medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Vaccine efficacy for COVID-19 outbreak in New York City

J. Demongeot[(a)], Q. Griette[(b)], P. Magal[(b),*], and G. F. Webb[(c)]

[(a)] *Université Grenoble Alpes, AGEIS EA7407, F-38700 La Tronche, France.*

[(b)] *Univ. Bordeaux, IMB, UMR 5251, F-33400 Talence, France.*

CNRS, IMB, UMR 5251, F-33400 Talence, France.

[(c)] *Mathematics Department, Vanderbilt University, Nashville, TN, United States.*

December 18, 2021

Abstract

This article aims to study the COVID-19 data for New York City. We use both the daily number of second does vaccination and the daily number of reported cases for New York City. This article provides a method to combine an epidemic model and such data. We explore the influence of vaccine efficacy on our results.

Keywords: Vaccine efficacy, Corona virus, reported and unreported cases, parameters identification, epidemic mathematical model.

1 Introduction

Developing vaccines against an infectious agent often requires years of research and testing to ensure efficacy and safety. In contrast, in the case of COVID-19, the vaccines took less than a year to develop and deploy. This rapid development has left many open questions whose answers may affect the usefulness of the epidemiological models proposed for COVID-19. In particular, vaccination efficacy rates are different for different populations in terms of the level and duration of vaccination immunization. We will study vaccinated efficacy according to the epidemic state of vaccinated individuals: susceptible, infected, or having received vaccine doses.

We will neglect both the duration of the immunization period (we will assume it is permanent) and the length of the acquisition phase of immunization after vaccine injection (we will assume it is identical for all vaccinated individuals). We will suppose the vaccination efficacy independent of age while noting that this hypothesis is very restrictive.

More careful considerations encourage caution with regard to the above simplifications and could lead to further studies. For example, vaccination efficacy is dependent on the appearance of variants and the existence of cross-immunization:

(1) Inducing an antibody response may help select variants [25], a phenomenon very difficult to control because it is impossible to know what exact dose of antigenic Spike protein is released after each vaccination and what is its pharmacokinetics and its bio-distribution over time.

(2) In the development of mRNA vaccines, cross-immunity was overlooked entirely [19]. There are anti-coronavirus antibodies and many epitopes common to the various

*Corresponding author. e-mail: pierre.magal@u-bordeaux.fr

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

endemic known coronaviruses, conserved with SARS-CoV-2. Vaccination ignores pre-existing cross-immunity, which is unfortunate, as the doses injected could be adjusted for a response via cross-immunity against epitopes common to coronaviruses. Young individuals are those whose cross-immunity is still active, and it would be helpful to design a vaccination policy to obtain the best efficacy per target population at risk.

Despite these limitations, a first simplifying approach can lead to a model making it possible to predict effective vaccination coverage at the population level that will prevent the appearance of successive epidemic waves.

Daily vaccination data, even if they are global and unrefined (for example, by age group or social classification), make it possible to better understand the effect of vaccination policy and test the consequences of changes in this policy to improve effectiveness. We use an epidemic model to understand the complex interactions between the epidemic dynamic and the epidemic data. Our model considers the changes in the public health policy, such as confinement, social distancing measures, etc., through the time-dependent transmission rate in the model. Data consist of the daily number of reported cases and the daily number of second doses of vaccine. We refer to [12, 13, 14, 15, 16, 18, 22, 23] for more results on the subject.

In the study, we propose a new model for vaccination implementation. We can connect the model with vaccination to a model without vaccination. We will find a simple transformation for the epidemic data to combine the daily reported case data and the cumulative number of vaccinated individuals.

We can use the model to explore controlling the dynamics of virus propagation, for example, by rapidly slowing down an epidemic wave. In this new model, we will take explicitly into account the variable corresponding to the size of the vaccinated population. We will simulate increasing efficacy of vaccination scenarios. We will apply our model to the COVID-19 epidemic in New York City.

