N_0$ infected individuals (protected):

$$\begin{aligned} n_S(0) &= x_S(0) = 0.35 & y_S(0) &= 0.0 \\ n_P(0) &= x_P(0) + y_P(0) = 0.65 & y_P(0) &= 0.000002 \end{aligned} \times N_0 \quad (23)$$

Again $\nu_{S/P} = 0$, $\mu_{S/P} = 0$ and $z_{S/P}(0) = 0$.

The situation is obviously more favorable; we can now intervene in a very soft way, for example keeping $r = 0.20$ and setting $q_S = 0.25$, $q_P = 0.10$ (when $t \geq 90$ D) without serious consequences; see **Fig. 2a**.

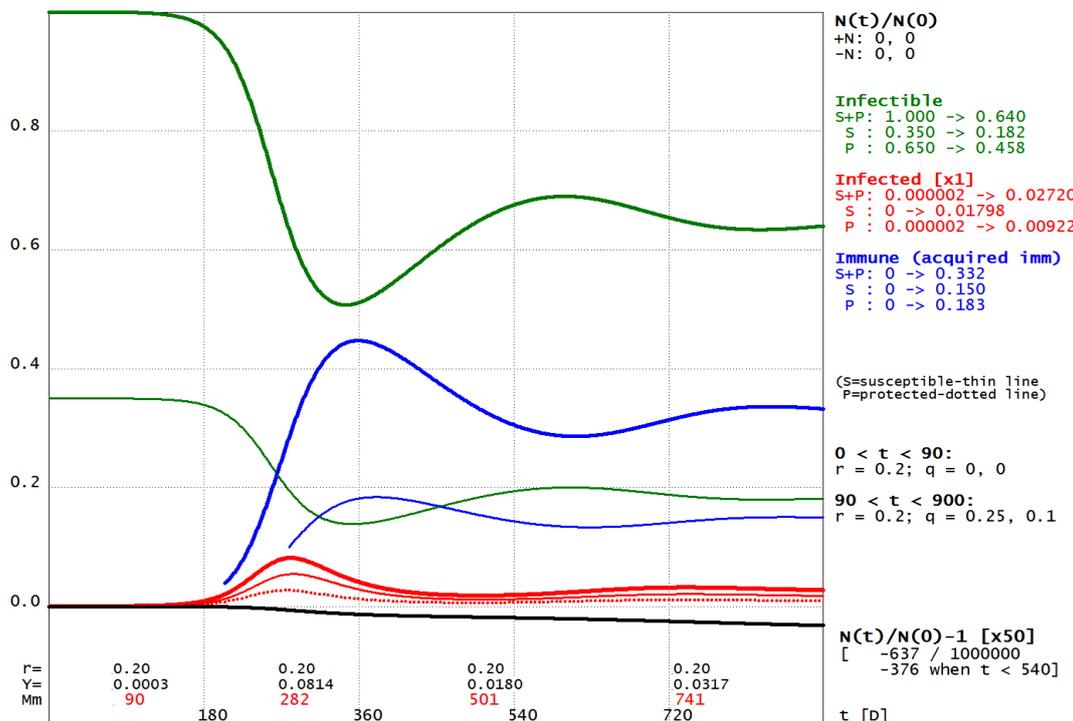


Fig. 2a - $x_S(0) = 0.35$, $x_P(0) = 0.65$. r is kept equal to 0.20, $q = 0.25, 0.1$.

A decidedly more energetic intervention, $r = 0.18$, $q_S = 0.40$ and $q_P = 0.20$ (always when $t \geq 90$ D) can in fact almost suppress the infection; see **Fig. 2b**. But there is a catch: when the initial conditions are restored, the number of immune is still small and the infected and immune curves will have a marked peak, as in **Fig. 2c**.

Observe that the initial peak can be, by a *very* energetic intervention, *completely eliminated*. If, in the example of **Fig. 2b** we set, when $t = t_1 = 90$ D, $r = 0.14$, then $Y(t) \leq Y(t_1) = 0.00029 N_0$ (for $t \geq t_1$) and the deaths number due to the disease is almost nil. But sooner or later you will have to reset r to its pre-disease value and thus generate the usual peak of infected – unless, in the meantime, a way is found to “produce” a substantial number of *immune* individuals.

