The final publication is available at

https://doi.org/10.1016/j.physa.2017.05.043

Additional Information

Probabilistic prediction of outbreaks of meningococcus W-135 infections over the next few years in Spain

L. Acedo *, C. Burgos, J.-C. Cortés, R.-J. Villanueva,

Instituto Universitario de Matemática Multidisciplinar,
Universitat Politècnica de València, Valencia (Spain)

Abstract

The genogroups of meningococcal and other bacteria are in competition in the ecosystem they form with the human hosts. Changes in vaccination strategies, prophylactic measures or usual habits, may also change the distribution of the genogroups in the ecosystem but, usually, this competition is ignored in most epidemiological models, despite it can be highly influential in the evolution of infection diseases and outbreaks.

Our goal is to propose a susceptible-carrier-susceptilbe (SCS) epidemiological model to determine the percentage of carriers in the population, and introduce a fractional Lokta-Volterra competition model to describe the evolution of the meningococcal genogroups in Spain among the carriers.

Using data from the distribution of the genogroups in Spain in 2011 and 2012, we find the model parameters and their uncertainties according to a probabilistic fitting approach. On this basis, we predict the evolution of the carriers of the different genogroups over the next few years and, in particular, the percentage of carriers of meningococcus W-135 with a 95% confidence interval. Then, we estimate the probability of having a possible outbreak of meningococcus W-135 in Spain over the next few years. According to our model and, under the present conditions, the risk of a serious outbreak of W-135 in Spain in the next 3 years is below 0.3%.

* Corresponding Author.
  Email addresses: `luiacrod@imm.upv.es` (L. Acedo),
`claburs i@posgrado.upv.es` (C. Burgos), `jccortes@imm.upv.es` (J.-C. Cortés),
`rjvillan@imm.upv.es` (R.-J. Villanueva).

1   Introduction

Meningococcal disease is caused by the bacterium Neisseria meningitidis, also called meningococcus. About 10% of people have this type of bacteria in the back of their nose and throat with no signs or symptoms of the disease, being called carriers. But sometimes Neisseria meningitidis bacteria can invade the body causing certain illnesses, which are known as meningococcal diseases [1].

The disease, Meningitis, is an infection of the brain and spinal cord and can even infect the blood. Nowadays, the main cause of Meningitis is the bacterium Neisseria meningitidis. This bacterium is transmitted exclusively among humans, mainly during adolescence. An individual may get infected by contact with a carrier individual, that is, healthy carriers transmit the bacteria. It is treated with specific antibiotics, however, even properly treated, there is up to 10% of mortality and 10% of survivors have sequelae [2,3].

On the one hand, there are 12 types or genogroups of Meningococcus based on the capsular polysaccharides: A, B, C, H, I, K, L, X, Y, Z, E and W, but 90% of all infections are caused by types A, B, C, Y and W, being A, B and C the most common (W-cases are only the 4% in US). On the other hand, the type A has been the most prevalent in Africa and Asia, but is rare in North America and Europe.

In Spain the major concern in the last decades has been the infection by the serogroups C and B. In particular, Neisseria meningitidis serogroup C (MenC) was the leading cause of meningococcal disease in Spain in the late 1990s. Men C mainly impacted on the infants and toddlers populations but also on adolescents and young adults, which also acted as carriers of the disease and spreaded the disease among other age groups. After controlling MenC by vaccination campaigns and recent revisions and booster doses, we can fairly say that MenC disease prevalence has greatly decreased in Spain [4]. On the other hand, Public Health responsables and Meningitis experts are now concerned about the fact that the ecological niche left by the waning serogroup could be occupied by another serogroup such as B or the even more lethal, W-135.

The serogroup W-135 origin has been traced back to Africa where it has competed with serogroup A and other in the so-called sub-Saharan Africa Meningitis belt. Since the early 2000s, the W-135 it is known to be associated with outbreaks of the Meningitis disease after the annual Hajj pilgrimage to Mecca [5]. The disease appeared also in Europe in 2000 and as early as April 2000 four children with W-135 Meningitis were treated in London. The situation became severe in Chile in 2012 when 133 cases of Invasive Meningococcal Disease were reported and the 58% of the cases the serogroup W-135 were identified. As a consequence, the fatality ratio peaked in Chile up to a 27%,

being the largest in the last twenty years [6]. Isolated clinical cases have also been reported in Spain since 2005, but no widespread epidemic by W-135 has occurred.

We must take into account that Neisseria meningitidis bacteria are spread through the exchange of respiratory and throat secretions like spit (e.g., living in close quarters, kissing, sharing drinks, etc). Fortunately, these bacteria are not as contagious as what causes the common cold or the flu. Besides, the bacteria are not spread by casual contact or by simply breathing the air where a person with meningococcal disease has been. Sometimes Neisseria meningitidis bacteria spreads to people who have had close or lengthy contact with a patient with meningococcal disease. People in the same household, roommates, or anyone with direct contact with a patient's oral secretions, meaning saliva or spit, such as a boyfriend or girlfriend, would be considered at increased risk of getting the infection [1].

Meningococcus is a part of the common flora in the nasopharynx of up to $5 - 15\%$ of adults and the genogroups are in competition by this ecosystem with humans. Thus, changes in health habits or prophylactic measures may change the distribution of the genogroups in this ecosystem. The study of the replacement of a genogroup by another as the most prevalent in the population is a fundamental problem in epidemiology. This is even more evident in the modern era of antibiotic resistant strains and vaccines selectively addressing a given genogroup and allowing other to proliferate.

For these reasons, in this paper we will study a Lotka-Volterra competition model coupled to a susceptible-carrier-susceptible (SCS) model for the transmission of the meningococcal bacteria. Standard Lotka-Volterra model simulates a predator-prey ecological system in which, the predator and prey populations interact and regulate each other. A different version of the model with a logistic term for the evolution of the free population of each genogroup is used for the interaction of different colonies, strains or genogroups competing for the same resources (or hosts, in the case of bacteria). Instead of the original model based upon ordinary differential equations, we propose also a generalization involving fractional derivatives. Fractional calculus is a tool to describe some physical systems exhibiting hysteresis and viscoelastic properties but also other behaviour such as subdiffusion [7–9]. The order of the fractional derivative is a parameter that could help us to estimate the effect of genetical changes in the strains and recombination which enhance the adaptation to the host of the dominant genogroups.

From the epidemiological and clinical point of view, there have been some recent interest in the W-135 specially after the outbreak in Chile [6], the subsequent outbreaks associated with the Hajj's pilgrims [5] and the vaccination against W-135 carried out in UK as a prevention after the initial cases of dis-

ease in which this strain was isolated [12]. Some models for the propagation of meningococcal C and B diseases [10,11], and also for the relation among prevalence of the infection and invasive disease [13] have also been studied recently. Pérez-Breva et al. [14] have also analyzed an agent-based model for the duration of immunity after vaccination of selected population age groups and the consequences of improved vaccination strategies against meningococcal C disease. Anyway, in the western countries, most of the cases are produced by groups B and C. Genogroups Y and W are less frequent although there are differences in their incidence in some countries. The genogroup W-135 is associated to cases and outbreaks after traveling to Mecca [5].