2 Materials and Methods

2.1 Data

The data are taken from the New York City Department of Health and Mental Hygiene [26]. The epidemic of Sars-CoV2 started in NYC on February 29, 2020. The first complete vaccination (i.e., vaccination with two doses) started in NYC on December 15, 2020. In Figure 1-(a), the green dots correspond to the day by day constant values of the function $\mathrm{CR}_{\mathrm{Data}}(t)'$ that is used in the model. In Figure 1-(b), this green curve corresponds to the value function $\mathrm{CR}_{\mathrm{Data}}(t) = \int_{t_0}^{t} \mathrm{CR}_{\mathrm{Data}}(\sigma)' d\sigma$ that is used in the model.

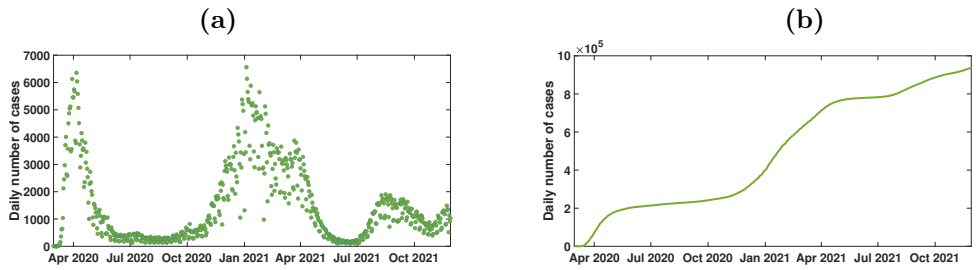

Figure 1: *In (a), we plot* $\mathrm{CR}_{\mathrm{Data}}(t)'$ *the daily number of reported cases of Sars-CoV2 for New York City. In (b), we plot* $\mathrm{CR}_{\mathrm{Data}}(t)$ *the cumulative number of reported cases of Sars-CoV2 for New York City.*

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

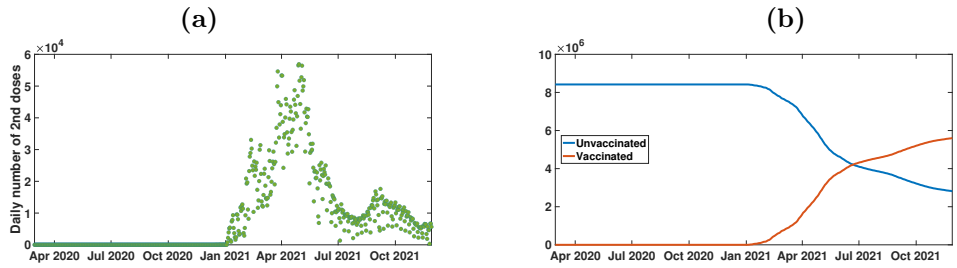

Figure 2: *In (a), we plot* $V_{\text{Data}}(t)$ *the daily number of second doses of vaccine for New York City. In (b), we plot* $\text{CV}_{\text{Data}}(t)$ *the cumulative number of second vaccine doses for New York City (red curve), and* $N - \text{CV}_{\text{Data}}(t)$ *the number of unvaccinated individuals for New York City (blue curve). The two curves in (b) cross when the number of vaccinated people reaches* 50%.

In Figure 2-(a), the green dots correspond to the day by day constant values of the function $V_{\text{Data}}(t)$ that is used in the model. In Figure 2-(b), the red curve corresponds the function $\text{CV}_{\text{Data}}(t) = \int_{t_0}^{t} V_{\text{Data}}(\sigma) d\sigma$ that is used in the model.

2.2 Epidemic model

Many epidemiological models are based on SIR or SEIR models, which are classical in epidemic modelling. We refer to [24, 20] for early articles devoted to such models and to [1, 2, 5, 3, 4, 6, 8, 17, 21] for later models. In this section, we compare the following SEIUR model to cumulative reported cases data

