Tuberculosis propagation

HARESS EL MEHDI

21/01/2019

1 Historical knowledge

In this report, we focus on the epidemiological study of four différentes diseases: Ebola, Dengue, Tuberculosis and HIV. In this section we make a synthesis of the origin of these diseases, their methods of transmission, the state of the disease today and possible treatments différents.

1.1 Tuberculosis

In 1867, tuberculosis was the leading cause of death in Canada. The bacterium that causes tuberculosis, the bacillus, was discovered by the German scientist Robert Koch in 1882. Evidence that tuberculosis was contagious led to coordinated efforts to isolate sufferers in sanatoriums, specialized hospitals where patients could rest, get fresh air and follow a good diet.

The bacillus mainly attacks the lungs, but can also affecter the central nervous system, circulatory system, genital urinary system, bones, joints and even the skin. Tuberculosis is spread from person to person through the air. When a person with pulmonary tuberculosis coughs, sneezes or spits, he or she throws TB bacilli into the air. It suffit to inhale only a few of them to become infected. It can also be spread through the use of utensils (dishes, glass, etc.).

à drinking) unsterilized from an infected person. In quite rare cases a pregnant woman souffrant with active tuberculosis can infect the fetus (vertical transmission). The common symptoms of active pulmonary tuberculosis are cough with sputum, sometimes tinged with blood, tho-rachic pain, weakness, weight loss, fever and night sweats. Tests have been developed to detect tuberculosis and drug resistance.

Tuberculosis can be treated and cured. If tuberculosis is active and therefore sensitive to medication, a standard 6-month course of 4 drugs is given along with assistance from a health care worker. However, when the drugs are not used properly, resistant strains have been found. Multi-resistant tuberculosis (MDR-TB) is a form caused by a bacillus that does not react to isoniazid and rifampicin, the two most common first-line drugs efficaces. However, MDR-TB can be treated and cured with second-line drugs. However, these options are more limited and require a long duration (up to two years of treatment) and expensive, toxic drugs.

Today, tuberculosis is one of the top 10 leading causes of death worldwide. In 2016, 10.4 million people contracted the disease and 1.7 million died from it (0.4 million of whom also had HIV). More than 0.95 of TB deaths occur in low- and middle-income countries. Seven countries account for 0.64 of cases, led by India, followed by Indonesia, China, the Philippines, Nigeria, Pakistan, and South Africa. Over their lifetime, people infected with the tuberculosis bacillus have a 0.05 lifetime risk of developing the disease. In contrast, the risk is much higher for those with weakened immune systems, such as people living with HIV, malnourished or with diabetes.

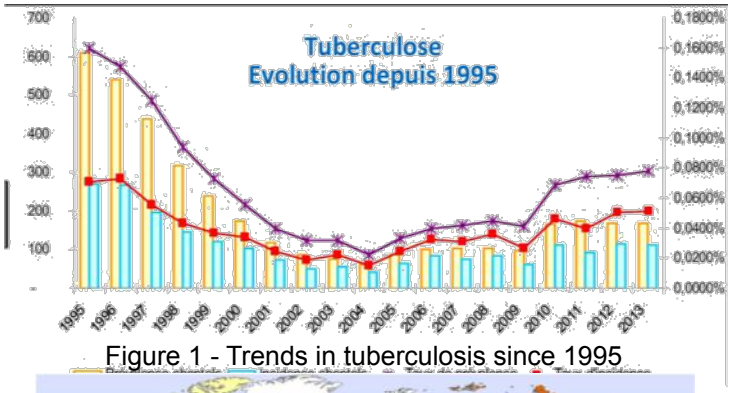

Figure 1 - Trends in tuberculosis since 1995

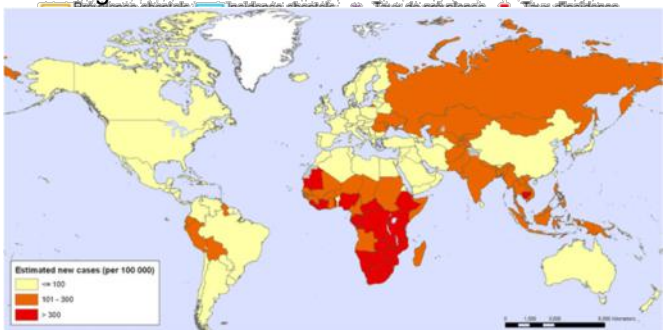

Figure 2 - Global distribution by country and prevalence of infection with My-cobacterium tuberculosis (MTB), the vector of tuberculosis. Source: WHO Report 2006

2 Models

Thus we have information to understand diseases and develop models to simulate their spread.

