Epidemic Progression and Vaccination in a Heterogeneous Population. Application to the Covid-19 epidemic

Vitaly Volpert[1,2,3], **Malay Banerjee**[4], **Swarnali Sharma**[5]

[1] Institut Camille Jordan, UMR 5208 CNRS, University Lyon 1, 69622 Villeurbanne, France
[2] INRIA Team Dracula, INRIA Lyon La Doua, 69603 Villeurbanne, France
[3] Peoples Friendship University of Russia (RUDN University)
6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation
[4] Department of Mathematics & Statistics, IIT Kanpur, Kanpur - 208016, India
[5] Department of Mathematics, Vijaygarh Jyotish Ray College, Kolkata - 700032, India

Abstract. The paper is devoted to a compartmental epidemiological model of infection progression in a heterogeneous population which consists of two groups with high disease transmission (HT) and low disease transmission (LT) potentials. Final size and duration of epidemic, the total and current maximal number of infected individuals are estimated depending on the structure of the population. It is shown that with the same basic reproduction number $\mathcal{R}_0$ in the beginning of epidemic, its further progression depends on the ratio between the two groups. Therefore, fitting the data in the beginning of epidemic and the determination of $\mathcal{R}_0$ are not sufficient to predict its long time behaviour. Available data on the Covid-19 epidemic allows the estimation of the proportion of the HT and LT groups. Estimated structure of the population is used for the investigation of the influence of vaccination on further epidemic development. The result of vaccination strongly depends on the proportion of vaccinated individuals between the two groups. Vaccination of the HT group acts to stop the epidemic and essentially decreases the total number of infected individuals at the end of epidemic and the current maximal number of infected individuals while vaccination of the LT group only acts to protect vaccinated individuals from further infection.

Key words: Covid-19, heterogeneous population, final size, duration of epidemic, vaccination

1 Introduction

Covid-19 epidemic has stimulated an unprecedented interest to the epidemiological models, mostly, compartmental ODE models. There are numerous recent works devoted to fitting the available data, calculating the basic reproduction number, and making predictions about the

further epidemic progression (see [1, 2, 4, 16, 23, 24, 27] and the references therein). These models give a good description of the evolution of the number of infected individuals and the sizes of other classes involved with the epidemiological models in the beginning of epidemic, and they take into account the influence of the measures of social distancing and some other measures to prevent the rapid epidemic spread. The situation is more complex with the prediction of the future epidemic progression because the parameters of the models are influenced by the measures of social distancing and other behavioral changes, and hence it is impossible to predict various scenario in advance.

At the end of the first year of the epidemic and during its second wave, sufficient amount of data are available to model the long time epidemic progression, including the development of collective immunity, the final size of epidemic, and the influence of vaccination on further epidemic growth profile. An important assumption here is that recovered and vaccinated individuals do not become susceptible any more. Though it is one of the most important open questions of the coronavirus disease, and immunological studies show that the quantity of antibodies in recovered individuals can be highly variable [25], we will adopt here this hypothesis.

The influence of heterogeneity of the population with respect to its role in the epidemic progression is largely discussed in the existing literature [3, 9, 17, 20, 22]. Different age and social groups can have different frequency of interactions and implementation of the measures of social distancing. Furthermore, the so-called superspreaders, a relatively small group of people with a large number of social interactions, play an important role in the coronavirus epidemic [13]. Consideration of two types of individuals, one having frequent social interaction and the other having restricted/cautious social interaction, together in a single group with an average infectivity can lead to erroneous predictions.

In this work we will study how the heterogeneity of the population influences its long time progression including the final size and duration of epidemic. In order to simplify the model and the interpretation of the results, we will consider only two groups in the population, one of them with high disease transmission (HT) and another one with low disease transmission (LT) potentials. If the total size of the population is N, the initial size of the first group (HT) is N_1 and the second group (LT) is N_2 such that $N = N_1 + N_2$, then we introduce the coefficient $k = N_1/N$ characterizing the structure of the population. If $k = 1$, then the whole population belongs to the first group, if $k = 0$, to the second group. In general, k adopts the values between 0 and 1. Two extreme values of k correspond to a single group epidemic.

The proportion between these two groups strongly influences the final size of epidemic (final number of susceptible individuals S_f) (cf. [6]) and the maximal current number of infected individuals I_m. This second parameter (I_m) is particularly important for the estimation of the necessary number of hospital beds to handle the worst case scenario. It appears that for the same basic reproduction number $\mathcal{R}_0$, the values S_f and I_m can differ several times depending on the parameter k. It is important to stress here that fitting the same data in the beginning of epidemic can be done for any value of k but further epidemic progression will crucially depend on it. Thus, the initial growth rate does not allow the

prediction of long time epidemic progression in a heterogeneous population.

We suggest a method to estimate the structure of the population in each given country during the Covid-19 epidemic on the basis of the available data before, during and after the first lockdown. Since there were no measures of social distancing (obligation of wearing masks, restriction on social gathering, etc.) before the lockdown, we assume that the whole population belonged to the first group (HT), and $k = 1$. On the other hand, during the lockdown, we assume that the whole population respected strict measures of social distancing and other restrictions, and $N_2 = N$, that is $k = 0$. In both cases, we fit the data and determine the parameters of the epidemiological model. After the lockdown when the restrictions are gradually lifted, the population splits into two groups, HT and LT, in certain proportion. We assign the first group the same parameters as before lockdown, and the second group the same parameters as during lockdown. Hence, we have all parameters characterizing each group, and one free parameter k which determines the proportion between the groups. The data after lockdown allow us to determine the value of this parameter and to characterize the heterogeneity of the population. Carrying out this analysis for several countries, we obtain $k \approx 0.1$ during the summer period and $k \approx 0.2$ in September-November 2020 [22]. This increase of the parameter k corresponds to the second wave of the epidemic. The rate of epidemic growth and the size of the second wave are determined by the value of k.

Clearly, suggested approach does not take into account the heterogeneity inside each group, possible exchange between the groups, and some other factors. However, it gives a single efficient parameter characterizing the structure of the population and the epidemic progression. We will call this parameter the coefficient of social interaction since $k = 0$ corresponds to low interaction during the lockdown and $k = 1$ to high interaction before the lockdown. Once the lockdown is relaxed, it is expected that people continue to follow some restrictions, either imposed by public authorities or self-imposed. However, different social and professional groups can have different levels of implementation of these restrictions and of the intensity of social contacts. As a result, the population splits into two groups, and the parameter k adopts some intermediate value between 0 and 1.

Having determined the structure of the heterogeneous population, we can study the influence of vaccination on the further epidemic progression. The results of the vaccination strongly depend on whether it is applied to HT group or to LT group. In particular, with only 5% of vaccinated individuals (of the whole population) for $k = 0.2$, the total number of infected individuals at the end of epidemic is almost 3 times less than without vaccination, if vaccination is applied to the HT group. If vaccination is applied to the LT group, the effect of vaccination is weak. Hence, vaccination of the first group acts to stop the epidemic while vaccination of the second group only protects vaccinated individuals. Though this result can be expected, the difference in the results of vaccination is quite striking.

The contents of the paper are as follows. In the next section, we introduce and study a model problem of a heterogeneous population. We determine various parameters of epidemic progression and show that with the same initial growth rate, its outcome can strongly differ depending on the structure of the population. In Section 3 we introduce a more complete epidemiological model of heterogeneous population. We apply it to Covid-19 in Section 4 in

order to determine the structure of the heterogeneous populations and to model the influence of vaccination on the epidemic progression. Finally, discussion of the model and of the result is presented in Section 5.

2 Model problem

We begin the study of epidemic progression in a heterogeneous population with the model problem consisting of susceptible and infected individuals with two sub-populations:

$$\frac{dS_1}{dt} = -\beta_{11}\frac{S_1}{N}I_1 - \beta_{12}\frac{S_1}{N}I_2, \tag{2.1}$$

$$\frac{dS_2}{dt} = -\beta_{21}\frac{S_2}{N}I_1 - \beta_{22}\frac{S_2}{N}I_2, \tag{2.2}$$

$$\frac{dI_1}{dt} = \beta_{11}\frac{S_1}{N}I_1 + \beta_{12}\frac{S_1}{N}I_2 - \sigma_1 I_1, \tag{2.3}$$

$$\frac{dI_2}{dt} = \beta_{21}\frac{S_2}{N}I_1 + \beta_{22}\frac{S_2}{N}I_2 - \sigma_2 I_2, \tag{2.4}$$

where β_{ij} are the rates of disease transmissions, σ_j are the clearance rates, and N is the total population. This model is similar to the model recently considered in [6]. We will present a more detailed analysis compared to the previous one. Along with basic reproduction number and the final size of epidemic, we will determine the maximal number of infected and will show that for the same value of basic reproduction number, the populations can strongly differ by their final size and the maximum of infected individuals. This effect occurs because of the heterogeneity of the population. Furthermore, we will study the influence of vaccination on the heterogeneous population.