$$\begin{cases} S'(t) = -\tau(t)\left[I(t) + \kappa\, U(t)\right] S(t) & -e\, V_{\text{Data}}(t) \dfrac{S(t)}{N_{UV}(t)}, \\ E'(t) = \tau(t)\left[I(t) + \kappa\, U(t)\right] S(t) - \alpha\, E(t) & -e\, V_{\text{Data}}(t) \dfrac{E(t)}{N_{UV}(t)}, \\ I'(t) = \alpha\, E(t) - \nu\, I(t) & -e\, V_{\text{Data}}(t) \dfrac{I(t)}{N_{UV}(t)}, \\ U'(t) = \nu\,(1-f)\, I(t) - \eta\, U(t) & -e\, V_{\text{Data}}(t) \dfrac{U(t)}{N_{UV}(t)}, \\ R'(t) = \nu\, f\, I(t) - \eta\, R(t), \end{cases} \tag{2.1}$$

where at time t, $S(t)$ is the number of susceptible uninfected individuals, $E(t)$ is the number of exposed individuals (infected, but not yet capable of transmitting the infection), $I(t)$ is the number of asymptomatic infectious individuals, $R(t)$ is the number of reported symptomatic infectious individuals, and $U(t)$ is the number of unreported symptomatic infectious individuals. $N_{UV}(t)$ is the number of unvaccinated individuals. In the model, $S(t)/N_{UV}(t)$ (respectively, $E(t)/N_{UV}(t)$, $I(t)/N_{UV}(t)$, and $U(t)/N_{UV}(t)$) is the fraction of susceptible (respectively, infected, reported, and unreported) in the population of unvaccinated individuals.

The system (2.1) is supplemented by the initial data

$$S(t_0) = S_0,\ E(t_0) = E_0,\ I(t_0) = I_0,\ U(t_0) = U_0,\ \text{and}\ R(t_0) = R_0. \tag{2.2}$$

The mathematical model corresponds to the flowchart in Figure 3.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

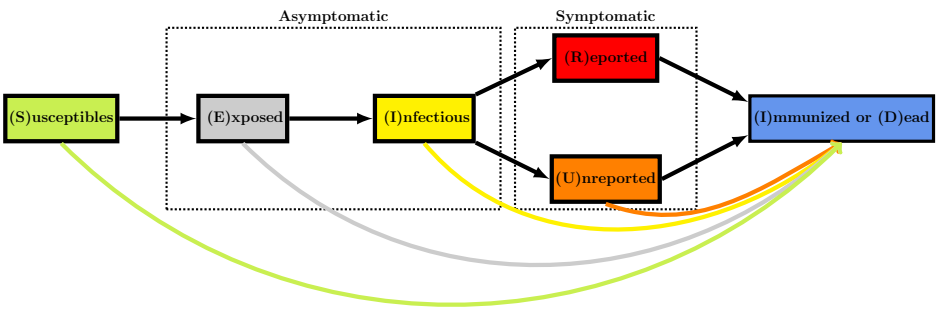

Figure 3: *Flowchart for the model. The colored arrows at the bottom represent the vaccination.*

In the model, $\tau(t)$ is the time-dependent rate of transmission, $1/\alpha$ is the average duration of the exposed period, $1/\nu$ is the average duration of the asymptomatic infectious period, and for simplicity, we subdivide the class of symptomatic infectious individuals into the fraction $0 \leq f \leq 1$ showing severe symptoms, and the fraction $1 - f$ showing mild symptoms, assumed to be undetected. The quantity $1/\eta$ is the average duration of the symptomatic infectious period for both unreported and reported symptomatic individuals. Asymptomatic infectious and unreported symptomatic infection individuals both contribute to the infection of susceptible individuals, with the parameter κ (≥ 1 or ≤ 1) corresponding to their relative contributions. It is assumed that reported symptomatic individuals do not contribute significantly to the transmission of the virus.

In the model, the parameter $0 \leq e \leq 1$ is the vaccine efficacy. This means that when $e = 0$ the vaccine is not effective at all, and if $e = 1$ the vaccine is fully effective. The cumulative number of removed individuals $t \to \mathrm{ID}(t)$, immunized (recovered or vaccinated), and dead, satisfies the equation

$$\mathrm{ID}'(t) = \eta\,[R(t) + U(t)] + e\,V_{\mathrm{Data}}(t)\left[\frac{S(t) + E(t) + I(t) + U(t)}{N_{UV}(t)}\right]. \tag{2.3}$$

In this model, $V_{\mathrm{Data}}(t)$ is the flux of new vaccinated individuals. This means that

$$\int_{t_1}^{t_2} V_{\mathrm{Data}}(\sigma)d\sigma,$$

is the total number of vaccinated individuals between t_1 and t_2.