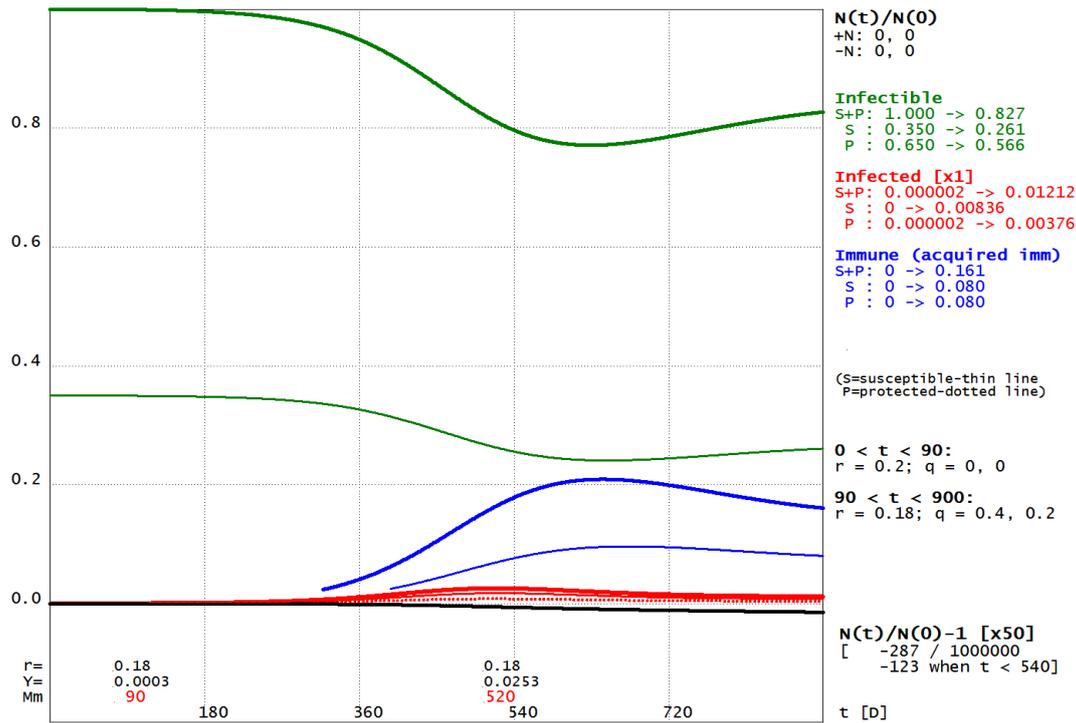


Fig. 2b - $x_S(0) = 0.35$, $x_P(0) = 0.65$. For $t \geq 90$, $r = 0.18$, $q = 0.4, 0.2$.

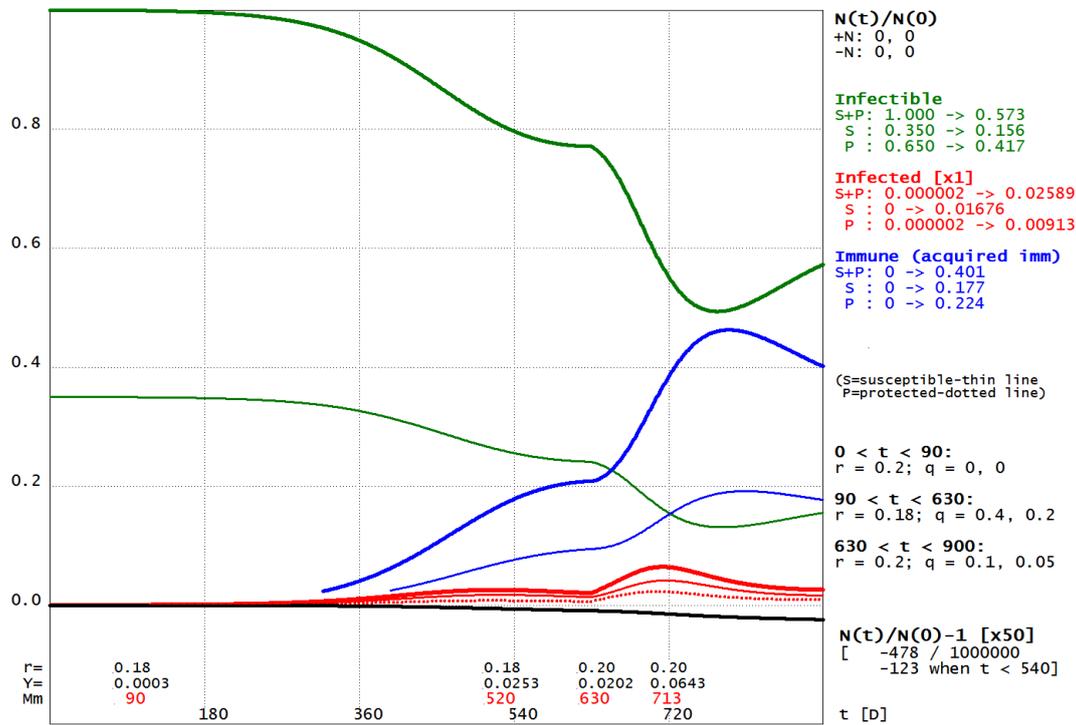


Fig. 2c - $x_S(0) = 0.35$, $x_P(0) = 0.65$. $r (= 0.20)$ will be restored when $t = 630$.

5.3 Deadly Diseases

If, for some $k \in K$, $a_k = 0$ and $d_k > 0$ (a *deadly disease*, for the x_k), then $z_k = 0$ (Appendix B); moreover, if ν_k and μ_k are ignored, x_k *monotonically decreases*. But now we are interested to a more realistic case so, in the following example (**Fig. 3a**), $K = \{S, P\}$ and

$$\nu_{S/P} = 0.00003 \quad \mu_{S/P} = 0.00003 \quad (24)$$

per day (= 0.0108 per year, Y; 1 Y = 360 D); γ is the same of Sections 5.1 and 5.2 while

$$\begin{aligned} a_S &= 0.005 & d_S &= 0.1 & \tau_S &= 9.5 \text{ D} \\ a_P &= 0.025 & d_P &= 0.001 & \tau_P &= 38.5 \text{ D} \end{aligned} \quad (25)$$

$$b_{S/P} = 0.00003 \quad \sigma_{S/P} = 92.6 \text{ Y}. \quad (26)$$

The dying probabilities (by the disease) of the susceptible and the protected are, respectively, equal to $d_S \tau_S = 0.95$ and to $d_P \tau_P = 0.038$.