2.1 Basic reproduction number

In the beginning of epidemic, S_1 and S_2 can be considered as constant. We set:

$$\frac{S_1}{N} = k\ , \quad \frac{S_2}{N} = 1-k, \quad 0 \le k \le 1.$$

The linearized matrix of the system (2.1) - (2.4) evaluated at disease free equilibrium point $(kN, (1-k)N, 0, 0)$ is given by

$$J = \begin{pmatrix} 0 & 0 & -k\beta_{11} & -k\beta_{12} \\ 0 & 0 & -(1-k)\beta_{21} & -(1-k)\beta_{22} \\ 0 & 0 & k\beta_{11}-\sigma_1 & k\beta_{12} \\ 0 & 0 & (1-k)\beta_{21} & (1-k)\beta_{22}-\sigma_2 \end{pmatrix}.$$

The non-zero eigenvalues of J can be obtained from the block matrix

$$A = \begin{pmatrix} k\beta_{11} - \sigma_1 & k\beta_{12} \\ (1-k)\beta_{21} & (1-k)\beta_{22} - \sigma_2 \end{pmatrix}.$$

We find maximal eigenvalues of J from the equation:

$$\lambda^2 - (k\beta_{11} + (1-k)\beta_{22} - \sigma_1 - \sigma_2)\lambda + k(1-k)(\beta_{11}\beta_{22} - \beta_{12}\beta_{21}) - k\beta_{11}\sigma_2 - (1-k)\beta_{22}\sigma_1 + \sigma_1\sigma_2 = 0.$$

In order to simplify the expression for basic reproduction number, we suppose that

$$\sigma_1 = \sigma_2 = \sigma,\ \beta_{12} = \beta_{21} = (\beta_{11} + \beta_{22})/2\ . \tag{2.5}$$

Then

$$2\lambda = k\beta_{11} + (1-k)\beta_{22} - 2\sigma + \sqrt{k\beta_{11}^2 + (1-k)\beta_{22}^2}\ . \tag{2.6}$$

The basic reproduction number $\mathcal{R}_0$ is as follows:

$$\mathcal{R}_0 = \left(k\beta_{11} + (1-k)\beta_{22} + \sqrt{k\beta_{11}^2 + (1-k)\beta_{22}^2} \right) / (2\sigma)\ .$$

If, moreover, $\beta_{11} = \beta_{22} = \beta$ for some β, then $\lambda = \beta - \sigma$, $\mathcal{R}_0 = \beta/\sigma$.

2.2 Final size of epidemic

Taking a sum of equations (2.1), (2.3) and (2.2), (2.4), we obtain the equalities:

$$\frac{dS_1}{dt} + \frac{dI_1}{dt} = -\sigma_1 I_1\ , \quad \frac{dS_2}{dt} + \frac{dI_2}{dt} = -\sigma_2 I_2\ . \tag{2.7}$$

Integrating them from 0 to ∞ and assuming that $I_j(0) = I_j(\infty) = 0$, $j = 1, 2$, we conclude that

$$S_1^0 - S_1^f = \sigma_1 \int_0^\infty I_1(t)dt\ , \quad S_2^0 - S_2^f = \sigma_2 \int_0^\infty I_2(t)dt\ .$$

Next, we divide equation (2.1) by S_1, equation (2.2) by S_2 and integrate from 0 to ∞:

$$-\ln\left(\frac{S_1^f}{S_1^0}\right) = \frac{\beta_{11}}{N\sigma_1}\left(S_1^0 - S_1^f\right) + \frac{\beta_{12}}{N\sigma_2}\left(S_2^0 - S_2^f\right)\ ,$$

$$-\ln\left(\frac{S_2^f}{S_2^0}\right) = \frac{\beta_{21}}{N\sigma_1}\left(S_1^0 - S_1^f\right) + \frac{\beta_{22}}{N\sigma_2}\left(S_2^0 - S_2^f\right)\ .$$

With the notation $x = S_1^f/S_1^0$, $y = S_2^f/S_2^0$, and assumptions (2.5) we obtain the system of two coupled transcendental equations

$$\beta_{11}k(1-x) + \beta_{12}(1-k)(1-y) = -\sigma \ln x, \tag{2.8}$$

$$\beta_{21}k(1-x)+\beta_{22}(1-k)(1-y)=-\sigma\ln y, \tag{2.9}$$

with respect to x and y. If $\beta_{11}=\beta_{22}=\beta$, then this system is reduced to the single equation $\mathcal{R}_0(1-x)+\ln x=0$ independent of k. Its solution gives the final size of susceptible population for the homogeneous population.

In the general case, the solution of this system depends on β_{11},β_{22}, and k. We will vary their values in such a way that the basic reproduction number does not change, and we will analyze the final size of epidemic.

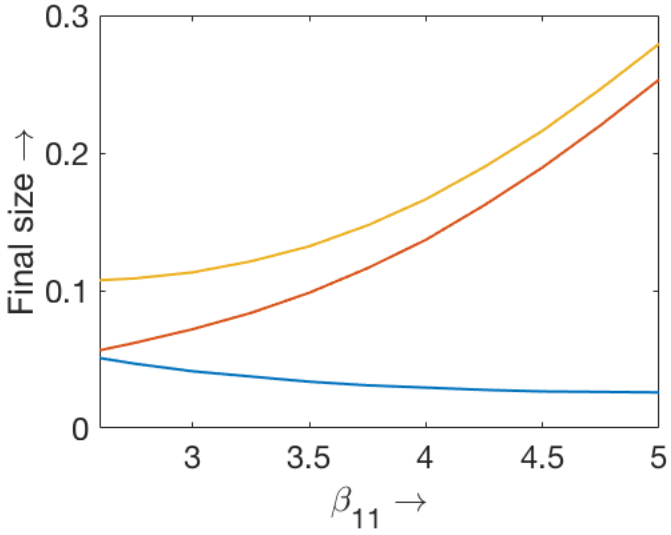

Figure 1: Final value of the total susceptible population and in the two sub-classes as functions of β_{11}, S_1^f - lower curve, S_2^f - middle curve, $S_1^f+S_2^f$ - upper curve. The values of parameters: $\beta_{22}=5-\beta_{11}$, $\beta_{12}=\beta_{21}=2.5$, $\sigma=1$.

Consider the following example: $\beta=2.5,\sigma=1$. Then $\lambda=1.5$, $\mathcal{R}_0=2.5$. For different values β_{11} and β_{22} such that $(\beta_{11}+\beta_{22})/2=\beta$, we find from (2.6) the value of k for which $\lambda=1.5$:

β_{11}	2.6	3	3.5	4	4.5	4.75	5
β_{12},β_{21}	2.5	2.5	2.5	2.5	2.5	2.5	2.5
β_{22}	2.4	2	1.5	1	0.5	0.25	0
k	0.495	0.475	0.45	0.427	0.403	0.393	3.82

For each of these combination of parameters, we determine x and y from system (2.8), (2.9), and the corresponding values S_1^f and S_2^f. The results of these calculations are shown in Figure 1. The final value of the first susceptible sub-population slowly decreases, for the second sub-population increases. The final value of the total susceptible population increases almost 3 times between the minimal value for the homogeneous population and the maximal value reached for $\beta_{22}=0$.

We consider another example where

$$\beta = 1.5,\ \sigma = 1,\ \lambda = 0.5,\ \mathcal{R}_0 = 1.5.$$

In this case the final size of susceptible for the homogeneous population is $S^f = 0.417$. The maximal total susceptible for the heterogeneous population is reached for $\beta_{11} = 3$ and equals $S^f = 0.587$. The ratio of the maximal and minimal value decreases with the decrease of $\mathcal{R}_0$ but their difference remains approximately the same as in the previous example.

2.3 Maximum number of infected

Maximal number of infected individuals for homogeneous population. If we assume that $\beta_{ij} = \beta$ for all $i, j = 1, 2$ for some β, and $\sigma_1 = \sigma_2$, then system (2.1)-(2.4) can be reduced to the system

$$\frac{dS}{dt} = -\beta \frac{S}{N} I\ , \quad \frac{dI}{dt} = \beta \frac{S}{N} I - \sigma I, \tag{2.10}$$

where $S = S_1 + S_2$, $I = I_1 + I_2$. Denote by t_m the time of maximum of $I(t)$, and by I_m the maximal value, $I_m = I(t_m)$, and $S_m = S(t_m)$. Integrating the sum of the equations in (2.10)

$$\frac{dS}{dt} + \frac{dI}{dt} = -\sigma I$$

from 0 to t_m, we get

$$S_m - S_0 + I_m - I_0 = -\sigma \int_0^{t_m} I(t)dt, \tag{2.11}$$

where $S_0 = S(0) = N$, $I_0 = I(0)$. Next, from the first equation in (2.10),

$$\ln(S_m) - \ln(S_0) = -\frac{\beta}{N} \int_0^{t_m} I(t)dt. \tag{2.12}$$

From equations (2.11), (2.12),

$$I_m - I_0 = S_0 - S_m + \frac{\sigma N}{\beta} \ln\left(\frac{S_m}{S_0}\right). \tag{2.13}$$

From the second equation in (2.10), since the derivative equals 0 at $t = t_m$, we find $S_m = \sigma N/\beta$. Using the notation $\mathcal{R}_0 = \beta/\sigma$, from (2.13) we find,

$$\frac{I_m - I_0}{N} = 1 - \frac{1}{\mathcal{R}_0} - \frac{\ln \mathcal{R}_0}{\mathcal{R}_0}\ , \quad S_m = \frac{N}{\mathcal{R}_0}\ . \tag{2.14}$$

If $\mathcal{R}_0 = 1$, then from the previous equations it follows that $I_m = I_0$, $S_m = N$. For $\mathcal{R}_0 > 1$, we get $I_m > I_0$ and $S_m < N$.