Since no individuals were vaccinated at the start of the epidemic (i.e. for $t = t_0$), we can assume that the total number of individuals N in the population at time t_0 is

$$N = S_0 + E_0 + I_0 + R_0 + U_0.$$

The cumulative number of vaccinated individuals is given by

$$\mathrm{CV}'_{\mathrm{Data}}(t) = V_{\mathrm{Data}}(t), \text{ and } \mathrm{CV}_{\mathrm{Data}}(t_0) = 0,$$

which is equivalent to

$$\mathrm{CV}_{\mathrm{Data}}(t) = \int_{t_0}^{t} V_{\mathrm{Data}}\,(\sigma)\,\mathrm{d}\sigma.$$

The number of unvaccinated individuals is

$$N_{UV}(t) = N - \mathrm{CV}_{\mathrm{Data}}(t).$$

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Therefore the model (2.1) can be rewritten as follows

$$\begin{cases} S'(t) = -\tau(t)\left[I(t) + \kappa\, U(t)\right] S(t) & -e\,\dfrac{\mathrm{CV}'_{\mathrm{Data}}(t)}{N - \mathrm{CV}_{\mathrm{Data}}(t)} S(t), \\ E'(t) = \tau(t)\left[I(t) + \kappa\, U(t)\right] S(t) - \alpha\, E(t) & -e\,\dfrac{\mathrm{CV}'_{\mathrm{Data}}(t)}{N - \mathrm{CV}_{\mathrm{Data}}(t)} E(t), \\ I'(t) = \alpha\, E(t) - \nu\, I(t) & -e\,\dfrac{\mathrm{CV}'_{\mathrm{Data}}(t)}{N - \mathrm{CV}_{\mathrm{Data}}(t)} I(t), \\ U'(t) = \nu\,(1-f)\, I(t) - \eta\, U(t) & -e\,\dfrac{\mathrm{CV}'_{\mathrm{Data}}(t)}{N - \mathrm{CV}_{\mathrm{Data}}(t)} U(t), \end{cases} \tag{2.4}$$

Remark 1 *We did not include an R equation in (2.4), because the R compartment is decoupled from the rest of the system and we will not use it in the following.*

At the end of the asymptomatic infectious period (corresponding to the I compartment), it is assumed that a fraction $f \in (0, 1]$ of infectious individuals is reported. Therefore, the cumulative number of reported cases $\mathrm{CR}(t)$ is connected to the epidemic model by the following relationship

$$\mathrm{CR}'_{\mathrm{Data}}(t) = \nu\, f\, I(t), \text{ for } t \geq t_0, \text{ and } \mathrm{CR}_{\mathrm{Data}}(t_0) = \mathrm{CR}_0. \tag{2.5}$$

Given and estimated parameters

In the model, the data are represented by $\mathrm{CR}'_{\mathrm{Data}}(t)$, the daily number of reported cases, and $V_{\mathrm{Data}}(t)$, the daily number of vaccinations.

In order to compare the model and the data, it is assumed that the known parameters are

$$S_0, U_0, \kappa, \alpha, \nu, f, \eta.$$

The three remaining parameters are estimated from the above quantities:

$$E_0, I_0, t \to \tau(t).$$

2.3 Identification problem

We define the fraction of not effectively vaccinated individuals at time t, starting from the time t_0, by

$$W(t) = \left(\frac{N - \mathrm{CV}_{\mathrm{Data}}(t)}{N - \mathrm{CV}_{\mathrm{Data}}(t_0)}\right)^e,$$

where $\mathrm{CV}_{\mathrm{Data}}(t)$ is the cumulative number of second dose vaccinated. Then $W(0) = 1$, and $t \to W(t)$ is a non-increasing function. Define

$$\widehat{\mathrm{CR}}_{\mathrm{Data}}(t) = \int_{t_0}^{t} \frac{\mathrm{CR}'_{\mathrm{Data}}(\sigma)}{W(\sigma)} d\sigma. \tag{2.6}$$