Now *three quarters* of the population are (somewhat) protected and, when $t = 0$, there are $2 \times 10^{-6} N_0$ infected individuals (protected):

$$\begin{aligned} n_S(0) &= x_S(0) = 0.25 & y_S(0) &= 0.0 \\ n_P(0) &= x_P(0) + y_P(0) = 0.75 & y_P(0) &= 0.000002 \end{aligned} \times N_0 \quad (27)$$

$z_{S/P}(0) = 0$. Moreover $r = 0.14$, when $t \leq 120$ D and $r = 0.12$, $q = 0.75, 0.2$, when $t > 120$ D; r is reset to 0.14 when $t = 3.5$ Y. Hence:

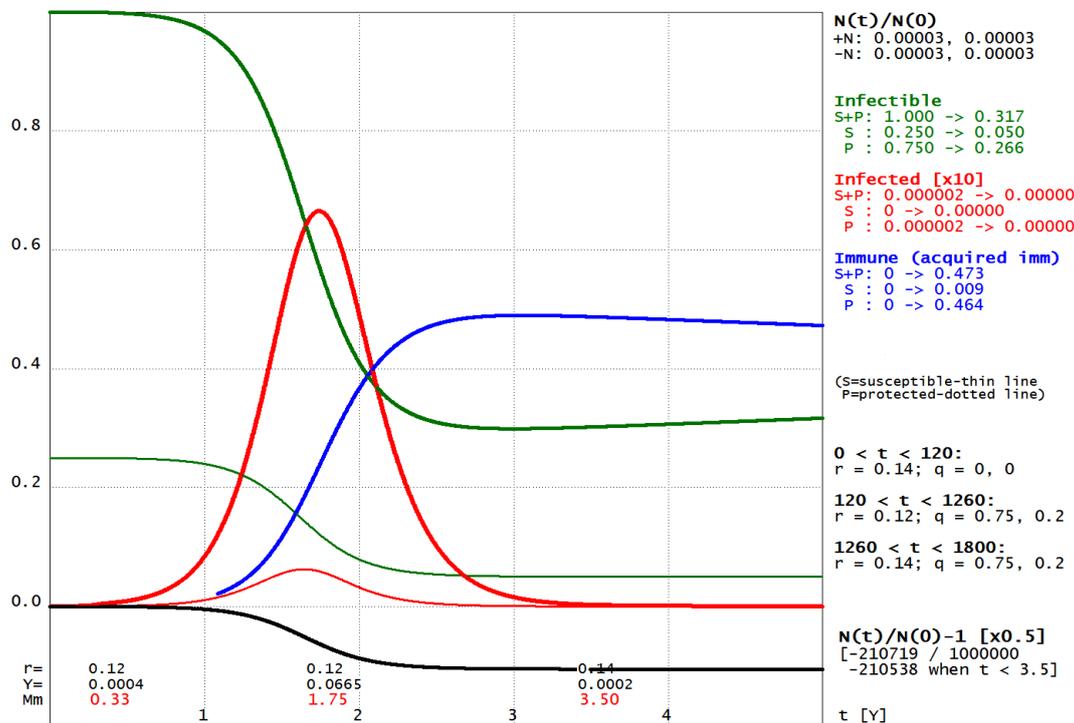


Fig. 3a - $x_S(0) = 0.25$, $x_P(0) = 0.75$. $r(0) = 0.14$; for $t \geq 120$, $r = 0.12$, $q = 0.75, 0.2$.

Now X and Z will *oscillate*, as Y . But, in *this case*, Y rapidly reduces to negligible values: after 10 years (from the epidemic onset) $Y \cong 7 \times 10^{-10} N_0$ and, at its first, absolute, minimum (after 15 years), $Y \cong 4 \times 10^{-11} N_0$; also if $N_0 \cong 10^9$ the epidemic is, “for all practical purpose”, *definitely over* – if it is not reintroduced from the outside.

5.4 Endemic Phases

Diseases with rapid and efficient diffusion, but relatively *low mortality* rates, generally quickly reach an *endemic phase*: see **Fig. 4a**, which shows a *five years* continuation of **Fig. 2a**. Here it is assumed that:

$$\nu_{S/P} = 0.00003 \quad \mu_{S/P} = 0.00003. \quad (28)$$

Things do not change significantly if, after an initial energetic intervention, the r and q are (inevitably) reset close to their initial values. See **Fig. 4b**, which shows a *five years* continuation of **Fig. 2c**: the endemic phase is only slightly *postponed*. And, when $t_M = 5 \text{ Y}$, the values of X , Y and Z are, in both cases, almost the same:

$$X(t_M) \cong 0.648N_0, \quad Y(t_M) \cong 0.028N_0, \quad Z(t_M) \cong 0.323N_0 \quad (29)$$

while $N(t_M) \cong 0.999N_0$.

Generally the *deadly diseases* do not enter in the endemic phase, they die out (as in **Fig. 3a**), but it may happen that, for *very low values* of the contact rate r , the first minimum of the infected curve is still high enough, then the following minima of the curve will be *even higher* and the disease will move towards an endemic phase; see the characteristic **Fig. 4c**.