The increasing number of cases in countries such as Chile and UK and the high risk of mortality associated with this strain implies that the statistical modelling of this pandemic is a hot topic in epidemiology. For these reasons we will estimate the probability for an outbreak of meningococcal W-135 disease [15] in Spain using a probabilistic fitting technique for the genogroup competition SCS model discussed above. Our approach allows us to predict the maximum increase of the number of carriers expected from the seroepidemiological studies of 2011 and 2012 at the Reference Laboratory for Meningococci of the Carlos III Institute of Health in Spain [4]. A different, but still unsolved question, is the relationship among the number of carriers and the clinical cases of invasive meningococcal disease. According to the experts this connection is still uncertain but it is considered that even a 4% or 5% of the total population being carriers of the W-135 serogroup could be alarming or, at least, highly worrying. The probabilistic fitting technique would allow us to show that, from the 2011 and 2012 seroepidemiological results the maximum number of carriers of W-135 in Spain would remain below 3% in the next three years (excluding increasing immigration of infected individuals). Consequently, an outbreak of meningococcal W disease is unlikely in the near future.

The paper is organized as follows: In Section 2 we describe the Lotka-Volterra competition model for the genogroups coupled to the SCS model for the human hosts as well as the data provided by the Spanish Reference Laboratory for Meningococci at the Carlos III Institute of Health. The probabilistic fitting technique for obtaining the statistical distribution of the model parameters in order to perform probabilistic predictions for the future is described in Section 3. The results and the evolution of the estimated number of carriers in the next decade is shown in Section 4. The paper ends with some conclusions drawn in Section 5.

4

2 Lotka-Volterra competition model and SCS epidemilogical transmission

In this section we will introduce the competition model for the meningococcal genogroups as well as the susceptible-carrier-susceptible (SCS) model. These equations describe the complex ecosystem composed by all the meningococci genogroups and the human hosts. Firstly, we will give a summary of the data obtained in the seroepidemiological study for all meningococcal genogroups and, in particular, the W-135 strain.

2.1 Data

In this study, we use data provided by the Reference Laboratory for Meningococci of the Spanish Institute of Health Carlos III, collected in Tables 1 and 2, for December, 2011 and December, 2012. These are currently the only available data.

	$t_1 = 2011$	$t_2 = 2012$
Sample size	3000	500
Susceptible population	2626 (87.53%)	409 (81.8%)
Carrier population	374 (12.47%)	91 (18.2%)

Table 1
Sample size, number and percentage of susceptible carriers of any meningococcus in Spain in Dec 2011 and Dec 2012.

Dates	Genotype W	Other genotypes
$t_1 = 2011$	4.3%	95.7%
$t_2 = 2012$	5.5%	94.5%

Table 2
Percentage of carriers of the genotype Men W-135 in Spain in Dec 2011 and Dec 2012 among the carriers.

2.2 Susceptible-Carrier-Susceptible (SCS) model

Firstly, we state an epidemiological model in order to describe the transmission dynamics of all the meningococci. This is a classical SIS type-model where the "infected" people is called "carrier" in this case. Carriers are people who carry meningococci bacteria and can transmit it to other people. Only when meningococci invades the host, they get infected. As we indicated, the

5

meningococci spreads through the exchange of respiratory and throat secretions. The carriers clear the bacteria after some months becoming susceptible again. The above transmission dynamics of meningococci can be modeled by the following system of non-linear difference equations

$$\begin{cases} S_{t+1} = S_t - \beta S_t C_t + \gamma C_t, \\ C_{t+1} = C_t + \beta S_t C_t - \gamma C_t, \end{cases} \tag{1}$$

where S_t and C_t are the percentage of susceptible and carriers in the month t, respectively, β is the transmission rate of meningococci and $1/\gamma$ is the average time in moths a carrier individual clears the bacteria. A diagram of the model can be seen in Figure 1.
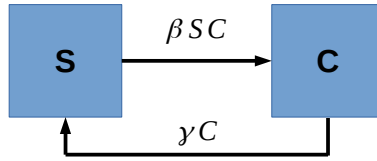

Fig. 1. Diagram of Susceptible-Carrier-Susceptible (SCS) model in Eq. (1).

Taking into account that S_t and C_t are percentages and $S_t + C_t = 1$ we can write the above system (1) with only the following equation

$$C_{t+1} = (1 + \beta - \gamma)C_t - \beta C_t^2 , \tag{2}$$

which is a non-linear difference equation for the percentage of carriers, C_t, at month t.

2.3 Lotka-Volterra's competition model

The above model will give us the percentage of carriers in every time instant (in months). But meningococci are in competition in their ecosystem. Therefore, in order to understand the competition dynamics of the meningococcus W-135 with respect to the others, we introduce the following Lotka-Volterra's continuous competition model of two species

$$\begin{cases} X_1'(t) = r_1 X_1(t)(K_1 - X_1(t)) - \alpha_{1,2} X_1(t) X_2(t), \\ X_2'(t) = r_2 X_2(t)(K_2 - X_2(t)) - \alpha_{2,1} X_2(t) X_1(t), \end{cases} \tag{3}$$

where $i = 1$ corresponds to W-135 genogroup and $i = 2$ corresponds to non-W-135 genogroup, and, for $i, j = 1, 2$, we have:

- $X_i(t)$ is the total amount of the genotype i meningoccocus bacteria at the time instant t,

6

- $r_i > 0$ is the growth rate of the genotype i meningoccocus bacteria,
- $K_i > 0$ is the carrying capacity of the genotype i meningoccocus bacteria,
- $\alpha_{i,j} > 0$ is the effect of the genotype j bacteria on the growth of the genotype i bacteria.

Notice that in our model we assign a label to the genogroup W-135 in which we are mostly interested, $i = 1$, but the collective composed by the other genogroups is treated as a single group with $i = 2$.

Taking into account that data collected in Table 2 are in percentages and, in order to combine the results of this model with those of the SCS model, we need scale the model (3). To this end, firstly, we introduce the following change of variables,

$$x_i(t) = \frac{X_i(t)}{K_i}, \qquad i = 1, 2, \tag{4}$$

which represents the percentage of ecosystem occupied by the i-th-genogroup. It is easy to check that $0 \leq x_i(t) \leq 1$ and $x_1(t) + x_2(t) = 1$.

Introducing the change of variable in Eq. (4) into Eq. (3), this latter system can be written as

$$\begin{cases} x_1'(t) = r_1 x_1(t) K_1 (1 - x_1(t)) - \alpha_{1,2} K_2 x_1(t) x_2(t), \\ x_2'(t) = r_2 x_2(t) K_2 (1 - x_2(t)) - \alpha_{2,1} K_1 x_1(t) x_2(t). \end{cases} \tag{5}$$

Finally, if we define $H_i = r_i K_i > 0$, $i = 1, 2$, $M_1 = \alpha_{1,2} K_2 > 0$ and $M_2 = \alpha_{2,1} K_1 > 0$, we have

$$\begin{cases} x_1'(t) = H_1 x_1(t)(1 - x_1(t)) - M_1 x_1(t) x_2(t), \\ x_2'(t) = H_2 x_2(t)(1 - x_2(t)) - M_2 x_1(t) x_2(t), \end{cases} \tag{6}$$

where H_i and M_i, $i = 1, 2$, are the new parameters of the system. As in the SCS model, taking into account that $x_1(t) + x_2(t) = 1$, we only consider the first equation of (6), related with the dynamics of the meningococcus W-135,

$$x_1'(t) = H_1 x_1(t)(1 - x_1(t)) - M_1 x_1(t)(1 - x_1(t)),$$

that can be simplified as

$$x_1'(t) = N_1 x_1(t)(1 - x_1(t)), \tag{7}$$

where $N_1 = H_1 - M_1$. So, the competition dynamics is basically governed by the Ricatti equation (7) in which the parameter N_1 is given by the difference among the reproductive factor for the W-135 genogroup, H_1, and the competition parameter with the other genogroups, M_1.