By applying the results in Griette, Demongeot and Magal [11] to the system (A.2) in Appendix A , we obtain the following formula for the rate of transmission expressed in function of the cumulative number of reported cases $t \to \mathrm{CR}_{\mathrm{Data}}(t)$, and the cumulative number of vaccinated individuals $t \to \mathrm{CV}_{\mathrm{Data}}(t)$.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Computation of the rate of transmission

The transmission rate is fully determined by the parameters $\kappa, \alpha, \nu, \eta, f, S_0, E_0, I_0, U_0$, and by using the five following equations for $t \geq t_0$

$$\tau(t) = \frac{\widehat{\tau}(t)}{W(t)}, \tag{2.7}$$

where for $t \geq t_0$,

$$\widehat{\tau}(t) = \frac{1}{\widehat{I}(t) + \kappa \widehat{U}(t)} \times \frac{\widehat{\mathrm{CE}}''(t) + \alpha \widehat{\mathrm{CE}}'(t)}{E_0 + S_0 - \widehat{\mathrm{CE}}'(t) - \alpha \widehat{\mathrm{CE}}(t)},$$

where

$$\widehat{I}(t) = \frac{\widehat{\mathrm{CR}}'_{\mathrm{Data}}(t)}{\nu f},$$

$$\widehat{\mathrm{CE}}(t) = \frac{1}{\alpha \nu f} \left[\widehat{\mathrm{CR}}'_{\mathrm{Data}}(t) - \nu f I_0 + \nu \widehat{\mathrm{CR}}_{\mathrm{Data}}(t) \right],$$

$$\widehat{U}(t) = \mathrm{e}^{-\eta(t-t_0)} U_0 + \int_{t_0}^{t} \mathrm{e}^{-\eta(t-s)} \frac{(1-f)}{f} \widehat{\mathrm{CR}}'_{\mathrm{Data}}(s) \mathrm{d}s.$$

The data that are represented by the functions $t \to \mathrm{CR}_{\mathrm{Data}}(t)$ cumulative number of reported cases, and $t \to V_{\mathrm{Data}}(t)$ the cumulative number of second doses of vaccine are involved in the formula (2.6) to define $\widehat{\mathrm{CR}}_{\mathrm{Data}}(t)$.

Computation of some initial values from the data

From (2.5) we obtain

$$I_0 = \frac{\mathrm{CR}'_{\mathrm{Data}}(t_0)}{\nu f},$$

and by using the I-equation of system (A.2) and (A.5), we obtain

$$E_0 = \frac{1}{\alpha} \left\{ \frac{1}{\nu f} \left[\frac{\mathrm{CR}'_{\mathrm{Data}}(t_0)}{W(t_0)} \right]' + \frac{\mathrm{CR}'_{\mathrm{Data}}(t_0)}{f} \right\}.$$

2.4 Data normalized by $W(t)$

In Figure 4-(a), we plot the daily number of reported cases normalized by $W(t)$ (the fraction of not efficiently vaccinated individuals at time t)

$$\widehat{\mathrm{CR}}_{\mathrm{Data}}(t)' = \frac{\mathrm{CR}_{\mathrm{Data}}(t)'}{W(t)} = \mathrm{CR}_{\mathrm{Data}}(t)' \left(\frac{N}{N - \mathrm{CV}_{\mathrm{Data}}(t)} \right)^{e}$$

for several values of $e = 0,\ 0.25,\ 0.5,\ ,0.75,\ 1$.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