In this case

$$\begin{aligned} \nu_S &= 0.000035 & \nu_P &= 0.000035 \\ \mu_S &= 0.000032 & \mu_P &= 0.000030 \end{aligned} \quad (30)$$

per day; they are chosen in such a way that, after 125 years, $N(t) \cong N(0)$. γ and the a , b and d are the same of Section 5.3, **Fig. 3a**.

Again *three quarters* of the population are protected and, when $t = 0$, there are $2 \times 10^{-6} N_0$ infected individuals, as in Section 5.3. But now it is always $r = 0.09$; $q_S = 0.75$, $q_P = 0.2$, for $t > 120 \text{ D}$. Hence, after 125 years, $N = 1.006N_0$ and

$$\begin{aligned} n_S &= 0.092 & x_S &= 0.090 & z_S &= 0.002 \\ n_P &= 0.914 & x_P &= 0.710 & z_P &= 0.204 \end{aligned} \times N_0, \quad (31)$$

you can see that the susceptible are in the process of disappearing.

It may be convenient to define a sort of *measure of the endemicity* of the epidemic, at time t . For example:

$$\omega_Y(t) = 2 \frac{\beta(t) - \alpha(t)}{\beta(t) + \alpha(t)} \quad (32)$$

where $\alpha(t) = \inf_{s \geq t} Y(s)/N(s)$ and $\beta(t) = \sup_{s \geq t} Y(s)/N(s)$. $\omega_Y(t)$ has to be “small”: in the case of **Fig. 4a**, $\omega_Y(t) = 0.02$ when $t = 3.3 \text{ Y}$ while, for **Fig. 4b**, $\omega_Y(t) = 0.025$ when $t = 3.9 \text{ Y}$.

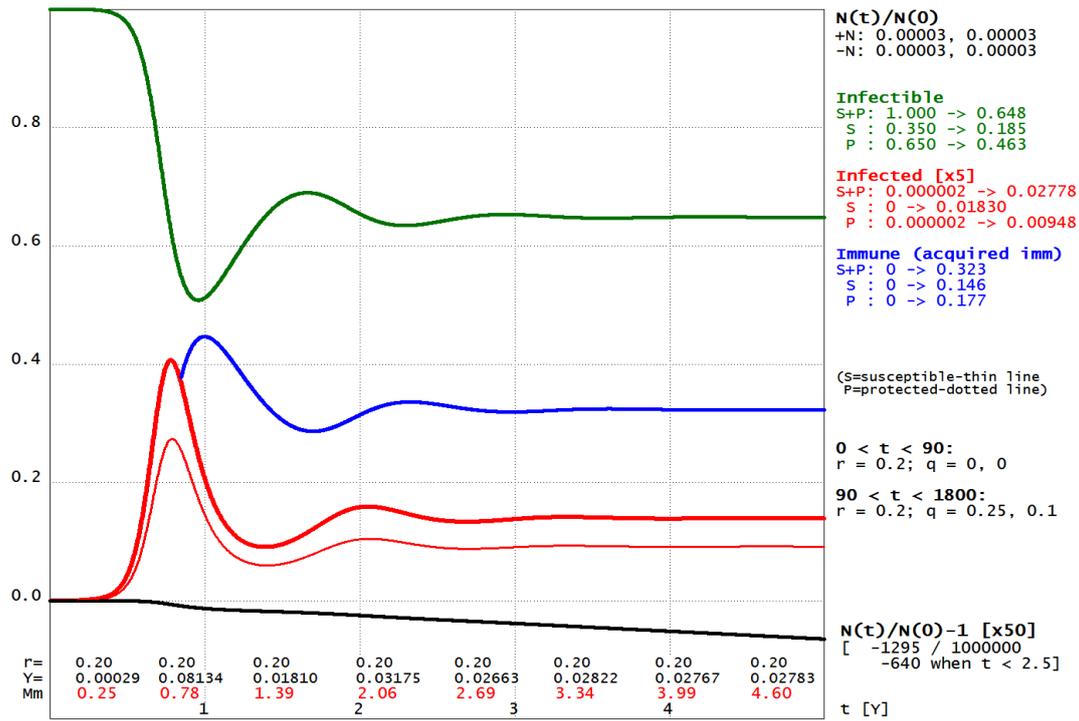


Fig. 4a - $x_S(0) = 0.35$, $x_P(0) = 0.65$. r is kept equal to 0.20, $q = 0.25, 0.1$.