2.1 Tuberculosis

2.1.1 Assumptions

While observing the evolution of tuberculosis in the world, we noticed that it is interesting to look at the spread of tuberculosis in the countries: Algeria, Mauritania and Morocco. These three countries present high levels of infection different and important dynamics of interaction (immigration/emigration). We have therefore sought to model this phenomenon as faithfully as possible. Based on a thesis by Berge TSANOU on a study of several epidemiological models [ref], we have chosen a SEIR model (S:healthy, E:latently infected, I:infected, R=healed) for each country. In effet, a treatment for tuberculosis exists in these three countries and the number of people cured of the disease is not negligible. Even if the number of infected persons represents only 5% of the number of individuals exposed to the disease, it is still necessary to represent it since it is this category of persons who present worrying symptoms and who require interventions to avoid the spread of the disease.

Here is a diagram that models the different interactions between the different model compartments for a given country. These interactions, as in many epidemiological models, are expressed by fixed parameters. We therefore introduce several factors:

- Disease-independent parameters :
  - Births per unit of time .
  - Mortality rate .
- Disease-dependent parameters :
  - h: Proportion of people exposed to the disease who become infected
  - Exposure rate to the disease .
  - Probability that an infected person becomes a latent infected person .

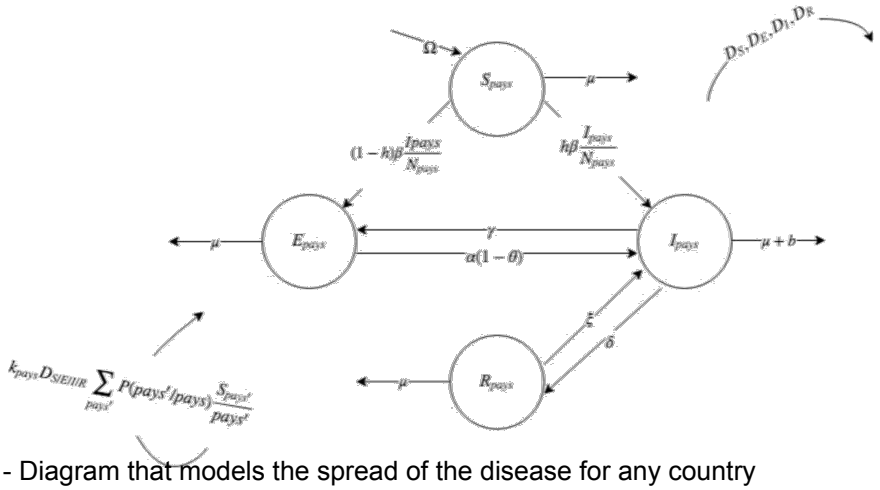

Figure 3 - Diagram that models the spread of the disease for any country

- Average length of time after which the disease in a latently infected person rises to the surface .
- : This parameter models the chemoprophylaxis. It represents the average duration of non-chemoprophylaxis that results in a fall from the latent infected to the infected.
- b: Mortality rate among infected persons (deaths caused only by tuberculosis)
- : Relapse rate from healed to infected people
- : Cure rate of the infected

- Immigration/Emigration Parameters :

- DS; DE; DI; DR: Emigration rates of healthy, latently infected, infected and cured individuals respectively.
- kpays: Parameter that represents the average number of individuals in a person's entourage.
- P(country'/country:) This is a probability that represents the relationship between countries. The probability is greater when the country: country' tends to accept immigrants from the country: country.

We can therefore write a system of differential equations describing the spread of tuberculosis in Algeria, Mauritania and Morocco. We decided to solve these equations with the programming language R. However, we first had to estimate values for all the parameters. For this, we mainly used [https :// www.who.int/tb/country/data/profiles/fr/].

In the following different simulations and results, we look at the spread of the disease between 2000 and 2016. The time scale is in years and the infected in millions.

2.1.2 Simulation and results

We first use R for to solve the equations and to maintain curves of the evolution of the number of infected people in the three countries. Then our algorithm tries different parameters and looks for those that create results that correspond as much as possible to the real data [https ://donnees.banquemondiale.org/indicator/ SH.TBS.INCD ?end=2016locations=DZ-MAstart=2000view=chart].