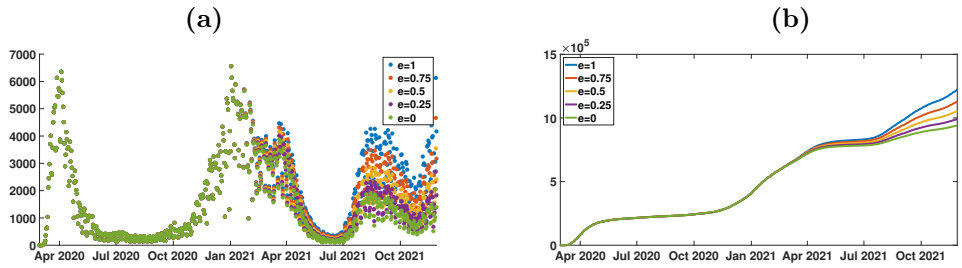

Figure 4: *We plot* $\widehat{\mathrm{CR}}_{\mathrm{Data}}(t)'$ *in (a), and we plot* $\widehat{\mathrm{CR}}_{\mathrm{Data}}(t)$ *in (b) for New York City, and* $e = 0, 0.25, 0.5, , 0.75, 1$. *In (a), the green dots for* $e = 0$ *corresponds to the original daily number of reported cases* $\mathrm{CR}_{\mathrm{Data}}(t)'$. *In (b), the green curve for* $e = 0$ *also corresponds to the original cumulative number of reported cases* $\mathrm{CR}_{\mathrm{Data}}(t)$.

In Figure 4-(b), we plot the daily number of cases normalized by $W(t)$, that is

$$\widehat{\mathrm{CR}}_{\mathrm{Data}}(t) = \int_{t_0}^{t} \frac{\mathrm{CR}_{\mathrm{Data}}(\sigma)'}{W(\sigma)} d\sigma = \int_{t_0}^{t} \mathrm{CR}_{\mathrm{Data}}(\sigma)' \left(\frac{N}{N - \mathrm{CV}_{\mathrm{Data}}(\sigma)} \right)^{e} d\sigma$$

for several values of $e = 0, 0.25, 0.5, , 0.75, 1$.

2.5 Phenomenological model

This section is devoted to the phenomenological model used to regularize the data. We refer to [10, 11] for more information. The phenomenological model is fitted to the cumulative reported cases data during the epidemic periods and extended by a lines in between. We regularize the junction point between the period where the phenomenological model has changed. The regularization is obtained using a convolution with a Gaussian function having a standard deviation equal to 7 days.

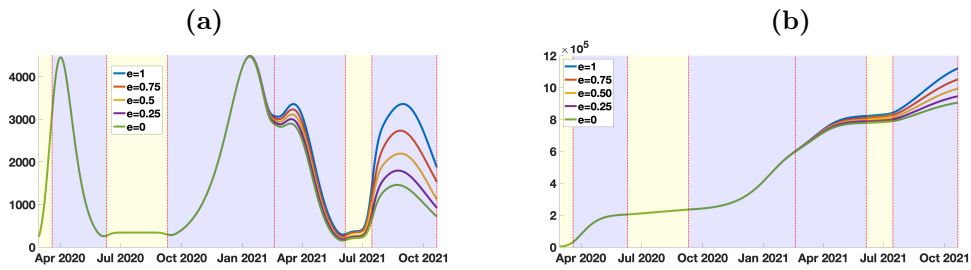

Figure 5: *In (a), we plot the phenomenological model used to represent the daily number of reported cases normalized by the efficient vaccinations. In (b), we plot the phenomenological model used to represent the cumulative number of reported cases normalized by the effective vaccinations. The blue background color regions correspond to epidemic phases, and the yellow background color regions to endemic phases.*

2.6 Instantaneous reproduction numbers

In order to compute the day by day transmission rate $t \to \tau(t)$, we use the formula (2.7). Then by using the model (A.2) we can consider the problem of the instantaneous reproduction numbers (see [11] for more information). To investigate the role of the vaccination for the COVID-19 data, we use our method to compute the transmission rate, and we consider the **instantaneous reproduction number with vaccination**

$$R_V(t) = \frac{\tau(t)S(t)}{\eta\nu} \times (\eta + \kappa\nu(1 - f)), \tag{2.8}$$

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

the **quasi-instantaneous reproduction number with vaccination**

$$R_V^0(t) = \frac{\tau(t)W(t)S_0}{\eta\nu} \times (\eta + \kappa\nu(1-f)), \tag{2.9}$$

and the **quasi-instantaneous reproduction number without vaccination**

$$R^0(t) = \frac{\tau(t)S_0}{\eta\nu} \times (\eta + \kappa\nu(1-f)). \tag{2.10}$$

3 Results

The parameters used in the simulations are listed in Table 1 in Appendix B.