As we have mentioned in the introduction, the competition dynamics may reinforce certain genogroups by DNA recombination or mutations and this would depend on the other genogroups coexisting with them as well as the time this coexistence lasts and their populations. In this spirit, we will extend the original Lotka-Volterra model to a generalized one in which ordinary derivatives are replaced by fractional derivatives. It is well-known that fractional differential equations may account for these memory effects as they are usually applied to visco-elastic materials and subdiffusive processes [7,8].

So, we replace the first derivative of model (7) by a fractional Caputo derivative of order α $(0 < \alpha < 1)$ [19]. This way, the model will be given by the following expression:

$$^{C}D^{\alpha}x_1(t) = N_1 x_1(t)(1 - x_1(t)), \tag{8}$$

where $^{C}D_0^{\alpha}f$ denotes the fractional Caputo derivative of order α of the function $f : [0, \infty[\; \rightarrow \mathbb{R}$, which is given by

$$\left(\, ^{C}D_0^{\alpha}f \right)(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{f'(s)}{(t-s)^{\alpha}} \mathrm{d}s, \tag{9}$$

where $\Gamma(\alpha)$ is the classical gamma function defined by $\Gamma(\alpha) = \int_0^{\infty} e^{-t}t^{\alpha-1}\mathrm{d}t$. Discretizing Eq. (8), we obtain the following expression,

$$\left(\Delta_*^{\alpha}x_1 \right)(t) = N_1 x_1(t + \alpha - 1)(1 - x_1(t + \alpha - 1)) \, , \tag{10}$$

where $t \in \mathbb{N}_{1-\alpha} := \{1 - \alpha, 2 - \alpha, \dots\}$, and Δ_{α}^* is the Caputo-like delta difference operator which, according to Atici and Eloe [19], represents the Caputo derivative in discrete time. This difference is defined by the following expression

$$\left(\Delta_*^{\alpha}x_1 \right)(t) = \frac{1}{\Gamma(1-\alpha)} \sum_{s=0}^{t-n+\alpha} \frac{\Gamma(t-s)}{\Gamma(t-s+1)}(\Delta x_1)(s), \tag{11}$$

where $\left(\Delta x_1 \right)(t)$ is the discretization of the first derivative of $x_1(s)$ in discrete time, i.e.,

$$(\Delta x_1)(s) = x_1(s) - x_1(s - 1) \, . \tag{12}$$

Moreover, by [18, Lemma 2.4], the difference system (10) can be rewritten as

$$\begin{aligned}
x_1(t) &= x_1(0) + \frac{1}{\Gamma(\alpha)} \sum_{s=1-\alpha}^{t-\alpha} \frac{\Gamma(t-s)}{\Gamma(t-s-\alpha+1)} N_1 x_1(s+\alpha-1)(1 - x_1(s+\alpha-1)) \\
&= x_1(0) + \frac{1}{\Gamma(\alpha)} \sum_{s=1}^{t} \frac{\Gamma(t-k+\alpha)}{\Gamma(t-k+1)} N_1 x_1(k-1)(1 - x_1(k-1)),
\end{aligned}$$
$$\tag{13}$$

where t is a positive integer. Denoting $x_1(t) = x_t$, the fractional Lotka-Volterra

model in discrete time is formulated as

$$x_t = x_0 + \frac{1}{\Gamma(\alpha)} \sum_{s=1}^{t} \frac{\Gamma(t-k+\alpha)}{\Gamma(t-k+1)} N_1 x_{k-1}(1-x_{k-1}), \qquad (14)$$

where N_1 and α are the model parameters to be determined and t is the time in months. Notice that our competition-epidemiological model involves now four parameters: the infection rate, β; the recovery rate, γ; the balance among the W-135 genogroup reproduction and its competition parameter, N_1; and also the fractional index, α, for the Caputo derivative in Eq. (14). In the next section we will show how to obtain probabilistic estimations of these four parameters from the data collected in Tables 1 and 2. From these fitting we will be able to predict reliable bounds on the evolution of W-135 prevalence with 95% confidence and, consequently, to quantify the risk of outbreaks in the near future.

3   Probabilistic Fitting

This technique, introduced in [15], consists of using information from surveys to assign probability distributions to the data. Then, we sample data values from these probability distributions and fit the model to the sampled data. Thus, we find model parameters that fit not only the data but also the uncertainty contained into the intrinsic survey error. Hence, these fitted model parameters will allow the model to capture the data uncertainty (with 95% confidence intervals).

We are going to apply probabilistic fitting to models (2) and (14).

3.1   Data 95% confidence intervals (95% CI)

Data in Table 1 correspond to the mean percentage of carriers in Dec 2011 and Dec 2012. Also, the sample sizes are 3000 and 500, respectively.

Assuming that the surveys are independent, for each one of the two available surveys, let us denote by $\mathbf{X}^j = (X_1^j, X_2^j)$, $0 \leq X_i^j \leq n_j$, $i, j = 1, 2$, a random vector whose entries are $X_1^j = \#$ Carriers, $X_2^j = \#$ Susceptible, and $n_1 = 3000$ and $n_2 = 500$ being the sample sizes of the surveys corresponding to December 2011 ($j = 1$) and December 2012 (j=2), respectively.

For the sake of clarity, let us fix $j = 1$ and let us consider the random vector $\mathbf{X}^1$, but the same findings to be presented in the following, apply to the

9

random vector $\mathbf{X}^2$ ($j = 2$). Initially, it is natural to assume that the random variable X_1^1 has a binomial distribution of parameters $n_1 = 3000$ and p_1, $X_1^1 \sim \text{Bi}(3000; p_1)$. As a consequence, $X_2^1 = n - X_1^1 \sim \text{Bi}(3000, 1 - p_1)$, see Appendix A. However, as the available information from the data surveys is limited to December 2011 ($j = 1$) and December 2012 ($j = 2$) only, we are going to consider that the parameter p_1 is a random variable rather than a deterministic value. Furthermore, since $p_1 \in (0, 1)$ we will assume that p_1 has a beta distribution of parameters β_1^1 and β_2^1, $p_1 \sim \text{Be}(\beta_1^1; \beta_2^1)$. At this time, it is worth pointing out that beta distribution is a two parametric probabilistic distribution whose domain is just the interval $(0, 1)$, then allowing for enough flexibility to describe, from a probabilistic standpoint, the parameter p_1. This approach leads us to assume that X_1^1 has a beta-binomial distribution

$$X_1^1 \sim \text{Bi}(3000; p_1), \ \ p_1 \sim \text{Be}(\beta_1^1; \beta_2^1),$$

whose probability mass function is given by

$$\mathbb{P}[X_1^1 = k] = \binom{3000}{k} \frac{B(k + \beta_1^1; 3000 - k + \beta_2^1)}{B(\beta_1^1; \beta_2^1)}, \ \ k = 0, 1, \ldots, 3000, \quad (15)$$

and $B(a; b)$ is the standard beta function defined by $B(a; b) := \int_0^1 t^{a-1}(1 - t)^{b-1}\mathrm{d}t$.