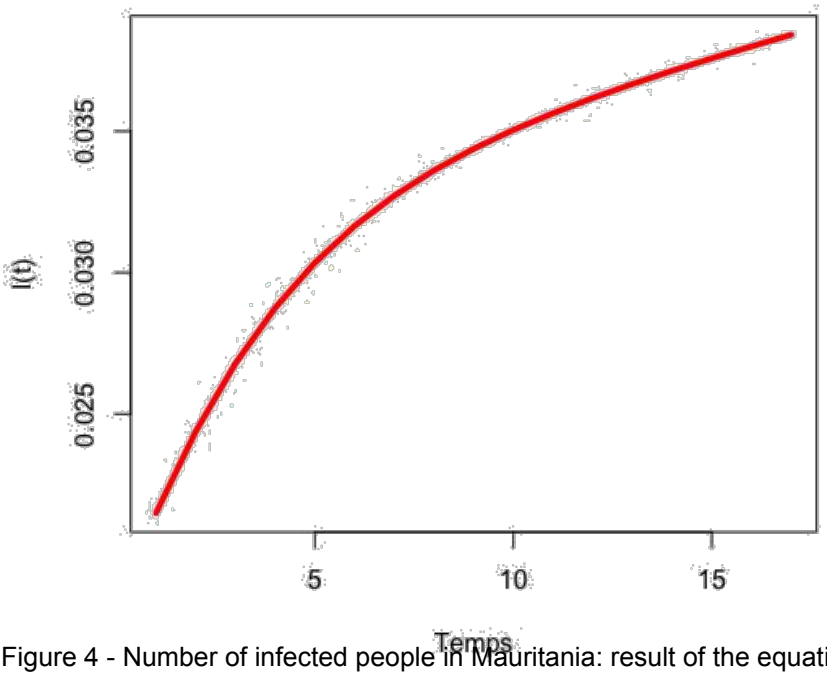

Figure 4 - Number of infected people in Mauritania: result of the equations.

The estimation of all parameters proved to be a failure. The algorithm does not converge. This can be explained in particular by the large number of parameters compared to the equations. As a solution, we decided to set a number of parameters, i.e. not to ask the algorithm to estimate them and thus to keep their first values. The new parameters to be estimated are :

h = 0:01; = 0:09; = 0:02; = 0:07; b = 0:07.

We notice that the estimated parameters are close to their starting values. Nevertheless, we have an error on the parameters which is very large. This problem may be due to the fact that we have set a few parameters and thus neglect their influence to some extent.

Then, in order to analyze the sensitivity of the model later, we first added noise to the data and then re-estimated the parameters so that the curves fit perfectly with the disturbed data. The goal is to find the same parameters as before to be sure that our model works even with data close to the real data. This time we have estimated the parameters for all three countries and not just Morocco. The curves obtained perfectly represent the expected behavior of the number of infected people in the different countries and the estimated parameters are the same as before.

Our model works very well despite the large errors in parameter estimation. It predicts that the rate of increase in the number of infected people in Mauritania decreases while the number of infected people in Morocco and Algeria increases, which is of course explained by immigration and its high rate from Mauritania to Morocco and Algeria.

2.1.3 Model sensitivity analysis

Using the code presented in the Mathematical Modeling for Biology course, we simulated the sensitivity of our model. In the graph below, we represent the influence of each parameter on the variation of the data. The first-order influence: expectation, and the second-order influence: variance, is represented.

On this graph, we observe that the parameter h has a great influence on the variation of the data with respect to the other parameters. The importance of h is obvious since it is the fraction of people exposed

à the disease that become infected. This parameter also greatly influences the calculation of R0, and thus decides

à how far the disease will spread. However, we believe that other parameters play a role.