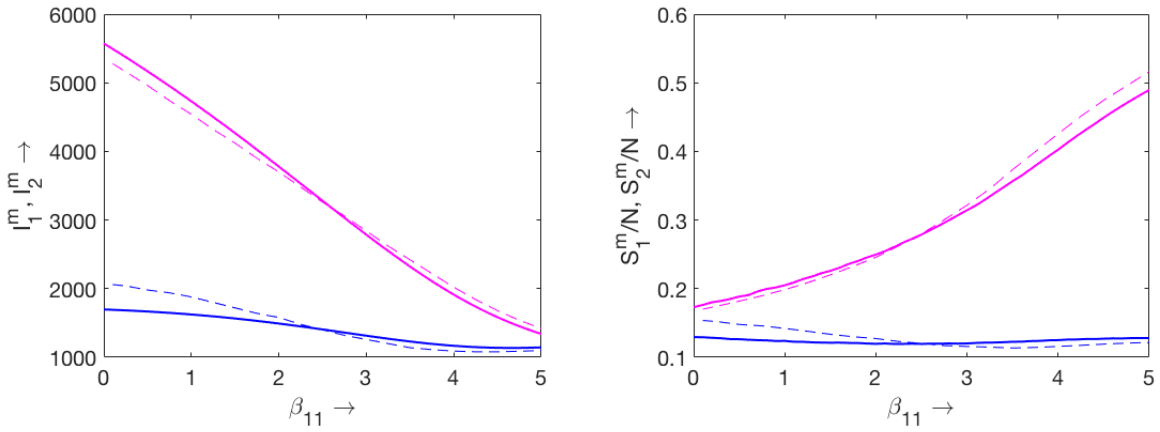

Figure 2: The maximal number of infected individuals (left figure) in direct numerical simulations of system (2.1)-(2.4) (solid lines) and as solution of system (2.19), (2.20) (dashed lines). The lower curves correspond to the first sub-population (I_1^m) and the upper curves to the second sub-population (I_2^m). The corresponding values of the number of susceptible individuals (right figure) in numerical simulations of system (2.1)-(2.4) (solid lines) and by formulas (2.18). The lower curves correspond to the fist sub-population (S_1^m) and the upper curves to the second sub-population (S_2^m). The values of parameters: $\beta_{22} = 5 - \beta_{11}$, $\beta_{12} = \beta_{21} = 2.5$, $\sigma = 0.1$. The value of k is chosen in such a way that the basic reproduction number $\mathcal{R}_0 = 2.5$ is the same in all simulations (see the explanation in the text).

Maximal number of infected individuals in a heterogeneous population. In order to find the maximal number of infected individuals in the heterogeneous population, we consider an approximation $t_m^1 = t_m^2$, where t_m^1 is the time of maximum of $I_1(t)$ and t_m^2 of $I_2(t)$. Numerical simulations show that these times to maximum are close to each other. Integrating equations (2.7) from 0 to t_m, we obtain:

$$S_1^0 - S_1^m + I_1^0 - I_1^m = \sigma_1 \int_0^{t_m} I_1(t)dt \ , \quad S_2^0 - S_2^m + I_2^0 - I_2^m = \sigma_2 \int_0^{t_m} I_2(t)dt \ , \tag{2.15}$$

where $S_j^0 = S_j(t_0)$, $S_j^m = S_j(t_m)$, $I_j^0 = I_j(t_0)$, $I_j^m = I_j(t_m)$, $j = 1, 2$. Next, we divide equation (2.1) by S_1, equation (2.2) by S_2 and integrate to get:

$$-\ln\left(\frac{S_1^m}{S_1^0}\right) = \frac{\beta_{11}}{N}\int_0^{t_m} I_1(t)dt + \frac{\beta_{12}}{N}\int_0^{t_m} I_2(t)dt \ ,$$

$$-\ln\left(\frac{S_2^m}{S_2^0}\right) = \frac{\beta_{21}}{N}\int_0^{t_m} I_1(t)dt + \frac{\beta_{22}}{N}\int_0^{t_m} I_2(t)dt \ .$$

Taking into account (2.15), we obtain

$$-\ln\left(\frac{S_1^m}{S_1^0}\right) = \frac{\beta_{11}}{N\sigma_1}(S_1^0 - S_1^m + I_1^0 - I_1^m) + \frac{\beta_{12}}{N\sigma_2}(S_2^0 - S_2^m + I_2^0 - I_2^m) \ , \tag{2.16}$$

$$-\ln\left(\frac{S_2^m}{S_2^0}\right) = \frac{\beta_{21}}{N\sigma_1}(S_1^0 - S_1^m + I_1^0 - I_1^m) + \frac{\beta_{22}}{N\sigma_2}(S_2^0 - S_2^m + I_2^0 - I_2^m)\ . \tag{2.17}$$

Assuming that $I_1'(t_m) = I_2'(t_m) = 0$, we get from (2.3), (2.4):

$$S_1^m = \sigma_1 N\frac{I_1^m}{\beta_{11}I_1^m + \beta_{12}I_2^m}\ , \quad S_2^m = \sigma_2 N\frac{I_2^m}{\beta_{21}I_1^m + \beta_{22}I_2^m}\ . \tag{2.18}$$

We suppose that $I_j^0 \ll I_j^m$, $j = 1, 2$ and $\sigma_1 = \sigma_2$. Set

$$x = I_1^m/N,\ \ y = I_2^m/N,\ \ \gamma_{ij} = \beta_{ij}/\sigma,\ \ i, j = 1, 2.$$

With this notation and (2.18), equations (2.16), (2.17) can be written in the following form:

$$\ln\left(k\frac{\gamma_{11}x + \gamma_{12}y}{x}\right) = (k - x)\gamma_{11} + (1 - k - y)\gamma_{12} - \frac{\gamma_{11}x}{\gamma_{11}x + \gamma_{12}y} - \frac{\gamma_{12}y}{\gamma_{21}x + \gamma_{22}y}\ , \tag{2.19}$$

$$\ln\left((1 - k)\frac{\gamma_{21}x + \gamma_{22}y}{y}\right) = (k - x)\gamma_{21} + (1 - k - y)\gamma_{22} - \frac{\gamma_{21}x}{\gamma_{11}x + \gamma_{12}y} - \frac{\gamma_{22}y}{\gamma_{21}x + \gamma_{22}y}\ . \tag{2.20}$$

Solving this system of equations, we find x and y and, consequently, I_j^m, $j = 1, 2$. We then use formulas (2.18) to determine S_j^m, $j = 1, 2$.

Figure 2 shows the comparison of the values I_j^m and S_j^m, $j = 1, 2$ obtained from direct numerical simulations and found by the approximate analytical method presented above. This approximation is more accurate for $\beta_{11} > \beta_{22}$ ($\beta_{11} + \beta_{22} = 5$) which corresponds to our main assumption that the first sub-population is smaller and spreads infection faster than the second sub-population. Let us recall that the population is homogeneous if $\beta_{11} = \beta_{22}$. The heterogeneity of the population increases with the increase of β_{11}. The maximal current number of infected individuals decreases with the increase of β_{11}. This effect is especially pronounced for the second sub-population.

2.4 Collective immunity

The notion of collective immunity implies that epidemic progression slows down due to the decrease of the number of susceptible individuals. Let us give a more precise definition for the homogeneous population implying that collective immunity begins at the moment of time when the number of infected individuals reaches its maximum, $t = t_m$, $I_m = I(t_m)$. The number of infected individuals begins to decrease after this time. The results of the previous section allow us to determine the exact value of I_m and S_m but not t_m. In order to find an approximate value of t_m we use the approximation $I(t) = I_0 e^{\lambda t}$ within the time interval $I_0 \le I(t) \le I_m$. Then

$$t_m \approx \frac{1}{\lambda}\ln\left(\frac{I_m}{I_0}\right) = \frac{1}{\beta - \sigma}\ln\left(1 + \frac{N}{I_0}\left(1 - \frac{1}{\mathcal{R}_0} - \frac{\ln\mathcal{R}_0}{\mathcal{R}_0}\right)\right)\ .$$

In the case of heterogeneous population, the total number of infected individuals $I(t) = I_1(t) + I_2(t)$ has a single maximum at some $t = t_m$ though the maxima of each component

$I_1(t)$ and $I_2(t)$ are reach at some close but different times t_m^1 and t_m^2, respectively. We consider that collective immunity begins at time $t = t_m$. In the approximate analytical solution considered above we assume that $t_m^1 = t_m^2$. Under this approximation, the time of the beginning of collective immunity can be determined similarly to the homogeneous population but with more cumbersome calculations related to solution of system (2.19) - (2.20).

2.5 Vaccination

The result of vaccination of a heterogeneous population essentially depends on the distribution of vaccinated individuals between different population groups. The epidemic is mainly spread by the first sub-population (HT), and their vaccination efficiently decreases the number of infected individuals.

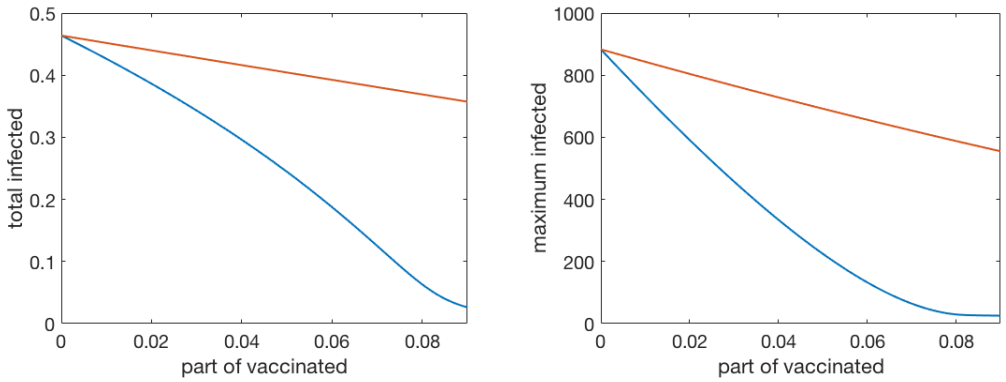

Figure 3: The total number of infected individuals (left figure) in numerical simulations of system (2.1)-(2.4) at the end of epidemic depending on the proportion of vaccinated individuals V to the total population N. The lower curve corresponds to the vaccination of the first sub-population ($\kappa = 1$) and the upper curve to the vaccination of the second sub-population ($\kappa = 0$) with the same total number of vaccinated individuals. The maximal current number of infected individuals (right figure) depending on the proportion of vaccinated individuals V to the total population N. The lower curve corresponds to the vaccination of the first sub-population ($\kappa = 1$) and the upper curve to the vaccination of the second sub-population ($\kappa = 0$) with the same total number of vaccinated individuals. The values of parameters: $\beta_{11} = 4, \beta_{22} = 1$, $\beta_{12} = \beta_{21} = 2.5$, $\sigma = 0.1$, $k = 0.1$, $t_0 = 5$.