As a consequence of the results collected in Appendix A for the second component of the random vector $\mathbf{X_1}$ one gets

$$X_2^1 = 3000 - X_1^1 \sim \text{Bi}(3000, 1 - p_1), \ \ 1 - p_1 \sim \text{Be}(\beta_2^1; \beta_1^1),$$

or equivalently, with $p_1 \sim \text{Be}(\beta_1^1; \beta_2^1)$. Hence, the probability mass function of X_2^1 is given by

$$\mathbb{P}[X_2^1 = k] = \binom{3000}{k} \frac{B(k + \beta_2^1; 3000 - k + \beta_1^1)}{B(\beta_1^1; \beta_2^1)}, \ \ k = 0, 1, \ldots, 3000, \quad (16)$$

where the symmetry property of the beta function $B(\beta_2^1; \beta_1^1) = B(\beta_1^1; \beta_2^1)$ has been used to express the denominator of (16).

An analogous development follows for the random vector $\mathbf{X}^2 = (X_1^2, X_2^2)$,

$$X_1^2 \sim \text{Bi}(500; p_2), \ \ X_2^2 \sim \text{Bi}(500; 1 - p_2), \ \ p_2 \sim \text{Be}(\beta_1^2; \beta_2^2),$$

$$\mathbb{P}[X_1^2 = k] = \binom{500}{k} \frac{B(k + \beta_1^2; 500 - k + \beta_2^2)}{B(\beta_1^2; \beta_2^2)}, \ \ k = 0, 1, \ldots, 500, \quad (17)$$

$$\mathbb{P}[X_2^2 = k] = \binom{500}{k} \frac{B(k + \beta_2^2; 500 - k + \beta_1^2)}{B(\beta_1^2; \beta_2^2)}, \ \ k = 0, 1, \ldots, 500. \quad (18)$$

In order to estimate the parameters $\{\beta_1^1, \beta_2^1\}$ of the random variable $p_1 \sim \text{Be}(\beta_1^1; \beta_2^1)$ (and hence of X_1^1 and X_2^1), firstly we observe, from Table 1 that in

10

December of 2011 ($j = 1$), the number of carriers was 374 out of 3000 individuals. Hence a reasonable punctual estimation of p_1 is $\frac{374}{3000}$. To provide estimates $(\hat{\beta}_1^1, \hat{\beta}_2^1)$ of (β_1^1, β_2^1), we will consider the synthetic sample $\{\frac{372}{3000}, \frac{373}{3000}, \frac{374}{3000}, \frac{375}{3000}, \frac{376}{3000}\}$ centered at $\frac{374}{3000}$ and then we apply the maximum likehood estimation (MLE) technique. This turns out

$$\hat{\beta}_1^1 = 6.122 \cdot 10^4 \quad \hat{\beta}_2^1 = 4.2984 \cdot 10^5. \tag{19}$$

These estimates satisfy the following desirable property

$$\mathbb{E}[p_1] = \frac{\beta_1^1}{\beta_1^1 + \beta_2^1} = \frac{6.122 \cdot 10^4}{6.122 \cdot 10^4 + 4.2984 \cdot 10^5} = \frac{374}{3000}.$$

Following the analogous strategy, we have constructed the estimates $(\hat{\beta}_1^2, \hat{\beta}_2^2)$ for the parameters (β_1^2, β_2^2) of the random variable $p_2 \sim \text{Be}(\beta_1^2; \beta_2^2)$ and hence of $(X_1^2$ and $X_2^2)$. Now, we have considered the sample $\{\frac{89}{500}, \frac{90}{500}, \frac{91}{500}, \frac{92}{500}, \frac{93}{500}\}$ centered at $\frac{91}{500}$. This figure comes from Table 1, since in December 2012 (j=2), the number of sample data being carrier was 91 out of 500 individuals. In this case, the application of MLE method turns out the estimates

$$\hat{\beta}_1^2 = 3.387 \cdot 10^3 \quad \hat{\beta}_2^2 = 1.5221 \cdot 10^4. \tag{20}$$

This estimates also satisfies $\mathbb{E}[p_2] = \hat{\beta}_1^2/(\hat{\beta}_1^2 + \hat{\beta}_2^2) = \frac{91}{500}$.

Now we compute the percentiles 2.5 and 97.5 in order to determine 95% CI (confidence intervals) of each random vector $\mathbf{X}^1 = (X_1^1, X_2^1)$ and $\mathbf{X}^2 = (X_1^2, X_2^2)$. These intervals are computed using the probability mass functions given by (15)–(18) and the estimates (19)–(20). The 95% CI are collected in Table 3.

Dates	Carriers (X_1^j)	Susceptible (X_2^j)
$t_1 = 2011$ ($j = 1$)	$[11.30\%, 13.67\%]$	$[86.33\%, 88.73\%]$
$t_2 = 2012$ ($j = 2$)	$[14.8\%, 21.60\%]$	$[78.20\%, 85.20\%]$

Table 3
95% CI of the data surveys using the probability mass function of Beta-Binomial distributions given by (15)–(18).

We apply the same method explained above but considering $Y_1^j = \#$ Carriers of Meningococcus W , $Y_2^j = \#$ Carriers of non-Meningococcus W and $n_1 = 3000$ and $n_2 = 500$. Therefore, we assume that

$$Y_1^j \sim \text{Bi}(n_j; p_j), \ p_j \sim \text{Be}(\gamma_1^j; \gamma_2^j), \quad j = 1, 2,$$

and consequently,

$$Y_2^j \sim \text{Bi}(n_j; 1 - p_j), \ 1 - pj \sim \text{Be}(\gamma_2^j; \gamma_1^j) \quad j = 1, 2.$$

Thus, their probability mass functions are given by

$$\mathbb{P}[Y_1^j = k] = \binom{n_j}{k} \frac{B(k + \gamma_1^j; n_j - k + \gamma_2^j)}{B(\gamma_1^j; \gamma_2^j)}, \quad k = 0, 1, \ldots, n_j, \quad j = 1, 2, \quad (21)$$

$$\mathbb{P}[Y_2^j = k] = \binom{n_j}{k} \frac{B(k + \gamma_2^j; n_j - k + \gamma_1^j)}{B(\gamma_1^j; \gamma_2^j)}, \quad k = 0, 1, \ldots, n_j, \quad j = 1, 2. \quad (22)$$

The parameters γ_i^j, $1 \leq i, j \leq 2$, are estimated using the MLE technique and we have obtained

$$\gamma_1^1 = 121.9, \quad \gamma_2^1 = 2727.7; \qquad \gamma_1^2 = 11.1872, \quad \gamma_2^2 = 192.4464.$$

In Table 4, we show the 95% CI for each the random variables $\{Y_1^1, Y_1^2, Y_2^1, Y_2^2\}$ computing the percentiles 2.5 and 97.5 of each random vector $\mathbf{Y}^1 = (Y_1^1, Y_2^1)$ and $\mathbf{Y}^2 = (Y_1^2, Y_2^2)$.

Dates	Genotype W (Y_1^j)	Genotype non-W (Y_2^j)
$t_1 = 2011$ ($j = 1$)	$[3.52\%, 4.97\%]$	$[95.03\%, 96.43\%]$
$t_2 = 2012$ ($j = 2$)	$[3.8\%, 7.7\%]$	$[92.30\%, 96.2\%]$

Table 4
95% CI of the data surveys using the probability mass function of Beta-Binomial distributions given by (21)–(22).