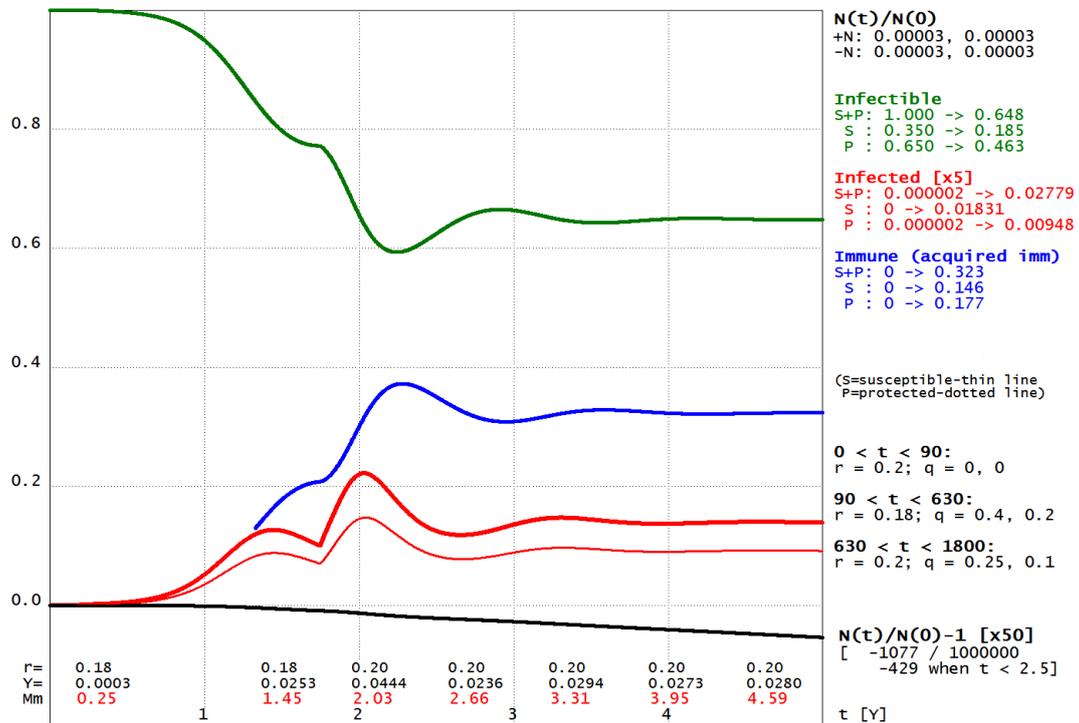


Fig. 4b - $x_S(0) = 0.35$, $x_P(0) = 0.65$. $r(0) = 0.20$; r will be restored when $t = 630$.

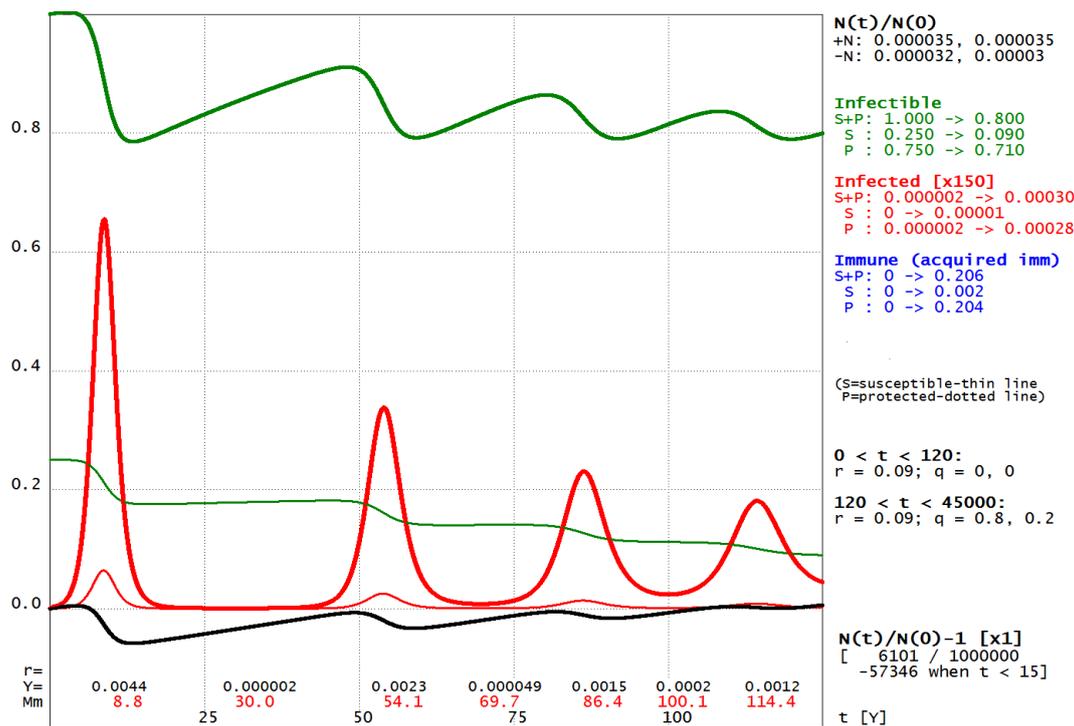


Fig. 4c - $x_S(0) = 0.25$, $x_P(0) = 0.75$. $r(0) = 0.09$; for $t \geq 120$, $r = 0.09$, $q = 0.8$, 0.2 .

The “final” endemic phase will begin after almost 300 years, depending on the values of the ν and μ . Note that a casual observer, who has come into contact with the disease in its final endemic phase (and without other information), may be tempted to attribute the origin of the disease to “unknown causes” – or similarly.

5.5 The Tolerant

Now we shall assume that the protected individuals (P) are, in fact, *tolerant*: they *cohabit* (collaborate) with the agents of the disease and both the recovery and the disease death rate are $= 0$.

The infected tolerant are assumed to be slightly infectious:

$$\gamma = \begin{pmatrix} 1.0 & 0.3 \\ 1.0 & 0.3 \end{pmatrix} \quad (33)$$

and

$$\begin{matrix} a_S = 0.04 & d_S = 0.0004 & \tau_S = 24.75 \text{ D} \\ a_P = 0 & d_P = 0 & - \end{matrix} \quad (34)$$

$$b_S = 0.01 \quad \sigma_S = 100 \text{ D} \quad (35)$$

The probability ($d_S\tau_S$) that a susceptible dies, after contracting the disease is $= 0.01$.