We assume that vaccination is fully efficient in the sense that vaccinated individuals do not become infected. In order to model the action of vaccination at time $t = t_0$, we set

$$S_1(t_0{+}0) = S_1(t_0{-}0){-}\kappa V, \quad S_2(t_0{+}0) = S_2(t_0{-}0){-}(1{-}\kappa)V, \quad I_j(t_0{+}0) = I_j(t_0{-}0), \quad j = 1, 2,$$

where V is the number of vaccinated, κ is part of vaccinated in the first sub-population, $(1 - \kappa)$ in the second sub-population. System (2.1)-(2.4) is considered for $t > t_0$ with the indicated initial conditions at $t = t_0$.

Table 1: The model under consideration is (2.1) - (2.4) with vaccine in S_1.

k	%	I_1^m/N	I_2^m/N	S_1^f/N	S_2^f/N	t_f
0.2	–	0.0007	0.0016	0.1209	0.6187	846
	10%	0.00004	0.0009	0.1249	0.6645	1113
	20%	0.00001	0.0003	0.1286	0.7168	1715
	30%	0.000008	0.00002	0.1321	0.7770	4828
0.3	–	0.0031	0.0045	0.1220	0.4334	488
	10%	0.0020	0.0031	0.1294	0.4742	582
	20%	0.0011	0.0019	0.1371	0.5216	736
	30%	0.0004	0.0008	0.1450	0.5772	1084
	40%	0.00007	0.0001	0.1530	0.6437	2068
0.4	–	0.0068	0.0069	0.1199	0.3086	365
	10%	0.0047	0.0051	0.1302	0.3435	422
	20%	0.0029	0.0034	0.1414	0.3846	510
	30%	0.0015	0.0019	0.1535	0.4337	664
	40%	0.0005	0.0007	0.1665	0.4935	1012

Figure 3 (left) shows the total number of infected individuals I_T depending on vaccination. The total number of infected is calculated by the formula: $I_T = N - V - S_1^f - S_2^f$. In numerical simulations we set $S_j^f = S_j(t_f)$, $j = 1, 2$, where t_f is the final time of epidemic defined as time when the number of infected individuals becomes less than 1. Let us recall, that $I_j(t)$ converge to 0 as $t \to \infty$ but these functions remain positive for any finite time. Taking into account that these variables signify the number of individuals, the epidemic can be considered as finished at $t = t_f$ defined above.

We compare two cases, where all vaccinated belong to the first sub-population ($\kappa = 1$) or all of them belong to the second sub-population ($\kappa = 0$). In the first case, the influence of vaccination on the total number of infected individuals at the end of epidemic is essentially stronger than in the second case. If we take as example V equal 5% of the total population ($V/N = 0.05$), then I_T reduces from $0.45N$ (without vaccination) to $0.25N$, that is almost twice. At the same time, if only the second population is vaccinated, then the reduction is only 5%, that is the same as the number of vaccinated.

This difference becomes even more essential for the maximal current number of infected individuals (Figure 3, right). More detailed data on the results of vaccination are presented in Tables 1-3. These results depend on the values of parameters, in particular on the parameter k characterizing the proportion between the two sub-populations. However, the general tendency is the same as presented above: vaccination of the first sub-population is much more efficient. It is also interesting to note that vaccination increases the final time of epidemic. These tables show the results of numerical simulations of the results of vaccination in three cases: vaccination is applied only to the first sub-population (HT), only to the second sub-population (LT), and in proportion 20% HT and 80% LT. In all cases the number of

Table 2: The model under consideration is (2.1) - (2.4) with vaccine in S_2.

k	%	I_1^m/N	I_2^m/N	S_1^f/N	S_2^f/N	t_f
0.2	–	0.0007	0.0016	0.1209	0.6187	846
	10%	0.0006	0.0014	0.1259	0.6152	901
	20%	0.0005	0.0011	0.1310	0.6114	965
	30%	0.0004	0.0009	0.1363	0.6073	1040
0.3	–	0.0031	0.0045	0.1220	0.4334	488
	10%	0.0028	0.0039	0.1291	0.4265	511
	20%	0.0025	0.0033	0.1366	0.4190	536
	30%	0.0022	0.0028	0.1446	0.4109	565
	40%	0.0020	0.0024	0.1530	0.4020	599
0.4	–	0.0068	0.0069	0.1199	0.3086	365
	10%	0.0063	0.0059	0.1289	0.2986	380
	20%	0.0057	0.0050	0.1387	0.2875	397
	30%	0.0052	0.0041	0.1491	0.2753	416
	40%	0.0046	0.0034	0.1605	0.2619	439

vaccinated individuals is the same. It is given in percentage of N_1. The case $k = 0.2$ and 10% of N_1, for example, corresponds to 2% of the total population N. In all cases the vaccine is administrated on day 5.

3 Full model

3.1 Model of heterogeneous population

We consider conventional compartmental approach to model the epidemic progression with the following classes of population: susceptible individuals S, exposed (with viral load) but not yet infectious E_1, exposed infectious (no yet symptoms) E_2, infected symptomatic I_s, infected asymptomatic I_a, quarantined Q, hospitalized J, recovered R. Exposed infectious and infected asymptomatic individuals are in some sense similar to each other because they are infectious but do not manifest symptoms. However, the rate of disease transmission and the duration of these stages for them can be different.

The susceptible population can be heterogeneous with respect to a variety of characteristics: age classes, their activity including education, professional, retired, medical workers [10, 19, 21]. We will study how the heterogeneity of the population can influence the final size and duration of epidemic. For simplicity of presentation and analysis we restrict ourselves to two subclasses of susceptible individuals S_1 and S_2. According to this separation on subclasses, we introduce the corresponding subclasses in the groups E_1, E_2, I_s, and I_a, while Q, J, and R remain homogeneous. Under these assumptions, we obtain the following

Table 3: The model under consideration is (2.1) - (2.4) with vaccine in S_1 and S_2.

k	%	I_1^m/N	I_2^m/N	S_1^f/N	S_2^f/N	t_f
0.2	–	0.0007	0.0016	0.1209	0.6187	846
	10%	0.0006	0.0013	0.1257	0.6247	935
	20%	0.0004	0.0009	0.1307	0.6306	1050
	30%	0.0003	0.0007	0.1359	0.6366	1203
0.3	–	0.0031	0.0045	0.1220	0.4334	488
	10%	0.0025	0.0037	0.1292	0.4402	529
	20%	0.0020	0.0029	0.1369	0.4471	581
	30%	0.0016	0.0022	0.1451	0.4541	647
	40%	0.0012	0.0016	0.1538	0.4610	734
0.4	–	0.0068	0.0069	0.1199	0.3086	365
	10%	0.0056	0.0056	0.1295	0.3158	395
	20%	0.0045	0.0044	0.1399	0.3231	433
	30%	0.0035	0.0034	0.1512	0.3304	483
	40%	0.0025	0.0024	0.1636	0.3378	549

equations for S_1 and S_2:

$$\frac{dS_1}{dt} = -\frac{S_1}{N}\left[\beta_{11}(I_{s_1} + p_{11}I_{a_1} + p_{12}E_{12}) + \beta_{12}(I_{s_2} + p_{21}I_{a_2} + p_{22}E_{22}) + \beta_Q Q + \beta_J J\right], \quad (3.1)$$

$$\frac{dS_2}{dt} = -\frac{S_2}{N}\left[\beta_{21}(I_{s_1} + p_{31}I_{a_1} + p_{32}E_{12}) + \beta_{22}(I_{s_2} + p_{41}I_{a_2} + p_{42}E_{22}) + \beta_Q Q + \beta_J J\right]. \quad (3.2)$$

Here E_{12}, I_{s_1}, I_{a_1} and E_{22}, I_{s_2}, I_{a_2} are the subclasses of the corresponding classes E_2, I_s, I_a; β_{ij} are the coefficients characterizing the intensity of infection transmission between the classes S_i and I_{s_j}, $i, j = 1, 2$; the coefficients p_{ij}, $i = 1, 2, 3, 4$, $j = 1, 2$ show how the coefficients of infection propagation change for the classes I_{a_1}, I_{a_2} and E_{21}, E_{22} in comparison with I_{s_1} and I_{s_2}. Finally, the coefficients β_Q and β_J characterize infection progression due to the interaction with the classes Q and J.

The corresponding equations for the classes E_{11}, E_{21} have the following form:

$$\frac{dE_{11}}{dt} = \frac{S_1}{N}\left[\beta_{11}(I_{s_1} + p_{11}I_{a_1} + p_{12}E_{12}) + \beta_{12}(I_{s_2} + p_{21}I_{a_2} + p_{22}E_{22}) + \beta_Q Q + \beta_J J\right] - \mu_1 E_{11}, \quad (3.3)$$

$$\frac{dE_{21}}{dt} = \frac{S_2}{N}\left[\beta_{21}(I_{s_1} + p_{31}I_{a_1} + p_{32}E_{12}) + \beta_{22}(I_{s_2} + p_{41}I_{a_2} + p_{42}E_{22}) + \beta_Q Q + \beta_J J\right] - \mu_2 E_{21}, \quad (3.4)$$

where μ_1 and μ_2 are the rates at which E_{11} and E_{21} progress to the infectious exposed compartments E_{12} and E_{22}, respectively. Next,

$$\frac{dE_{12}}{dt} = \mu_1 E_{11} - \delta_1 E_{12}, \quad (3.5)$$

$$\frac{dE_{22}}{dt} = \mu_2 E_{21} - \delta_2 E_{22}. \tag{3.6}$$

The coefficients δ_1 and δ_2 characterize the transition from exposed to infected classes. The equations for the infected classes are as follows:

$$\frac{dI_{a_1}}{dt} = (1-\sigma_1)\delta_1 E_{12} - \eta_1 I_{a_1}, \tag{3.7}$$

$$\frac{dI_{a_2}}{dt} = (1-\sigma_2)\delta_2 E_{22} - \eta_2 I_{a_2}, \tag{3.8}$$

$$\frac{dI_{s_1}}{dt} = \sigma_1\delta_1 E_{12} - (\rho_{11} + \zeta_{11} + \zeta_{12} + \zeta_{13})I_{s_1}, \tag{3.9}$$

$$\frac{dI_{s_2}}{dt} = \sigma_2\delta_2 E_{22} - (\rho_{21} + \zeta_{21} + \zeta_{22} + \zeta_{23})I_{s_2}, \tag{3.10}$$

where σ_1 and σ_2 determine the the proportions between the classes of symptomatic and asymptomatic individuals, $0 < \sigma_j < 1$, $j = 1, 2$.