3.2 Probabilistic estimation

Let $\mathbb{M}_1(t; \beta, \gamma)$ and $\mathbb{M}_2(t; \alpha, N_1)$ denote short representations of model (2) and (14), respectively, where $\{\beta, \gamma\}$, and $\{\alpha, N_1\}$ are the models parameters of $\mathbb{M}_1$ and $\mathbb{M}_2$, respectively, and t is the time in months.

Also, we have the data (95% CI) collected in Tables 3 and 4 obtained by sampling the probability mass functions given by (15)–(18) and (21)–(22), respectively, and calculating the percentiles 2.5 and 97.5.

For the probability distributions given in Eq. (15)–(18), we take a sample $d_{t_{ij}}^*$, $i, j = 1, 2$ and we search for the values of the model parameters β^*, γ^* such that

$$\left\| \begin{bmatrix} \mathbb{M}_1(t_1; \beta^*, \gamma^*) \\ \mathbb{M}_1(t_2; \beta^*, \gamma^*) \end{bmatrix} - \begin{bmatrix} d_{t_11}^*, d_{t_12}^* \\ d_{t_21}^*, d_{t_22}^* \end{bmatrix} \right\|_2 \tag{23}$$

is as minimum as possible, being $\| \cdot \|_2$ the $2-$norm [20].

This procedure is a classic optimization problem that has been carried out through the Particle Swarm Optimization (PSO) algorithm [21], with the difference that we will not fit the data itself, but samples taken from the data's

probability distributions. We shall perform the fitting N times (N being a large number), storing both the parameter values β^*, γ^* as well as the calculated errors e^*, ordered from smallest to largest values. The result of this procedure is a list of model parameters fitted to a sample of the data with their corresponding errors, represented in Table 5.

Error	Parameters	Model
e_1^*	β^1, γ^1	$\mathbb{M}_1(t; \beta^1, \gamma^1)$
e_2^*	β^2, γ^2	$\mathbb{M}_1(t; \beta^2, \gamma^2)$
$\vdots$	$\vdots$	$\vdots$
e_N^*	β^N, γ^N	$\mathbb{M}_1(t; \beta^N, \gamma^N)$

Table 5
Model fitting to N samples of the data's probability distributions.

The value N should be a large number in order to capture as more data uncertainty as possible during the sampling process and this uncertainty could be fitted by the model. In this case, $N = 10\,000$.

Now, we take $\mathbb{M}_1(t; \beta^1, \gamma^1)$ and $\mathbb{M}_1(t; \beta^2, \gamma^2)$ in Table 5 and calculate the outputs for times $t_1 = 0$ (Dec 2011) and $t_2 = 11$ (Dec 2012), the time instants (months) where data are available. For each time instant, we shall calculate percentiles 2.5 and 97.5 for carriers. Hence, we will name m_2 the sum of:

- the $2-$norm of the difference between the percentiles 2.5 from the model output and from the data percentiles 2.5 in Table 3, and
- the $2-$norm of the difference between the percentiles 97.5 from the model output and from the data percentiles 97.5 in Table 3.

We will repeat the above process with the outputs from $\mathbb{M}_1(t; \beta^1, \gamma^1)$, $\mathbb{M}_1(t; \beta^2, \gamma^2)$ and $\mathbb{M}_1(t; \beta^3, \gamma^3)$, obtaining m_3, the measure between the confidence intervals from the outputs and the data. The same for m_4 and so on, until $\mathbb{M}_1(t; \beta^1, \gamma^1)$, $\mathbb{M}_1(t; \beta^2, \gamma^2)$, ..., $\mathbb{M}_1(t; \beta^N, \gamma^N)$, obtaining m_N as the measure between the confidence intervals from the outputs and the data.

Taking $m_k = \min\{m_1, \ldots, m_N\}$, we ensure that the 95% CI bands of the outputs from $\mathbb{M}_1(t; \beta^1, \gamma^1)$, $\mathbb{M}_1(t; \beta^2, \gamma^2)$, ..., $\mathbb{M}_1(t; \beta^k, \gamma^k)$ is the closest to the 95% confidence intervals from the data, with which our model will capture the maximum possible uncertainty of the data from the output of the models $\mathbb{M}_1(t; \beta^1, \gamma^1)$, $\mathbb{M}_1(t; \beta^2, \gamma^2)$, ..., $\mathbb{M}_1(t; \beta^k, \gamma^k)$.

This procedure should be repeated for model $\mathbb{M}_2(t; \alpha, N_1)$ by sampling the probability mass functions given by Eq. (21)–(22).

4 Results

For the SCS model $\mathbb{M}_1(t; \beta, \gamma)$, we obtain $m_k = 5211$. With the outputs of $\mathbb{M}_1(t; \beta^1, \gamma^1)$, ..., $\mathbb{M}_1(t; \beta^{5211}, \gamma^{5211})$ for $t = 0, 1, \ldots, 119$ (from Dec 2011 until Dec 2020), we can predict, probabilistically, the transmission dynamics of the carriers. Then, for each time instant t we take the model output and we calculate the 95% CI and the mean of the expectation. The results are shown in Figure 2.
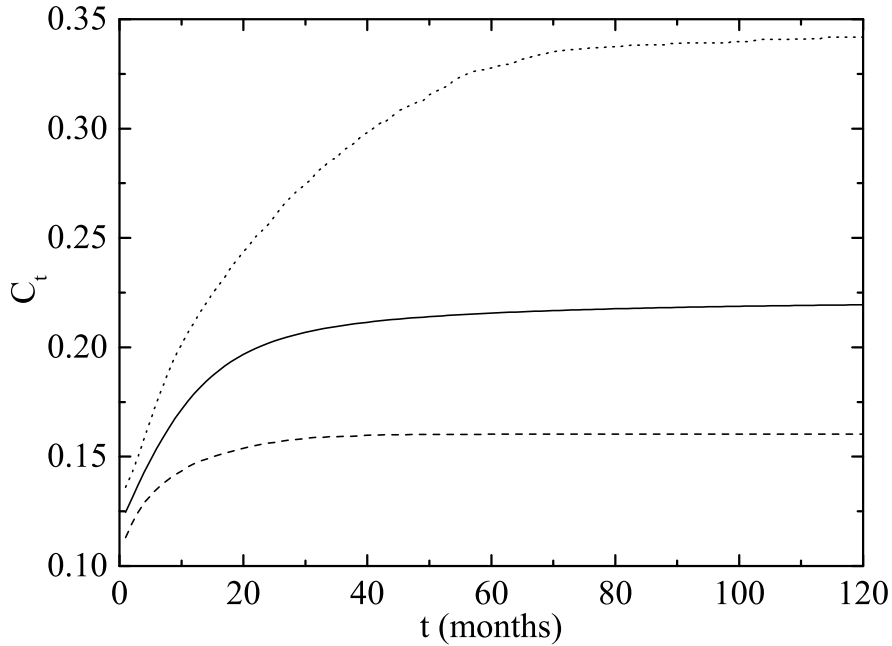

Fig. 2. 95% CI of the solution stochastic process of the percentage of carriers given by model (2). The dotted line represents the percentile 97.5, the dashed line, the percentile 2.5 and the solid line, the mean, for each month, t, from Dec 2011 until Dec 2021.