One quarter of the population is assumed to be tolerant; and, when $t = 0$, there are $10^{-6}N_0$ infected (tolerant):

$$\begin{aligned} n_S(0) &= x_S(0) = 0.75 & y_S(0) &= 0.0 \\ n_P(0) &= x_P(0) + y_P(0) = 0.25 & y_P(0) &= 0.000001 \end{aligned} \times N_0 \quad (36)$$

$z_S(0) = 0$ and, in this case ($a_P = 0$), z_P is always $= 0$. The contact rate r is now assumed to be *very low*, $= 0.02$; $q_S = 0.50$, $q_P = 0.05$, for $t > 720$ D.

The birth rates are quite high (but not enough to compensate for the disease death rate, for the susceptible):

$$\begin{aligned} \nu_S &= 0.00004 & \nu_P &= 0.00004 \\ \mu_S &= 0.00003 & \mu_P &= 0.00003 \end{aligned} \quad (37)$$

per day (0.0144 and 0.0108 per year); they are chosen in such a way that, after 300 years, $N(t) \cong N(0)$, see the **Fig. 5a**. Observe that $y_P = x_P = 0.134N_0$ when $t = 19.8$ Y and $y_P = x_S = 0.426N_0$ when $t = 152.3$ Y.

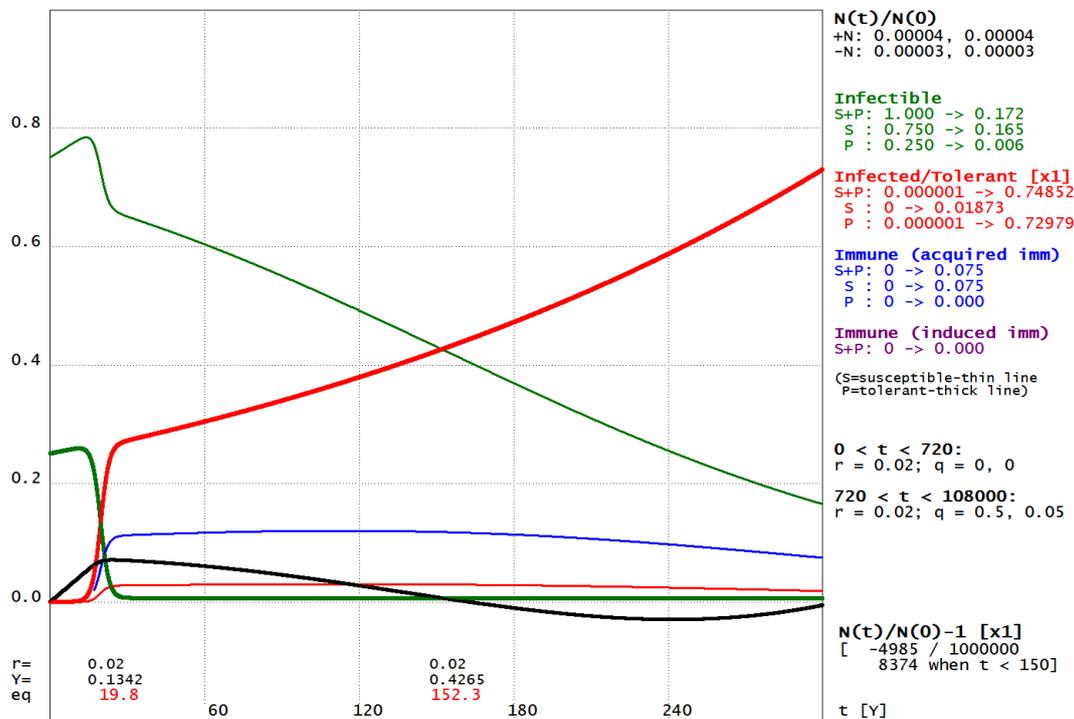


Fig. 5a - $x_S(0) = 0.75$, $x_P(0) = 0.25$. r is kept equal to 0.02, $q = 0.5, 0.05$.

Tolerance is assumed to be *congenital*, but the tolerant newborns are here assumed to be *uninfected*. On the other hand, if all of them are instead infected, the differential equations must be slightly modified:

$$x'_P = \dots + \lambda_P x_P, \quad y'_P = \dots + \lambda_P y_P; \quad (38)$$

however things do not change significantly.

The presence of infectious tolerant obviously poses some social and ethical dilemma.

Appendices

A The infinite-dimensional case

The model can be extended in a obvious way assuming that K is a generic compact (then complete) metric space, that dk is a *positive measure* on K and that $t \in I \mapsto x_t, y_t$ and z_t are continuously differentiable maps $I \rightarrow E = \mathbf{L}^1(K, dk)$, the Banach space of (equivalence classes of) dk -summable functions $K \rightarrow \mathbb{R}$. $n_t = x_t + y_t + z_t$ and, if H is a measurable subset of K ,

$$N_H(t) = \int_H n_t(k) dk \quad (39)$$

is the “number” of individuals characterized by H . $X_H(t) = \int_H x_t(k) dk$, $X(t) = X_K(t)$ and likewise for Y and Z ; $N(t) = X(t) + Y(t) + Z(t)$.