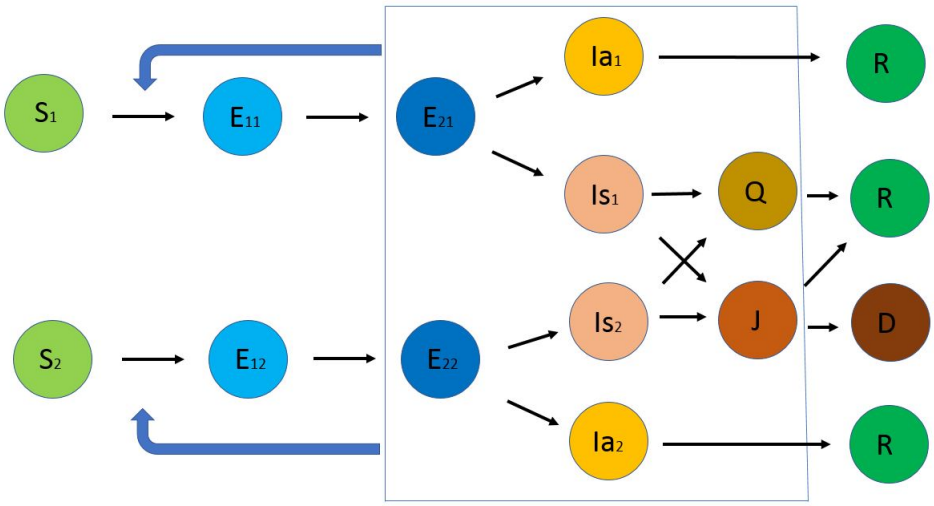

Figure 4: Schematic representation of the model with different classes of individuals. Two subclasses of susceptible give respectively exposed non-infectious, exposed infectious, infected symptomatic and asymptomatic. There are unique classes of quarantined, hospitalized, recovered, and dead.

Symptomatic infected individuals can become quarantined, hospitalized or recover, and asymptomatic infected recover without hospitalization. Symptomatic individuals move to the quarantine class, and the rate of their transfer to hospital and recovered class are ξ_1 and ξ_2, respectively. The equation for the quarantined class writes:

$$\frac{dQ}{dt} = \zeta_{11} I_{s_1} + \zeta_{21} I_{s_2} - (\xi_1 + \xi_2)Q. \tag{3.11}$$

Hospitalized individuals can recover or die with the rates ν and ρ_2:

$$\frac{dJ}{dt} = \zeta_{12} I_{s_1} + \zeta_{22} I_{s_2} + \xi_1 Q - (\rho_2 + \nu) J, \tag{3.12}$$

$$\frac{dR}{dt} = \eta_1 I_{a_1} + \eta_2 I_{a_2} + \zeta_{13} I_{s_1} + \zeta_{23} I_{s_2} + \xi_2 Q + \nu J. \tag{3.13}$$

This model was introduced in [22] where the basic reproduction number and the final size of epidemic were found. It was also used to fit the data on the Covid-19 epidemic in some countries before and during the lockdown and to determine the parameters by fitting the numerical simulation with the epidemiological data. Furthermore, most sensitive parameters were estimated. We will use this model and parameter values in the next section in order to study the influence of vaccination on epidemic progression, as illustrative example.

3.2 Characterization of epidemic progression

In this section we present the results of numerical simulations of system (3.1)-(3.13). Figure 5 shows the evolution in time of the two susceptible populations $S_1(t)$ (left figure) and $S_2(t)$ (right figure) for different values of parameter $k = N_1/N$. Larger values of k corresponds to the increase of the proportion of the first sub-population for which the disease transmission rate is more intensive. We observe from the figure that increasing k leads to the decrease of the final time of epidemic. The final size of the first sub-population remains approximately constant. This is due to the fact that epidemic is basically transmitted by the first sub-population, and it is finished when this sub-population reaches collective immunity. The final size of the second sub-population decreases with the increase of k (Figure 6, left) but it remains above the level of collective immunity. These final sizes can be found from the analytical formulas obtained in [22].

The maximal values of infected individuals increase with the increase of k (Figure 6, right). We consider here the sum of symptomatic and asymptomatic classes,

$$I_j(t) = I_{a_j}(t) + I_{s_j}(t)\ , \quad j = 1, 2.$$

Though the first sub-population $N_1 = kN$ is essentially less than the second one, $N_2 = (1-k)N$ for small k, the maximal values of infected individuals, I_m^1 and I_m^2 are close to each other because disease transmission occurs faster in the first sub-population.

4 Application to the Covid-19 epidemic

4.1 Coefficient of social interaction

The heterogeneity of the population with respect to the disease transmission is related to multiple factors. Among them different age, professional and social groups, various religious

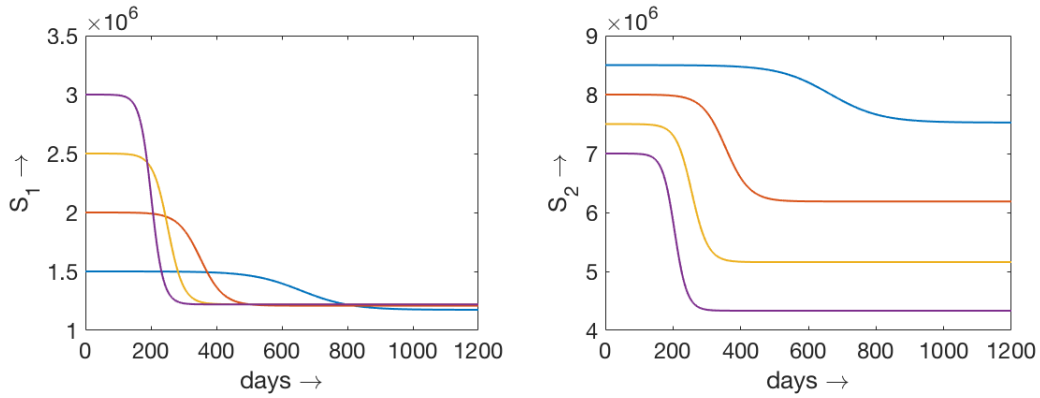

Figure 5: Numerical simulations of system (3.1)-(3.13). The evolution of the sub-populations $S_1(t)$ (left) and $S_2(t)$ in time for different values of k. The values of the coefficients β_{ij} are as follows: $\beta_{11} = 4, \beta_{22} = 1, \beta_{12} = \beta_{21} = 2.5$. The values of other parameters are given in Table 4 in the appendix.

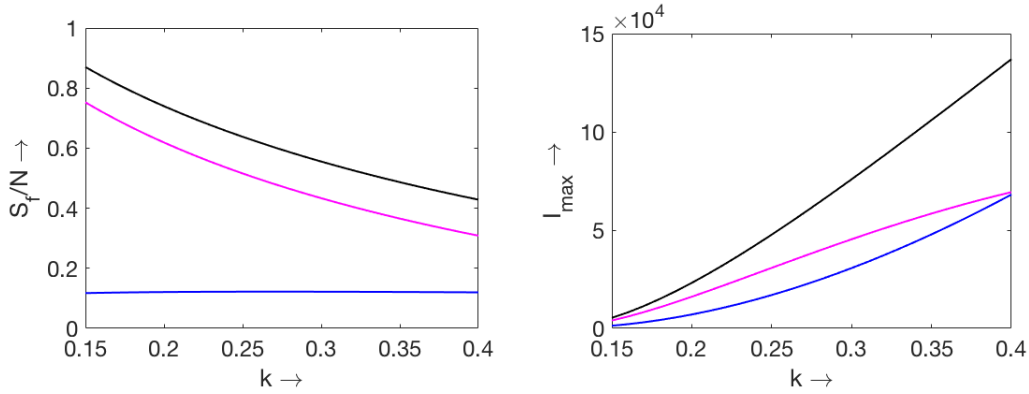

Figure 6: The final size of susceptible classes for different values of k (left). The upper curve shows the total number of susceptible, the middle curve corresponds to $S_2(t)$ and the lower curve to $S_1(t)$. The maximal number of infected individuals for different values of k (right). The upper curve shows the total number of infected (symptomatic plus asymptomatic), the middle curve corresponds to the second sub-population and the lower curve to first sub-population. The values of the coefficients β_{ij} are as follows: $\beta_{11} = 4, \beta_{22} = 1, \beta_{12} = \beta_{21} = 2.5$. The values of other parameters are given in Table 4 in the appendix.

and cultural traditions which can influence people behavior with respect to the measures of social distancing and vaccination. Detailed description of all these different groups would essentially complicate the model and would increase the number of parameters difficult to estimate. Therefore, we propose to consider only two cumulative groups. One of them includes people with high disease transmitting potential (HT) and another one with low disease transmitting potential (LT). According to the models we consider here, these two classes are represented by S_1 and S_2 respectively. They differ by the values of parameter

β_{ij} in the expression $\beta_{ij} I_i S_j/N$, $i, j = 1, 2$. These are effective parameters characterizing the frequency of contacts between infected and susceptible individuals and the rate of infection transmission. For the two groups N_1 (HT) and N_2 (LT) of the whole population N, there are the corresponding subclasses of susceptible S_1 and S_2, exposed and infected. Hence, there are four different coefficients: β_{11} characterizes the interaction inside HT, β_{22} inside LT, and β_{12}, β_{21} between the groups.