The probabilistic fitting technique allows us to estimate the density functions for model parameters applying the so-called kernel technique to the 5211 samples of parameters β, γ obtained [22, ch. 8]. These density functions can be seen in Figure 3.

Now, let us observe Figure 3. This plot gives us an idea about the time an individual needs to clear meningococcus bacteria looking at parameter γ. The 95% confidence interval of γ is $[0.1007, 0.7895]$, and taking into account that $1/\gamma$ is the average time to clear the meningococcus, we have that the estimated time needed to clear any meningococcus bacterium is $[1/0.7895, 1/0.1007] = [1.26662, 9.93049]$ months with 95% CI. These figures are in accordance with
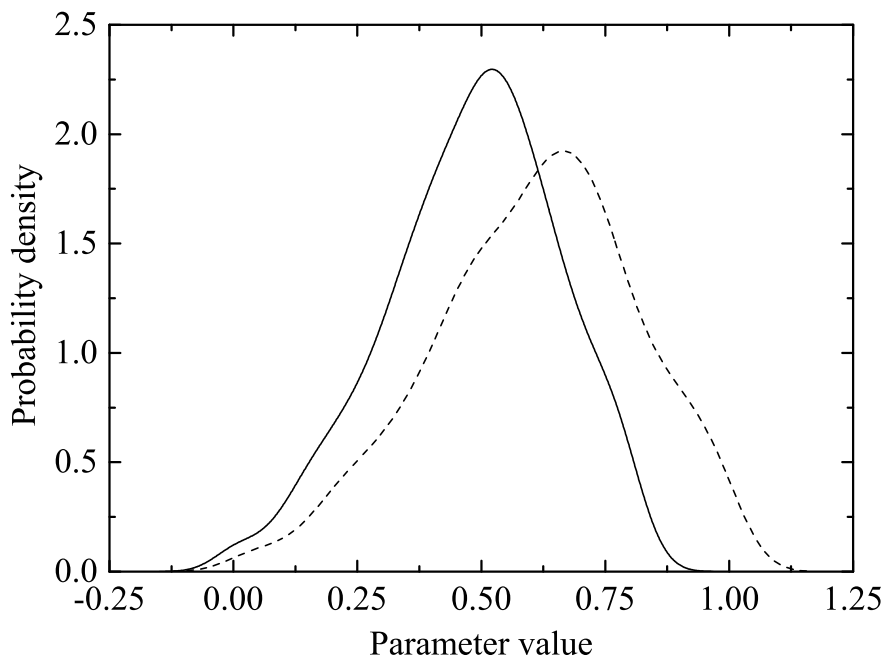
Fig. 3. Probability density functions for the parameters β (dashed line) and γ (solid line) of model (2). Notice that there are non-zero probabilities for negative values of these parameters, although in any reasonable epidemiological model they are considered positive. This feature is a consequence of the kernel technique and these negative unrealistic values can be overlooked.

the ones in the extant literature [10].

We can perform similar calculations for the Lotka-Volterra model $\mathbb{M}_2(t; \alpha, N_1)$. In this case, we have obtained $m_k = 651$ with $N = 10\,000$.

Notice that the order of the fractional derivative, α is far from 1 (that corresponds to the classical derivative) as we can see in Figure 5. Specifically, the mean of the pdf of α shown in this plot is 0.371326. Its 95% CI is $[0.004352, 0.7383]$. This supports the idea of the necessity of using fractional derivatives against classical derivatives in order to explain properly the competitiveness dynamics among the different genogroups of meningococci.

Finally, the product of both solution stochastic processes, C_t and x_t, of models (2) and (14), respectively, will give us the percentage of Men W-135 carriers among all the population. In a personal communication, Dr. Julio Vázquez, from the Neisserias Reference Laboratory for Meningococcus of the Spanish Institute of Health Carlos III, told us that a percentage of $4\% - 5\%$ or greater of carriers of Men W-135 in Spain would begin to be concerning.
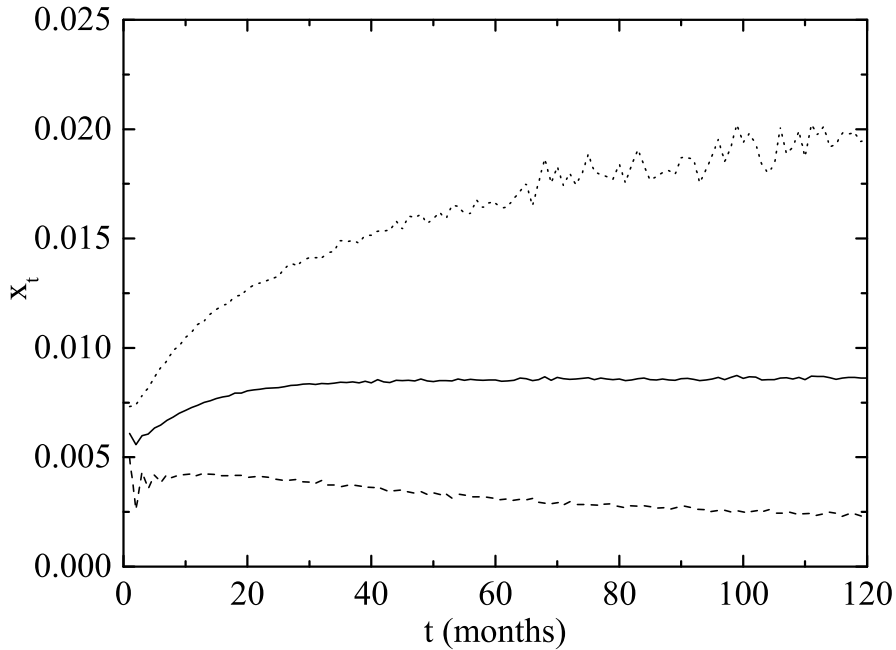
15

Fig. 4. Confidence intervals of the solution stochastic process for the percentage of meningococcus W-135 computed by the numerical scheme given in Eq. (14). The dotted line represents the percentile 97.5, the dashed line, the percentile 2.5 and the solid line, the mean, for each time, t, from Dec 2011 until Dec 2020.

Finally, taking into account the probabilistic information obtained previously for all time instants t, we are able to calculate the probability that the percentage of carriers of Men W-135 be greater than 5%. In Table 6 we can see the obtained results. In Figure 6 a graph with the evolution of the probabilities of having 5% of carriers of Men W-135 or more from Dec 2011 until Dec 2021.

with the combined

Date	Probability W-135 $\geq 5\%$
December 2017	2×10^{-4}
December 2018	0.0014
December 2019	0.0015
December 2020	0.0018

Table 6
Probability that the percentage of carriers of Men W-135 be greater than 5%.