Furthermore, for every $t \in I$, r_t is now a continuous map $K \times K \rightarrow \mathbb{R}$, $r_t(h, k) \geq 0$, $\int_{K \times K} r_t(h, k) dhdk > 0$ and

$$c(t, \hat{y})(h) = \int_K r_t(h, k) \hat{y}(k) dk, \quad C(t, x, \hat{y}) = \int_{K \times K} r_t(h, k) x(h) \hat{y}(k) dhdk; \quad (40)$$

r and \hat{y} (as a, b, d, q) are defined as in the finite-dimensional case. We arrive at a system of differential equations in a infinite-dimensional Banach space; the theory of such systems does not differ significantly from the finite-dimensional theory; see, for example, L. Schwartz [9].

B Differential inequalities

I is a real, open interval; a and b are *bounded* and *continuous* functions $I \rightarrow \mathbb{R}$, $a \leq b$; u is a *continuous differentiable* function $I \rightarrow \mathbb{R}$, $u \geq 0$. Then put $\alpha(t, t_0) = \int_{t_0}^t a(s) ds$, $\beta(t, t_0) = \int_{t_0}^t b(s) ds$ ($t_0, t \in I$; $t_0 \leq t$) and assume that, on I ,

$$ua \leq u' \leq ub. \quad (41)$$

Now the functions $t \mapsto e^{-\alpha(t, t_0)} u(t)$ and $t \mapsto e^{-\beta(t, t_0)} u(t)$ are *continuous differentiable* and their derivatives are respectively equal to $e^{-\alpha(t, t_0)} (u'(t) - u(t) a(t)) \geq 0$ and to $e^{-\beta(t, t_0)} (u'(t) - u(t) b(t)) \leq 0$ so, for $t \geq t_0$:

$$e^{-\alpha(t, t_0)} u(t) - u(t_0) \geq 0, \quad e^{-\beta(t, t_0)} u(t) - u(t_0) \leq 0 \quad (42)$$

or

$$e^{\alpha(t, t_0)} u(t_0) \leq u(t) \leq e^{\beta(t, t_0)} u(t_0). \quad (43)$$

If $u(t_0) > 0$, $u(t)$ is > 0 , for $t \geq t_0$; otherwise $u(t)$ is always $= 0$.

Note that, if $u' < ub$ or $ua < u'$,

$$u(t) > e^{\alpha(t, t_0)} u(t_0) \quad \text{or} \quad u(t) < e^{\beta(t, t_0)} u(t_0), \quad (44)$$

now only for $t > t_0$.

In our case (Section 4) observe that ($h, k \in K, \lambda_k = \nu_k - \mu_k, x_k \leq n_k, y_k \leq n_k$):

$$(-d_k(t) + \lambda_k) n_k \leq n'_k \leq \lambda_k n_k, \quad (45)$$

hence

$$e^{\lambda_k t - \delta_k(t)} n_k(0) \leq n_k(t) \leq e^{\lambda_k t} n_k(0) \quad (46a)$$

$$e^{\lambda_M t - \delta_M(t)} N(0) \leq N(t) \leq e^{\lambda_M t} N(0) \quad (46b)$$

where $\delta_k(t) = \int_0^t d_k(s) ds$ and, generally in the following, for $u \in E$, $u_m = \min_{k \in K} u_k$, $u_M = \max_{k \in K} u_k$. Consequently, if $n_k(0) = 0$, $n_k(t)$ is always $= 0$; otherwise it is > 0 (as for N); obviously only the latter case is here of any interest. Note that, if $\lambda_k = 0$,

$$n_k(t) = n_k(0) - \int_0^t d_k(s) y_k(s) ds \quad (47)$$

while, if $d_k = 0$, $n_k(t) = e^{\lambda_k t} n_k(0)$, as expected.

The infectible. Observe now that, if $N > 0$, $\sum_k \hat{y}_k = (Y - \sum_k q_k y_k) / (N - \sum_k q_k y_k)$ is $\leq Y/N$, so

$$c_h(t, \hat{\mathbf{y}}) \leq r_{hM}(t) \frac{Y}{N} \leq r_{hM}(t) \quad (48)$$

and

$$x'_h \geq -(c_h(t, \hat{\mathbf{y}}) + \mu_h) x_h \geq -(r_{hM}(t) + \mu_h) x_h. \quad (49)$$

Hence

$$e^{-\rho_{hM}(t) - \mu_h t} x_h(0) \leq x_h(t) \leq e^{\lambda_h t} n_h(0) \quad (50a)$$

$$e^{-\rho_M(t) - \mu_M t} X(0) \leq X(t) \leq e^{\lambda_M t} N(0) \quad (50b)$$

where $\rho_{h,k}(t) = \int_0^t r_{h,k}(s) ds$, $r_M(t) = r_{M,M}(t) = \max_{h,k \in K} r_{h,k}(t)$. Consequently $x_h(t)$ is always > 0 if $x_h(0) > 0$ (as for X); here it will be assumed that it is always so.