We consider the parameter $k = N_1/N$ already used in the previous sections. For $k = 0$, there is a single group LT, for $k = 1$ another single group HT. We will consider the values of k between 0 and 1. This parameter characterizes the distribution of total population into two groups and influences the intensity of social interactions.

We determine the values of parameters β_{ij} and k from the Covid-19 data. In the beginning of the epidemic, before lockdown, there were no measures of social distancing. We suppose that the whole population belonged to the first group (HT). We neglect here the heterogeneity of the population with respect to the frequency of contacts. Fitting the data on the epidemic progression allows us to determine the coefficient β_{11}. Its value can be different in different countries. Next, we suppose that during the first lockdown the whole population respected the measures of social distancing and belonged to the second group (LT). As before, fitting the data allows us to determine β_{22}. In the data fitting we used the model presented in Section 3 [22]. The values of the other parameters were determined from the available data.

After the first lockdown, the measures of social distancing were partially preserved. These restrictions differed between the countries and evolved in time. They were less strict than during the first lockdown allowing the emergence of two cumulative groups N_1 and N_2 described above. Simplifying this characterization of the population, we suppose that the first group (HT) is similar to the population before lockdown, and it is characterized by the coefficient β_{11} described above. The second group (LT) is similar to the population during the lockdown, and it is characterized by the coefficient β_{22}. We set the values of the coefficients β_{12} and β_{21} characterizing the interactions between the groups according to the formula $\beta_{12} = \beta_{21} = (\beta_{11} + \beta_{22})/2$. It is an empiric relation which cannot be determined from the data. We will discuss it below.

Next, we determine the value of the coefficient k fitting the data after lockdown. Figure 7 shows consecutive stages of epidemic progression in Germany with the first stage (before lockdown), second stage (during lockdown), third stage (June-July, 2020), and the fourth stage (September-November, 2020). According to the method described above, we get $\beta_{11} = 3.95$, $\beta_{22} = 1.05$, $\beta_{12} = \beta_{21} = 2.5$ (the values of other parameters are given in Table 4 in the appendix). Fitting the data in June-July, we find $k = 0.1$, that is the first group (HT) represents 10% of the whole population. We then continued the simulation for the period September-November with two different values of k: the same as before, $k = 0.1$ (left figure), and $k = 0.2$ (right figure). Increase of k shows a rapid growth of the number of infected. Let us note that this simulation was done in July, 2020 [22], and it gave a reasonably good prediction of the epidemic progression during the fourth stage. Increase of the coefficient of social interaction k during the fourth stage is related to the beginning of the academic year and the intensification of professional activity after summer vacation.

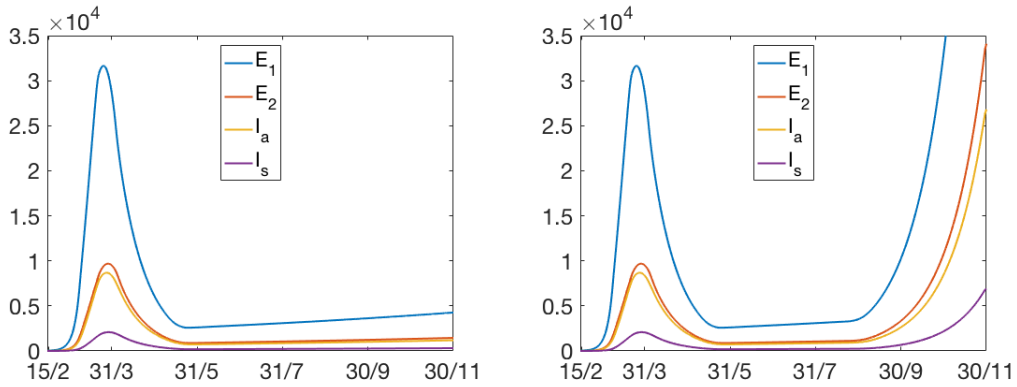

Figure 7: Numerical simulations of epidemic progression in Germany with system (3.1)-(3.13). The values of the coefficients β_{ij} are as follows: $\beta_{11} = 3.95, \beta_{22} = 1.05, \beta_{12} = \beta_{21} = 2.5$ (see the explanation in the text). The values of other parameters are given in the appendix. The values of k: 1 in February-March 2020 (before lockdown), 0 in April-May (lockdown), 0.1 in June-August (after lockdown), 0.1 in September-November (left) and 0.2 in September-November (right). Reprinted from [22] with permission.

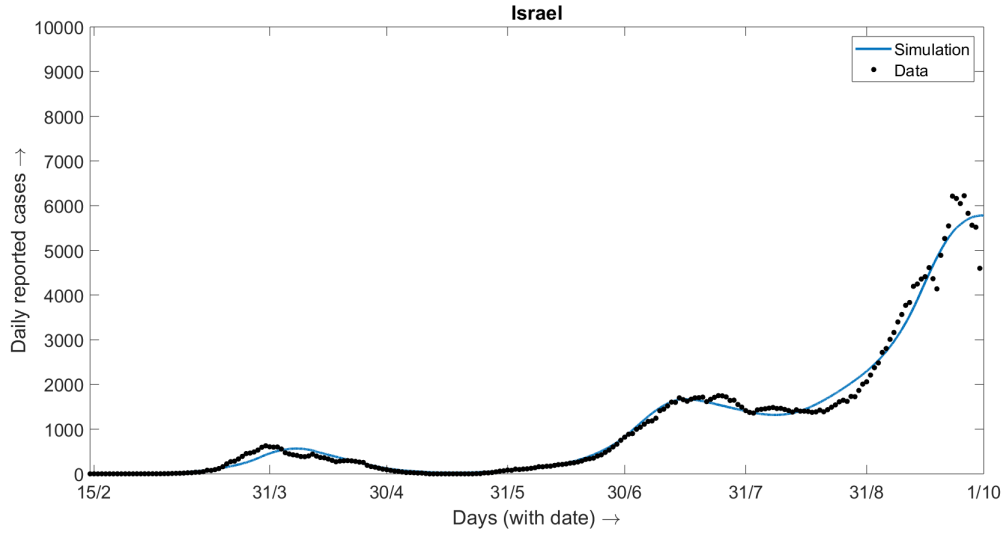

Figure 8: Numerical simulations of epidemic progression in Israel with system (3.1)-(3.13). The values of the coefficients β_{ij} are as follows: $\beta_{11} = 2.91, \beta_{22} = 0.3, \beta_{12} = \beta_{21} = 1.17$. The values of other parameters are given in the appendix. The values of k after lockdown changes from 0.1 to 0.3.

Another example is shown in Figure 8. Fitting of data for Israel gives $\beta_{11} = 2.91$, $\beta_{22} = 0.3$, $\beta_{12} = \beta_{21} = 1.17$. The value of k after lockdown varied from 0.1 to 0.3. The estimates of the coefficients β_{ij} for some European countries are presented in [22]. The values around $\beta_{11} = 4$ and $\beta_{22} = 1$ are quite specific, and we used them in the previous sections of

this work. The value of k after the first lockdown usually changes between 0.1 and 0.3. We will use these characteristic values of parameters in the next subsection in order to study the influence of vaccination on the epidemic progression.

4.2 Vaccination

We proceed to the effect of vaccination on the epidemic progression. Similar to the modelling approach considered in Section 2, we apply vaccination at some time $t = t_0$ and model it by decreasing the number of susceptible individuals:

$$S_1(t_0 + 0) = S_1(t_0 - 0) - \kappa V \ , \quad S_2(t_0 + 0) = S_2(t_0 - 0) - (1 - \kappa)V,$$

while all other classes do not change. Here κ is the proportion of all vaccinated individuals V in the first sub-population and $(1 - \kappa)$ in the second one.

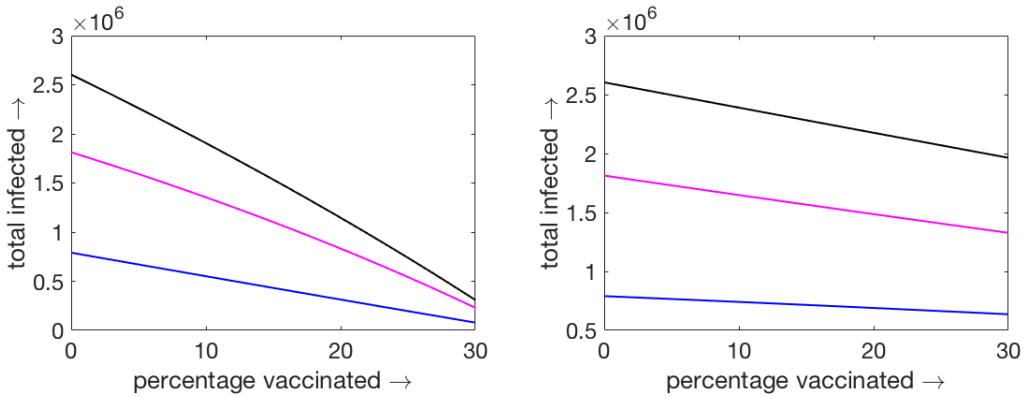

Figure 9: Total number of infected individuals at the end of epidemic as a function of the number vaccinated individuals applied only to the first sub-population ($\kappa = 1$, left figure) or only to the second sub-population ($\kappa = 0$, right figure). The percentage of vaccinated individuals is counted with respect to N_1 in both cases. The lower curve shows the total number on infected individuals in the first sub-population, the middle curve in the second sub-population, and the upper curve their sum.

Figures 9 and 10 show the influence of the number of vaccinated individuals and of their distribution among the two sub-population on the total and maximum number of infected individuals. The total number of infected individuals in the first sub-population is determined as $I_1^T = N_1 - S_1^f - \kappa V$ and in the second sub-population $I_2^T = N_2 - S_2^f - (1-\kappa)V$, where S_j^f, $j = 1, 2$ are the final numbers of susceptible individuals, $S_j^f = S_j(t_f)$, $j = 1, 2$ where t_f is time at which $I_{a_1}(t) + I_{a_2}(t) + I_{s_1}(t) + I_{s_2}(t) < 1$. The maximal current number of infected individuals in each sub-class are defined as the maximum of the function $I_m^i(t) = I_{s_j}(t) + I_{a_j}(t)$, $j = 1, 2$.