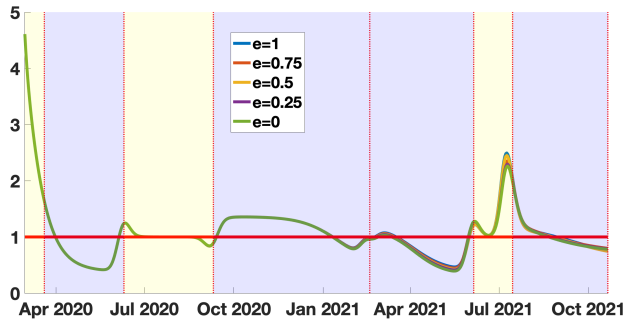

Figure 6: *In this figure we plot the instantaneous reproduction number with vaccination* $R_V(t) = \frac{\tau(t)S(t)}{\eta\nu} \times (\eta + \kappa\nu(1-f))$. *The blue background color regions correspond to epidemic phases, and the yellow background color regions to endemic phases.*

In Figure 6, we observe almost no influence of the vaccine efficacy e on the basic reproduction number. This is due to some compensatory effects between $\tau(t)$ and $S(t)$, because $\tau(t)$ and $S(t)$ are evaluated to adjust the number of cumulative reported cases, which is fixed.

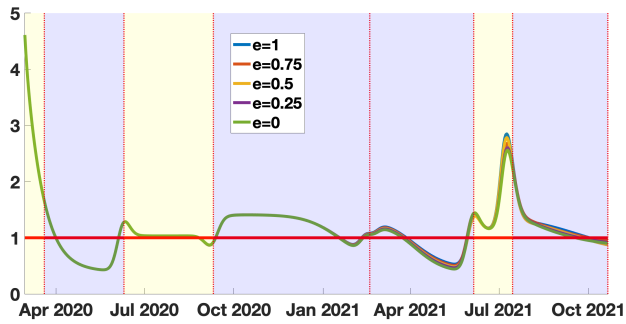

Figure 7: *In this figure we plot the quasi-instantaneous reproduction number with vaccination* $R_V^0(t) = \frac{\tau(t)W(t)S_0}{\eta\nu} \times (\eta + \kappa\nu(1-f))$. *The blue background color regions correspond to epidemic phases, and the yellow background color regions to endemic phases.*

In Figure 7, we see almost no difference with Figure 6. It means that the cumulative number of infected is so small compared to the total size of the population of New York City that the $R_V^0(t)$ is almost unchanged compared to $R_V(t)$. This means that the cumulative number of infected is too small to have a significant impact to reduce the basic reproduction number.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

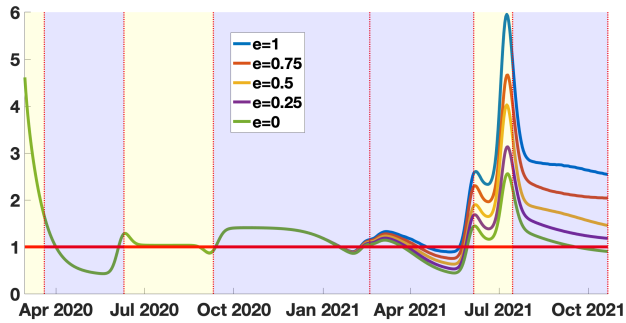

Figure 8: *In this figure we plot the quasi-instantaneous reproduction number without vaccination* $R^0(t) = \frac{\tau(t)S_0}{\eta\nu} \times (\eta + \kappa\nu(1-f))$. *The blue background color regions correspond to epidemic phases, and the yellow background color regions to endemic phases. One may observe that* $R^0(t)$ *is a multiple of the time dependent transmission rate.*