The infected. We need an upper bound for $C(t, \mathbf{x}, \hat{\mathbf{y}})$, depending on Y :

$$C(t, \mathbf{x}, \hat{\mathbf{y}}) \leq r_M(t) \frac{XY}{N} \leq r_M(t) Y \quad (51)$$

and $(\bar{a}_h(t) := a_h(t) + d_h(t) + \mu_h)$

$$-\bar{a}_M(t) Y \leq Y' \leq (r_M(t) - \bar{a}_m(t)) Y. \quad (52)$$

Hence

$$e^{-\alpha_M(t)} Y(0) \leq Y(t) \leq e^{\rho_M(t) - \alpha_m(t)} Y(0) \quad (53)$$

where $\alpha_h(t) = \int_0^t \bar{a}_h(s) ds$. Consequently, if $Y(0) = 0$, $Y(t)$ is always $= 0$; otherwise it is always > 0 . Obviously only the latter case is here of interest; then some $y_h(0)$ must be > 0 .

Generally $y'_h + \bar{a}_h(t) y_h = c_h(t, \hat{\mathbf{y}}) x_h \geq 0$, hence, if $y_h(0) > 0$,

$$y_h(t) \geq e^{-\alpha_h(t)} y_h(0) > 0 \quad (54)$$

for $t > 0$. But if $y_h(0) = 0$ and $c_h(t, \hat{\mathbf{y}}) > 0$ (for example, if all $r_{h,k}(t)$, $k \in K$, are > 0), $y'_h + \bar{a}_h(t) y_h > 0$ and

$$y_h(t) > e^{-\alpha_h(t)} y_h(0) > 0 \quad (55)$$

for $t > 0$. If instead $y_h(0) = 0$ and $c_h(t, \hat{\mathbf{y}}) = 0$ (all $r_{h,k}(t)$, $k \in K$, are $= 0$, say), $y'_h + \bar{a}_h(t) y_h = 0$ and

$$y_h(t) = e^{-\alpha_h(t)} y_h(0) = 0 \quad (56)$$

for $t > 0$. In this case also $z_h(t)$ is always $= 0$ (see below, The immune) and $x_h(t) = n_h(t) = e^{\lambda_h t} n_h(0)$.

Finally note that ($\hat{X} := X/N_q$, $\bar{r}_{h,k}(t) := r_{h,k}(t)(1 - q_k(t))$):

$$\hat{X} \bar{r}_m - \bar{a}_M \leq \frac{Y'}{Y} \leq \hat{X} \bar{r}_M - \bar{a}_m \quad (57)$$

which shows that if, for some t ,

$$R_m(t) = \hat{X}(t) \bar{r}_m(t) / \bar{a}_M(t) > 1 \quad (R_M(t) = \hat{X}(t) \bar{r}_M(t) / \bar{a}_m(t) < 1),$$

Y is strictly increasing (decreasing), at t . But for an *inhomogeneous* population (with respect to the disease) $\bar{r}_m \ll \bar{r}_M$, $\bar{a}_M \gg \bar{a}_m$ and, generally, $R_m(t) - 1$ and $R_M(t) - 1$ have a *different* sign.

The immune. Posing $\bar{b}_h(t) := b_h(t) + \mu_h$, $z'_h + \bar{b}_h(t) z_h = a_h(t) y_h$. Now, if $z_h(0) = 0$ and $a_h(t) y_h(t) > 0$,

$$z_h(t) > e^{-\beta_h(t)} z_h(0) > 0, \quad (58)$$

where $\beta_h(t) = \int_0^t \bar{b}_h(s) ds$. Consequently if, for some $h \in K$, $a_h(t) y_h(t) > 0$, $Z(t)$ is > 0 , for $t > 0$.

But if $z_k(0) = 0$ and $a_h = 0$ (a deadly disease, if $d_h > 0$) or $y_h = 0$ (fully protected individuals)

$$z_h(t) = e^{-\beta_h(t)} z_h(0) = 0, \quad (59)$$

as expected.

References

- [1] Birkhoff G, Rota G-C 1989 *Ordinary Differential Equations* (New York: John Wiley & Sons)
- [2] Carter R, Mendis K N 2002 "Evolutionary and Historical Aspects of the Burden of Malaria" *Clinical Microbiology Reviews* **15** (4) 564-594
- [3] Haldane J B S 1932 *The Causes of Evolution* (London: Longmans and Green). Reprinted 1993 (Princeton: Princeton University Press)

- [4] Haldane J B S 1949 “Disease and evolution” *La Ricerca Scientifica - Supplemento* **19** 1-11. Reprinted 1992 *Current Science* **63** (9/10) 599-604
- [5] Kermack W O, McEndrick A G 1927 “A Contribution to the Mathematical Theory of Epidemics” *Proc. R. Soc. Lond. A* **115** 700-721
- [6] Liang K C 1999 “Historical review of malaria control program in Taiwan” *The Kaohsiung Journal of Medical Science* **7** (5) 271-277
- [7] Murray J D 1993 *Mathematical Biology* (Berlin: Springer-Verlag)
- [8] Sansone G, Conti R 1956 *Equazioni differenziali non lineari* (Roma: Edizioni Cremonese). Translated 1964 *Non-linear Differential Equations* (Oxford: Pergamon Press)
- [9] Schwartz L 1997 *Analyse II - Calcul différentiel et équations différentielles* (Paris: Hermann)
- [10] Snowden F M 2005 *The Conquest of Malaria: Italy, 1900-1962* (London: Yale University Press)
- [11] Volterra V 1931 *Leçons sur la Théorie Mathématique de la Lutte pour la Vie* (Paris: Gauthier-Villars). Reprinted 1990 (Paris: Editions Jacques Gabay)
- [12] World Malaria Report 2019 (Geneva: World Health Organization)