The result of the vaccination strongly depends on its distribution between the two sub-populations. If vaccination is applied to the first sub-population ($\kappa = 1$), it decreases the

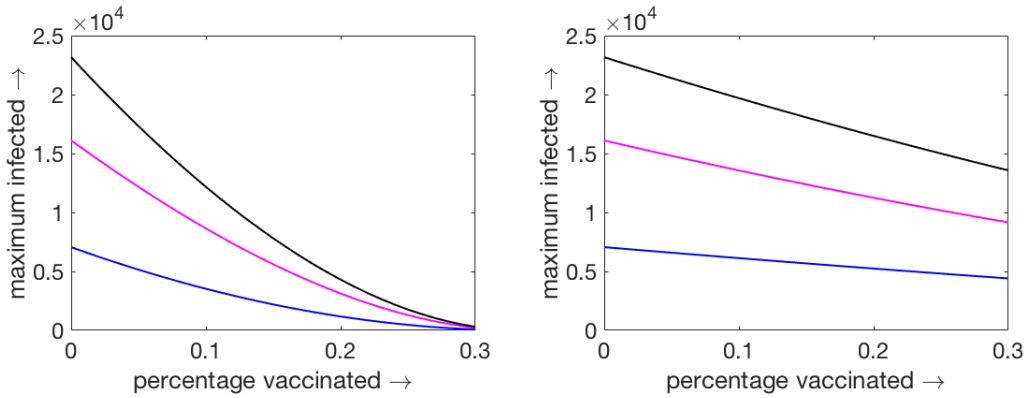

Figure 10: Maximal current number of infected individuals as a function of the number vaccinated individuals applied only to the first sub-population ($\kappa = 1$, left figure) or only to the second sub-population ($\kappa = 0$, right figure). The percentage of vaccinated individuals is counted with respect to N_1 in both cases. The lower curve shows the maximal number on infected individuals in the first sub-population, the middle curve in the second sub-population, and the upper curve their sum.

total number of infected individuals much stronger than if it is applied to the second sub-population ($\kappa = 0$) (Figure 9). In this example, $k = 0.2$ and $N_1 = 0.2N$. The percentage of vaccinated individuals is measured with respect to N_1. So, 30% of vaccination in HT class correspond to 6% of the total population N. In this case, the total number of infected individuals at the end of epidemic decreases 6 times if the vaccination is applied to the first sub-population, compared to the same number of vaccination (6% of total population) is applied to the second sub-population. This striking difference shows that in the first case vaccination acts to stop epidemic progression while in the second case it only protects vaccinated individuals from infection. This difference is even more essential for the current maximal number of infected individuals (Figure 10).

5 Discussion

Actual stage and epidemic waves. At the end of the first year of the Covid-19 epidemic, some of its properties are already sufficiently well understood. Among them, it is now clear that the epidemic progression obeys the usual epidemiological laws, and it can be described by conventional epidemiological models. There is a large body of research devoted to the description of the first stage of epidemic with such models, to the determination of the basic reproduction number and to related questions, basically with ODE models (see [14] and the references therein) but also with individual based models [5].

Another important observation is that, due to relatively high mortality rate and proportion of severe cases, the epidemic cannot be left to follow its natural route to collective immunity and extinction. National health systems become rapidly saturated and fail to

treat not only coronavirus patients but the whole population. Therefore, the only available method up to now consists in the introduction of the measures of social distancing, confinement, wearing mask and even strict lockdown [12].

On the other hand, these measures impose a heavy burden on the economy, and they are relaxed as soon as the epidemiological situation improves. Some time later, the epidemic progression restarts, and it becomes necessary to introduce these measures again. We now observe the second wave of epidemic spread at different countries, and they will certainly continue unless the vaccination will stop them. These oscillations in the epidemic progression can be described by the simple delay differential equation

$$\frac{dI}{dt} = \beta(I_d(t-\tau))\frac{S}{N}I(t) - \sigma I(t) \tag{5.1}$$

for the number of infected individuals I assuming that the number of susceptible individuals S is constant in the beginning of epidemic and $\beta(I_d)$ is a decreasing function of the number of new daily cases I_d taken with some time delay. This function describes the measures of social distancing depending on the epidemiological situation. Since the number of new daily cases is usually taken proportional to the product SI, we obtain a closed equation with respect to $I(t)$. During further epidemic progression, when S cannot be considered as constant, the combination of equation (5.1) with the other equations of the epidemiological models will describe the interaction of this oscillatory dynamics with collective immunity. Let us note that equation (5.1) is a particular case of a more general information-related SIR model [7, 8].

Further epidemic progression depends on the structure of the population. At the actual stage of epidemic progression, serological tests in some countries show that there are of the order of 10% of population with antibodies [15, 18]. We are yet far from the collective immunity but we need already to take into account the variation of S. Moreover, vaccination expected during the next year will also change the number of susceptible and, consequently, will influence the pattern of epidemic progression.

In the beginning of epidemic, observed exponential growth of the number of infected individuals can be described for any population structure [1, 11, 17, 26]. The heterogeneity of the population becomes important at the later stages of the epidemic development when it deviates from the exponential growth and approaches the stage of collective immunity and when it decays approaching the final time, defined as time when the number of infected individuals becomes less than 1. Essentially, it indicates that no one remain in the system who can spread the epidemic.

We study the influence of the heterogeneity of the population with the model problem in Section 2. This relatively simple model allows us to determine the final size of epidemic, the total and the maximal current number of infected individuals. The latter is particularly important for the estimation of available hospitals beds. The main conclusion here is that the data on the initial epidemic stage are now sufficient to predict its further progression. In the case of the homogeneous population, conventional SIR model allows the determination

of the final size of epidemic and of the maximal current number of infected individuals solely on the basis of the basic reproduction number, that is, on the basis of the initial growth rate. However, this is not the case for the heterogeneous population any more [6]. With the same initial growth rate, the final size of epidemic and the maximal current number of infected individuals strongly depend on the structure of the heterogeneous population. Moreover, the total and maximal numbers of infected individuals can change several times for realistic values of parameters determining the distribution of population into two groups.

How to estimate the structure of the heterogeneous population. Thus, we come to the question about the estimation of the structure of a heterogeneous population. In the context of COVID-19, it is needless to mention that the number of reasonable grouping seems to be greater than two but for the simplicity of mathematical modelling we restricted oursleves to HT and LT classes only. We use here the data on the Covid-19 epidemic for different countries. The main idea of our approach is to present the population as a combination of two groups, with high and low disease transmission potentials. The first group is related to the period before lockdown without measures of social distancing and some other control measures, and the second group to the period during lockdown when these measures were strict. In a simplified representation, these two groups can be identified by two factors: a) the number of interactions with other individuals, b) respect of the measures of social distancing (masks, sanitizers, and so on). For example, people who have their normal (as before lockdown) average interaction belong to the HT group, those who have reduced interaction (as during lockdown) belong to the LT group. Certainly, this is a simplified representation of the population because each group is heterogeneous itself, and some individuals can change their groups in different time periods. Furthermore, we approximate a gradual distribution of interactions by a binary function. However, these simplifications allow us to obtain tractable analytical results and give a simple description of the population characterized by a single parameter k defined as a proportion of the HT group to the whole population.

The population consisting of two groups is characterized by four parameters β_{ij}, $i, j = 1, 2$ describing the intensity of disease transmission in the groups $S_i I_j$, $i, j = 1, 2$. The coefficient β_{11} for the disease transmission inside the HT group is obtained by fitting the data before lockdown. The coefficient β_{22} for the disease transmission inside the LT group is obtained by fitting the data during lockdown. However, the cross-group coefficients β_{12} and β_{21} cannot be determined from the data. We impose the assumption that $\beta_{12} = \beta_{21} = (\beta_{11} + \beta_{22})/2$. The justification of this assumption is based on the physical example of two groups of balls moving with different speeds, v_1 and v_2. The number of their collisions is proportional to the average speed $(v_1 + v_2)/2$. We are aware that this representation of the population is too simplified, and further analysis of these coefficients is needed.

Let us note that in the model problem (2.1)-(2.4), only parameters β_{ij} are unknown, while σ_j can be estimated from the data on disease duration. Therefore, assuming that the population is homogeneous before lockdown, that is all β_{ij} are equal to each other, we have only one parameter β_{11} to determine by fitting the data. Similarly, a single parameter β_{22}

should be determined from the data during the lockdown, and the single parameter k from the data after lockdown. A similar situation occurs for a more complete model considered in Sections 3 and 4. If we consider more detailed models with different sub-classes inside HT and LT groups and the corresponding contact matrices, then there are more parameters β_{ij}, and they cannot be uniquely determined from the data.

The influence of heterogeneity on vaccination. Knowing the structure of the population, we can investigate how its heterogeneity influences the results of vaccination. Vaccination is modeled as a decrease of the number of susceptible individuals. We assume here that vaccinated individuals cannot become infected, that is, that vaccination is fully efficient. The results of the vaccination strongly differ depending on whether it is applied to the HT group or to the LT group. In the first case, a relatively small part of vaccinated individuals (5% of the total population) can reduced several times the total number of infected individuals at the end of epidemic, and even more, the maximal current number of infected individuals. Therefore, vaccination of the HT group strongly contributes to stop the epidemic in case of limited number of available doses of vaccine. Vaccination of the second group has much weaker influence on the epidemic progression. It is basically reduced to the protection of vaccinated individuals from infection. These results can be qualitatively expected but we need to determine first the structure of the population in order to give quantitative analysis of the results of vaccination. Future directions of this work will include vaccination taking into account age structure and comorbidities in order to reduce the death toll.