As we can see, the probability of having worrying percentages of carriers of Men W-135 is very low and, then, an outbreak is not expected.
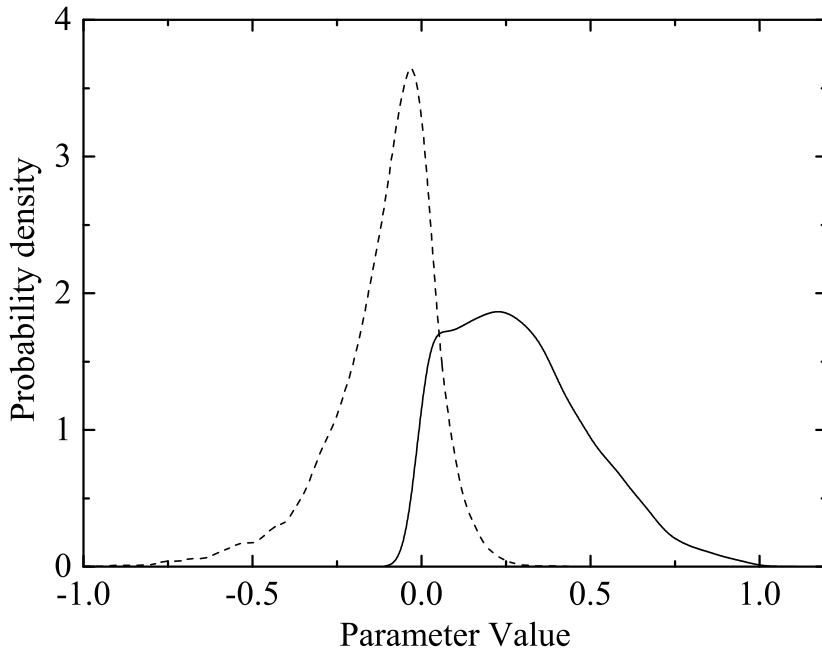
16

Fig. 5. Probability density functions (pdf's) for the parameters of the Lotka-Volterra model. The dashed line corresponds to the parameter N_1 and the solid line to the fractional derivative order, α. The short piece of negative values of the α parameter that appears in the plot, is a feature of the application of the so-called kernel technique used to construct the pdf's. This piece must been neglected since $0 < \alpha < 1$.

5   Conclusions

In this paper we have proposed a probabilistic model for the complex ecosystem composed by several meningococcal genogroups and the human hosts carriers. The competition among the different genogroups has been discussed in the framework of a Lotka-Volterra model with fractional derivatives. The Lotka-Volterra system is a well-known tool for simulating different population of the same species competing for the same resources and it has been applied in ecology and evolutionary models. Our model is generalized taking advantage of fractional calculus in order to consider the effects of the memory of this competition in the form of genetic changes of the strains.

The interaction with the human hosts is described in terms of a discrete susceptible-infectious-susceptible model in which the infectious are called carriers. This SCS model involves two parameters corresponding to the average time of recovery from the carrier state, $1/\gamma$, and the infective rate, β. We must also notice that meningococcal disease is very difficult to analyze through seroepidemiological studies because of the low carriage at a given time and the
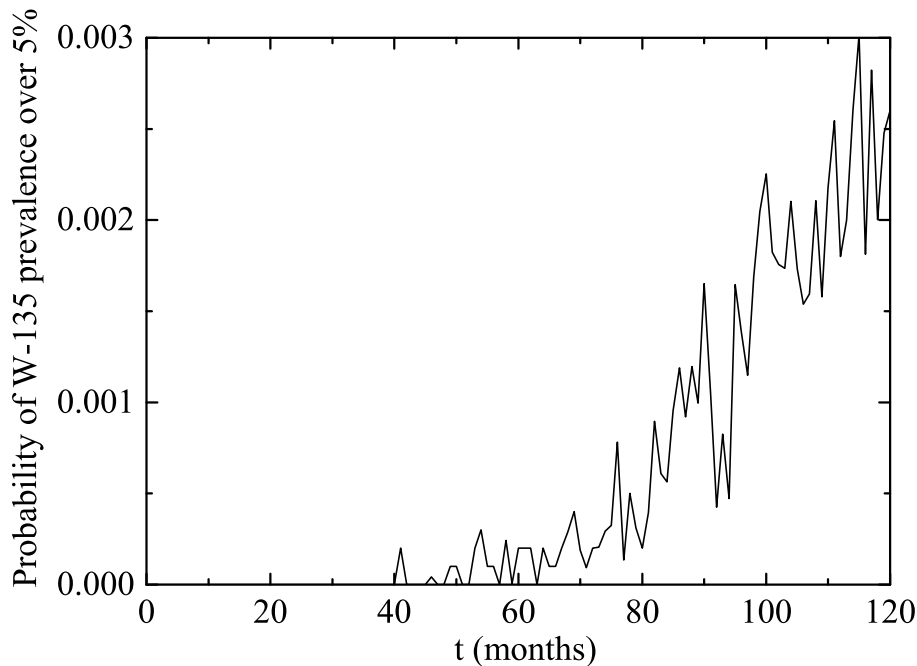
17

Fig. 6. Probability of having 5% of carriers of Men W-135 or more in Spain. Prediction from Dec 2011 until Dec 2021.

continuous mutation and recombination of DNA in the strains. For that reason, only a few mathematical models have been considered and, in particular, very little is known about the epidemiology of the W-135 genogroup which has caused recent outbreaks in Chile and the United Kingdom. Immigration from Africa is thought to be the main cause of the emergence of these strains in America and Europe where they were almost nonexistent or very rare even a decade ago. This is a reason of major concern for Public Health policies because of the high mortality rate for cases of invasive W-135 meningococcal disease.

In Spain there have been two seroepidemiological studies in 2011 and 2012 in which the W-135 genogroup was found in a small percentage of individuals. By fitting the results using a probabilistic approach, we have been able to determine the probability densities for the parameters of the Lotka-Volterra and SCS model. On the other hand, the prediction for the increase of W-135 in the near future shows that the probability for the carrier population to surpass the critical threshold of 5% is below a 0.3% and so we can fairly say that an outbreak of W-135 in Spain, under the present population and immigration conditions, is negligible. On the other hand, our model does not take into account an increase of W-135 carrier proportion in Africa, UK or South America and the corresponding risk of further contact of Spanish population with some of these carriers who could arrive to Spain.

Finally, we want to point out that, the effects of outbreaks in other countries in the present globalized world which takes into account immigration rates and the possibility of increased arrival of W-135 carriers as well as their contact with the native population is a problem that should be addressed in a future extended model to be discussed elsewhere.

Acknowledgements

Conflict of Interest Statement

The authors declare that there is no conflict of interests regarding the publication of this article.

References

[1] Centers for Disease Control and Prevention, `http://www.cdc.gov/meningococcal/index.html` (Accessed January, 12th, 2017).

[2] K. Cartwright, Meningococcal carriage and disease. In: Cartwright K., editor. Meningococcal disease. Chichester, UK: John Wiley & Sons, 1995, pp.71- 114.

[3] P. De Walls, Immunization strategies for the control of serogroup C meningococcal disease in developed countries, Expert. Rev. Vaccines 5 (2006) 269–75.

[4] Protocolos de la red nacional de vigilancia epidemiolgica (protocols for the national network for epidemiological surveillance), `http://www.isciii.es/ISCIII/es/contenidos/fd-servicios-cientifico-tecnicos/fd-vigilancias-alertas/fd-procedimientos/PROTOCOLOS_RENAVE-ciber.pdf`

[5] J. F. Aguilera, A. Perrocheau, C. Meffre, S. Hahné, and the W135 Working Group, Outbreak of Serogroup W135 Meningococcal Disease after the Hajj Pilgrimage, Europe, 2000, Emerg. Infect. Dis. 8(8) (2002) 761-777.

[6] M. T. Valenzuela, G. Moreno, A. Vaquero, M. Seoane, J. C. Hormazábal, M. P. Bertoglia, D. Gallegos, V. Sotomayor, J. Díaz, Emergence of W135 meningococcal serogroup in Chile during 2012, Rev. Méd. Chile 141(8) (2013) 959-967. `http://dx.doi.org/10.4067/S0034-98872013000800001`

[7] I. Podlubny, Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications, Academic Press, London, 1998.