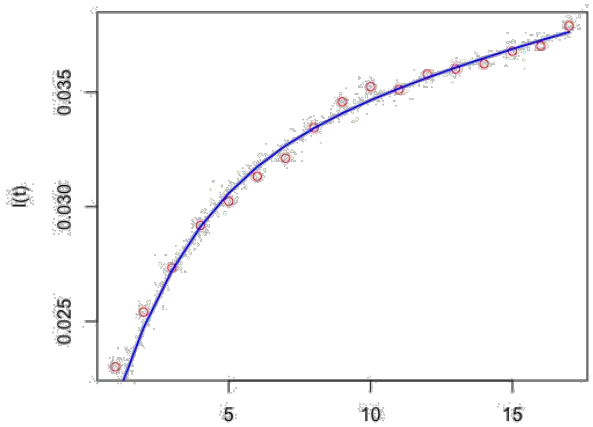

Figure 5 - Running the algorithm to estimate parameters that correspond to the actual data (number of infected in Mauritania)

Parameters:

	Estimate	Std. Error	t value	Pr(>|t|)
h	6.581e-02	1.893e+00	0.035	0.973
gamma	8.091e-02	2.519e+00	0.032	0.975
alpha	2.041e-02	1.897e+02	0.000	1.000
theta	5.592e-02	8.775e+03	0.000	1.000
b	7.765e-02	5.167e-01	0.150	0.883

Figure 6 - Estimated Parameters for Morocco

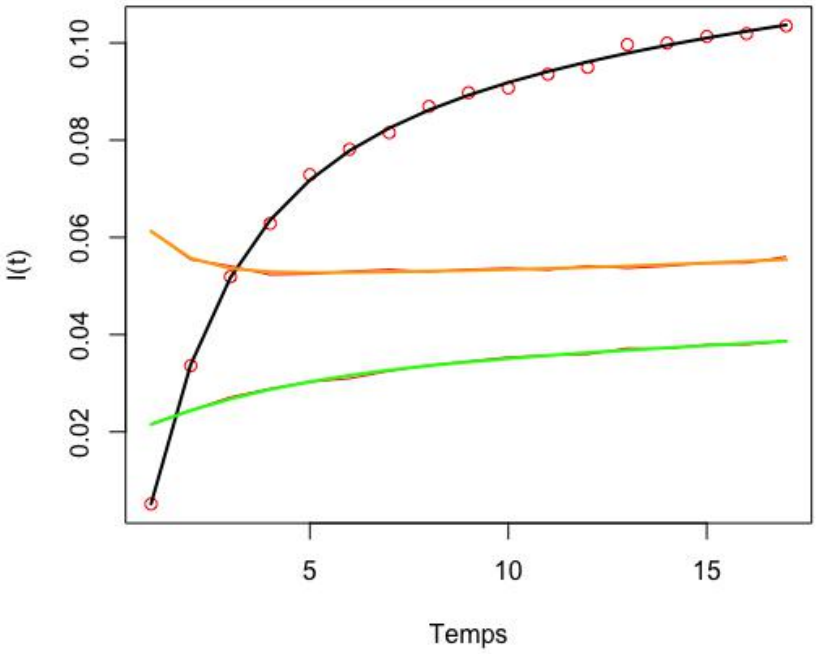

Figure 7 - Running the algorithm to estimate parameters that correspond to the actual data (number of infected in Algeria: green, Mauritania: black and Morocco: orange)

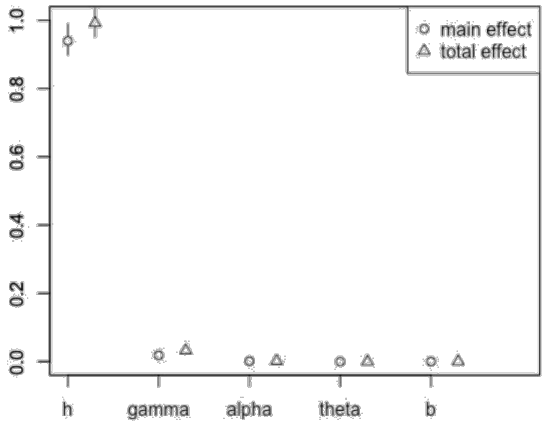

Figure 8 - Model Sensitivity Analysis

not negligible too. We formulate two hypotheses to explain the absence of any influence on the graph.

- The other parameters are small in front of h (apart from) and therefore since we have not standardized the data, it is possible that the scale on the graph is unsuitable.
- Setting the other parameters may have resulted in h being the only factor that decides the spread of the disease.

2.1.4 Opening

What we were going to do if we had more time, is to first try to reproduce the same method but by setting other parameters and find a "balance" between the parameters that will not cause any more problems (errors too big + sensitivity analysis that is not relevant enough). Then, we also want to try other types of algorithms for resolution (a genetic algorithm for example) or other types of models.