Let us also note that vaccination increases the final time of epidemic. In the case of vaccination of the HT group with 6% of vaccinated individuals of the total population and $k = 0.2$, the final time of epidemic increases almost 6 times in comparison with the case without vaccination. More complete results of vaccination are presented in Tables 1-3.

Acknowledgements

This work is supported by the Ministry of Science and Higher Education of the Russian Federation: agreement no. 075-03-2020-223/3 (FSSF-2020-0018).

References

[1] M. Aguiar, E. M. Ortuondo, J. B. Van-Dierdonck, J. Mar, N. Stollenwerk. Modelling COVID 19 in the Basque Country from introduction to control measure response. *Sci. Rep.*, **10** (2020), 17306.

[2] J. Arino, S. Portet. A simple model for COVID-19. *Infect. Dis. Model.*, **5** (2020), 309 – 315.

[3] U, Avila-Ponce de Leon, A, G. C. Perez, E, Avila-Vales. An SEIARD epidemic model for COVID-19 in Mexico: mathematical analysis and state-level forecast. *Chaos Solitons & Fractals*, **140** (2020), 110165.

[4] Y. Belgaid, M. Helal, E. Venturino. Analysis of a model for Coronavirus spread. *Mathematics*, **8**(5), (2020), 820.

[5] A. Bouchnita, A. Jebrane. A multi-scale model quantifies the impact of limited movement of the population and mandatory wearing of face masks in containing the COVID-19 epidemic in Morocco. *Math. Model. Nat. Phenom.*, **15**, (2020), 31.

[6] J. Dolbeault, G. Turinici. Heterogeneous social interactions and the COVID-19 lockdown outcome in a multigroup SEIR model. *Math. Model. Nat. Phenom.*, **15**, (2020), 36.

[7] A. d'Onofrio, P. Manfredi. Information-related changes in contact patterns may trigger oscillations in the endemic prevalence of infectious diseases. *J. Theor. Biol.*, **256**(3), (2009), 473 – 478.

[8] A. d'Onofrio, P. Manfredi. The interplay between voluntary vaccination and reduction of risky behavior: a general behavior-implicit SIR model for vaccine preventable infections. Current Trends in Dynamical Systems in Biology and Natural Sciences. Springer, 185 – 203, 2020.

[9] S. E. Eikenberry, M. Mancuso, E. Iboi, T. Phan, K. Eikenberry, Y. Kuang, E. Kostelich, A. B. Gumel. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infect. Dis. Model.*, **5**, (2020), 293 – 308.

[10] R. Gauchon, N. Ponthus, C. Pothier, C. Rigotti, V. Volpert, S. Derrode, J.-P. Bertoglio, A. Bienvenue, P.-O. Goard, A. Eyraud-Loisel, S. Pageaud, S. Loisel, P. Roy, and group CovDyn. Lessons learnt from the use of compartmental epidemic models over the French lockdown period. *medRxiv*, 2021.

[11] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. D. Filippo, A. D. Matteo, M. Colaneri, Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, **26**(6), (2020), 855 – 860.

[12] L. Gosce, A. Phillips, P. Spinola, R. K. Gupta, I. Abubakar. Modelling SARS-COV2 Spread in London: Approaches to Lift the Lockdown. *J. Infect.*, **81** (2020), 260 – 265.

[13] M. Kochanczyk, F. Grabowski, T. Lipniacki. Super-spreading events initiated the exponential growth phase of COVID-19 with $\mathcal{R}_0$ higher than initially estimated. *R. Soc. Open Sci.*, **7**, (2020), 200786.

[14] M. Kochanczyk, F. Grabowski, T. Lipniacki. Dynamics of COVID-19 pandemic at constant and time-dependent contact rates. *Math. Model. Nat. Phenom.*, **15** (2020), 28.

[15] P.C. Hallal, F.P. Hartwig, B.L. Horta, M.F. Silveira, C.J. Struchiner, L.P. Vidaletti, N.A. Neumann, L.C. Pellanda, O.A. Dellagostin, M.N. Burattini, G.D. Victora, A.M.B. Menezes, F.C. Barros, A.J.D. Barros, C.G. Victora. SARS-CoV-2 antibody prevalence in Brazil: results from two successive nationwide serological household surveys. *The Lancet*, **8** (2020), 1390 – 1398.

[16] S. G. Krantz, P. Polyakov, A. S. R. S. Rao. True epidemic growth construction through harmonic analysis. *J. Theor. Biol.*, **494** (2020), 110243.

[17] A. J. Kucharski, T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, R. M. Eggo. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect. Dis.*, **20** (2020), 553 – 558.

[18] S. Le Vu, G. Jones, F. Anna, T. Rose, J.-B. Richard, S. Bernard-Stoecklin, S. Goyard, C. Demeret, O. Helynck, C. Robin, V. Monnet, L. Perrin de Facci, M.-N. Ungeheuer, L. Leon, Y. Guillois, L. Filleul, P. Charneau, D. Levy-Bruhl, S. van der Werf, H. Noel. Prevalence of SARS-CoV-2 antibodies in France: results from nationwide serological surveillance. *medRxiv*, 2020.

[19] M. Prague, L. Wittkop, A. Collin, Q. Clairon, D. Dutartre, P. Moireau, R. Thiebaut, and B. P. Hejblum. Population modeling of early COVID-19 epidemic dynamics in French regions and estimation of the lockdown impact on infection rate. *medRxiv*, 2020.

[20] J. H. Rojas, M. Paredes, M. Banerjee, Olcay Akman, Anuj Mubayi. Mathematical Modeling & the Transmission Dynamics of SARS-CoV-2 in Cali, Colombia: Implications to a 2020 Outbreak & public health preparedness. *medRxiv*, 2020.

[21] J. Roux, C. Massonnaud, P. Crepey. COVID-19: One-month impact of the French lockdown on the epidemic burden. *medRxiv*, 2020.

[22] S. Sharma, V. Volpert, M. Banerjee. Extended SEIQR type model for COVID-19 epidemic and data analysis. *Math. Biosci. Eng.*, **17**(6) (2020), 7562 – 7604.

[23] S. Sinha. Epidemiological dynamics of the COVID-19 pandemic in India: an interim assessment. *Stat. Appl.*, **18** (2020), 333 – 350.

[24] M. Supino, A. d'Onofrio, F. Luongo, G. Occhipinti, A. Dal Co. The effects of containment measures in the Italian outbreak of COVID-19. *BMC Public Health*, **20**(1) (2020), 1 – 8.

[25] F. Wu, A. Wang, M. Liu, Q. Wang, J. Chen, S. Xia, Y. Ling, Y. Zhang, J. Xun, L. Lu, S. Jiang, H. Lu, Y. Wen, J. Huang. Neutralizing antibody responses to SARS-CoV-2 in a COVID-19 recovered patient cohort and their implications. medRxiv preprint doi: https://doi.org/10.1101/2020.03.30.20047365

[26] J. Yuan, M. Li, G. Lv, Z. K. Lu. Monitoring transmissibility and mortality of COVID-19 in Europe. *Int. J. Infec. Dis.*, **95** (2020), 311 – 325.

[27] S. Zhao, Q. Lin, J. Ran, S. S. Musa, G. Yang, W. Wang, M. H. Wang. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *Int. J. Infect. Dis.*, **92** (2020), 214 – 217.

6 Appendix

The values of parameters used in the simulations presented in Figure 7 are given in the following table.

Table 4: Best fitted values for the parameters of the full model for Germany [22].

Parameter	**Best fit values for 1st 30 days (95% CI) (Sensitivity Index)**	**Best fit values for next 12 days (95% CI) (Sensitivity Index)**	**Best fit values upto 95th day (95% CI) (Sensitivity Index)**
β	3.98 (3.92-4.04)(1)	1.77 (1.74-1.79)(1)	1.05 (1.04-1.06)(1)
p_1	0.177 (0.175-0.179)(0.2383)	0.16 (0.158-0.162)(0.2204)	0.34 (0.336-0.344)(0.5099)
p_2	0.3 (0.296-0.304)(0.4039)	0.3 (0.296-0.304)(0.4133)	0.05 (0.049-0.051)(0.0914)
p_3	0.05 (0.049-0.051)(0.0048)	0.05 (0.049-0.051)(0.0049)	0.05 (0.049-0.051)(0.0053)
p_4	0.05 (0.049-0.051)(0.0192)	0.05 (0.049-0.051)(0.0196)	0.05 (0.049-0.051)(0.0213)
δ	1 (0.987-1.013)(-0.4039)	1 (0.987-1.013)(-0.4133)	0.82 (0.811-0.829)(-0.0914)
σ	0.1 (0.099-0.101)(0.3314)	0.1 (0.099-0.101)(0.3418)	0.1 (0.099-0.101)(0.342)
η	0.9 (0.889-0.911)(-0.2383)	0.9 (0.889-0.911)(-0.2204)	0.9 (0.889-0.911)(-0.5099)
ρ_1	0.1 (0.099-0.101)(-0.0873)	0.1 (0.099-0.101)(-0.0893)	0.1 (0.099-0.101)(-0.0972)
ρ_2	0.07 (-0.0144)	0.07 (-0.0148)	0.07 (-0.0161)
ζ_1	0.07 (-0.0067)	0.07 (-0.0069)	0.07 (-0.0075)
ζ_2	0.1 (-0.0873)	0.1 (-0.0893)	0.1 (-0.0972)
ζ_3	0.14 (-0.0715)	0.14 (-0.0732)	0.14 (-0.0796)
ξ_1	0.14 (-0.00047)	0.14 (-0.00048)	0.14 (-0.00052)
ξ_2	0.1 (0.099-0.101)(-0.0043)	0.1 (0.099-0.101)(-0.0044)	0.1 (0.099-0.101)(-0.0048)
ν	0.05 (-0.0103)	0.05 (-0.0105)	0.05 (-0.0115)
μ	0.28 (0.277-0.283)(−−)	0.28 (0.277-0.283)(−−)	0.28 (0.277-0.283)(−−)