[8] S. B. Yuste, L. Acedo, K. Lindenberg, Reaction front in an A+B→C reaction-subdiffusion process, Phys. Rev. E 69 (2004) 036126.

[9] S. B. Yuste, L. Acedo, An explicit finite difference method and a new von Neumann-type stability analysis for fractional diffusion equations, SIAM J. Numer. Anal, 42(5) (2005) 1862-1874.

[10] C. L. Trotter, N. J. Gay, W. J. Edmunds, Dynamic Models of Meningococcal Carriage, Disease, and the Impact of Serogroup C Conjugate Vaccination, Am. J. Epidemiol. 162(1) (2005) 89-100, doi:10.1093/aje/kwi160.

[11] H. Christensen, M. Hickman, W. J. Edmunds, C. L. Trotter, Introducing vaccination against serogroup B meningococcal disease: An economic and mathematical modelling study of potential impact, Vaccine 31(23) (2013) 2638–2646. `http://dx.doi.org/10.1016/j.vaccine.2013.03.034`.

[12] S. N. Ladhani, M. Ramsay, R. Borrow, A. Riordan, J. M. Watson, A. J. Pollard, Enter B and W: two new meningococcal vaccine programmes launched, Arch. Dis. Child. 101(1) (2016) 91-95. DOI: 10.1136/archdischild-2015-308928 PMID: 26672098

[13] G. Guzzetta, P. Manfredi, R. Gasparini, D. Panatto, W. J. Edmunds, On the relationship between meningococcal transmission dynamics and disease: Remarks on humoral immunity, Vaccine 27(2526) (2009) 3429–3434. `http://dx.doi.org/10.1016/j.vaccine.2009.01.092`.

[14] L. Pérez-Breva, R. J. Villanueva, J. Vilanueva-Oller, L. Acedo, F. Santonja, J. A. Moraño, R. Abad, J. A. Vázquez, J. Díez-Domingo, Optimizing strategies for meningococcal C disease vaccination in Valencia (Spain), BMC Infectious Diseases 14(280) (2014). doi: 10.1186/1471-2334-14-280.

[15] J. C. Cortés, F. J. Santonja, A. C. Tarazona, R. J. Villanueva, J. Villanueva-Oller, A probabilistic estimation and prediction technique for dynamic continuous social science models: The evolution of the attitude of the Basque Country population towards ETA as a case study, Applied Mathematics and Computation 264 (2015) 13–20. doi: 10.1016/j.amc.2015.03.128.

[16] D. Navarro, A. Perfors, An introduction to the Beta-Binomial model, `https://www.cs.cmu.edu/~10701/lecture/technote2_betabinomial.pdf` (Accessed January, 13th, 2017).

[17] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Section 6.1 Gamma Function, Beta Function, Factorials, Numerical Recipes: The Art of Scientific Computing (3rd ed.), New York: Cambridge University Press, 2007.

[18] F. Chen, X. Luo, Y. Zhou, Existence results for nonlinear fractional difference equation, Advances in Difference Equations 2011 ID713201, 1-12. `https://doi.org/10.1155/2011/713201`

[19] Atici, F. M., Eloe, P. W. (2009). Initial Value Problems in Discrete Fractional Calculus, 137(3), 981989.

[20] G .H. Golub, C. F. Van Loan, Matrix Computations, Johns Hopkins University Press, 4th Edition, 2012.

[21] C. Jacob, N. Khemka, Particle Swarm Optimization in *Mathematica*. An exploration kit for evolutionary optimization, IMS'04, Proc. Sixth International Mathematica Symposium, Banff, Canada, 2004.

[22] Dirk P. Kroese, Thomas Taimre, Zdravko I. Botev, Handbook of Monte Carlo Methods, Series in Probability and Statistics, Wiley, 2011.

[23] P. Bolt, J. Britto, S. Nadel, M. Levin, Meningococcal disease due to W135: fresh public health concerns, Arch. Dis. Child. 84 (2001) 89–93.

[24] B. García-Vila, P. Albert-De la Cruz, E. Bermejo-López, T. Grau-Carmona, Meningitis fulminante por Neisseria meningitidis serotipo W-135 y déficit de complemento, Med. Intensiva, 29(2) (2005) 106–109.

[25] A. Lamelas et al., Emergence of a New Epidemic Neisseria meningitidis Serogroup A Clone in the African Meningitis Belt: High-Resolution Picture of Genomic Changes That Mediate Immune Evasion, mBio 5(5) (2014) e01974–14

[26] A. J. Pollard, M. Santamaría, C. J. Maiden et al., W135 Meningococcal Disease in Africa, Emerging Infectious Diseases 9(11) (2003) 1503–1504.

[27] A. Papoulis, Random Variables and Stochastic Processes, 3rd Ed., McGraw-Hill, New York, 1991.

Appendix A

For the sake of completeness, in this appendix we show two properties related with the binomial and beta distributions that have been required throughout the manuscript.

A key property of the binomial distribution. Let X be a binomial random variable of parameters n and p, $X \sim \text{Bi}(n; p)$. Then, it results as the sum of n independent and identically distributed (iid) Bernoulli random variables of parameter p, $X_i \sim \text{Ber}(p)$, i.e.,

$$X = \sum_{i=1}^{n} X_i, \quad X_i = \begin{cases} 0 & \text{with probability} \quad 1-p, \\ 1 & \text{with probability} \quad p, \end{cases} \quad 1 \leq i \leq n.$$

Observe that $Y_i = 1 - X_i \sim \text{Ber}(1 - p)$, $1 \le i \le n$. Therefore,

$$Y = \sum_{i=1}^{n} Y_i = n - \sum_{i=1}^{n} X_i = n - X \sim \text{Bi}(n; 1 - p).$$

A key property of the beta distribution. Let X be a beta distribution of parameters $\beta_1 > 0$ and $\beta_2 > 0$, $X \sim \text{Be}(\beta_1; \beta_2)$. Then, its probability density function (pdf) is given by

$$f_X(x) = \frac{\Gamma(\beta_1 + \beta_2)}{\Gamma(\beta_1 + \beta_2)} x^{\beta_1 - 1} (1 - x)^{\beta_2 - 1}, \quad 0 < x < 1.$$

Let us consider the random variable $Y = 1 - X$. Applying the random variable transformation technique [27, ch.5], we can derive the pdf of Y, $f_Y(y)$. Observe that the mapping $r : [0, 1] \to [0, 1]$ defined by $y = r(x) = 1 - x$ is bijective, its inverse is $s(y) = 1 - y$ and its Jacobian is non-zero, $\frac{ds(y)}{dy} = -1 \neq 0$. Then, according to the so-called Fundamental Theorem [27, p. 93] one gets,

$$f_Y(y) = f_X(s(y)) \left| \frac{ds(y)}{dy} \right| = \frac{\Gamma(\beta_1 + \beta_2)}{\Gamma(\beta_1)\Gamma(\beta_2)} (1 - y)^{\beta_1 - 1} y^{\beta_2 - 1}, \; 0 < y < 1.$$

As a consequence, we have proved that if $X \sim \text{Be}(\beta_1; \beta_2)$, then $1 - X \sim (\beta_2; \beta_1)$.