In Figure 8, the blue curve corresponds the $R^0(t)$. The instantaneous reproduction number should be interpreted as the instantaneous reproduction number in the absence of vaccination, conditionality to the fact that the vaccine is fully efficient. The same interpretation holds for $e = 0.75, 0.5, 0.25, 0$. This means that vaccination has a strong influence on the dynamic of the epidemic. This influence indeed strongly depends on the vaccine efficacy e. We can see that during the most recent epidemic wave, the situation in New York City would have been much worst in the absence of vaccination. We can see that, for $e = 0.75$ (red curve), the last peak of $R^0(t)$ is around 4.5.

We cannot estimate the value of the vaccine efficacy. But conditionally to the value of e, we can draw some conclusions. If vaccine efficacy is above 0.75, it means that New York City escapes an epidemic wave as bad (or worse) than the first one (this corresponds to the region between the blue curve and the red curve in Figure 8). If the vaccine efficacy is between 0.25 and 0.75, we can see a significant gain compared to the green curve (this corresponds to the region between the purple curve and red curve in Figure 8).

4 Discussion

In this article, we developed a new method to model the COVID-19 epidemic by using the daily reported cases and vaccination data. We use phenomenological models to get an exploitable reconstruction of the history of the epidemic and develop a new method to identify the parameters of an epidemic model with vaccination that reproduces the exact behavior of the data.

Since the first efficient vaccines against SARS-CoV-2 appeared at the end of 2020, many countries have implemented vaccination policies to protect their population. As a result, a non-negligible fraction of the population has acquired at least a partial immunity against the disease. This means that the number of susceptible hosts has been significantly reduced. Several studies, including some authors' works, developed methods to connect the data with epidemic models. In the presence of vaccination, these models may overestimate the number of susceptible hosts, and their conclusions should therefore be taken with precaution. We correct this flaw by including vaccination data in our model in the present study. We construct, in particular, the transmission rate $\tau(t)$ and the instantaneous reproduction number $R_V(t)$ of the disease.

In Figures 6 and 7, we present our computations concerning the instantaneous and quasi-instantaneous reproduction number $R_V(t)$ and $R_V^0(t)$. The explicit formula developed in this paper allows us to investigate the role of the vaccine efficacy parameter e. Surprisingly, the instantaneous reproduction number reconstructed from the data does not depend very

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

much on the vaccine's efficacy. We observe that the five curves presented in Figure 6 are almost equal (the same can be said about the quasi-instantaneous reproduction number in Figure 7). We understand this phenomenon as a balance between the number of susceptible hosts $S(t)$ and the transmission rate $\tau(t)$. This is because, since the data is unchanged, the increase in the efficacy of the vaccine reduces the number of available hosts, so the transmission rate must be increased to recover the data. In this process, the product $\tau(t)S(t)$ (and therefore $R_V(t)$ and $R_V^0(t)$) is almost unchanged.

Immunization data can help to understand better the effectiveness of a city, a state, or a country's immunization policy. Strictly speaking, this policy should be adapted to the populations at risk and, for example, be different according to the age groups and take into account the progressive degradation of the immunity conferred by the vaccines, due to the appearance of variants of the initial virus.The data used in this article do not consider any specificity of sub-populations for New York City. Therefore the age groups were not considered, nor the extinction of immunity over time, which could be taken into account by considering an efficacy e dependent on the time and a flux of the vaccinated back into the susceptible compartment.

Despite these shortcomings, the model clearly shows the impact of the vaccine policy on the epidemic dynamics, thanks to the explicit formulas allowing the calculation of important parameters, such as the transmission rate. In addition, the model allows for the introduction of additional elements, when documented by observed data, such as age, loss of immunization, and cross-immunization. This last phenomenon causes weights on the vaccine policy's effectiveness, which is interesting for further investigation.

Vaccination confers a new immunity, which is in addition to a possible pre-existing cross-immunity [19], and the vaccination policy could be adjusted in relation to the response via cross-immunity against epitopes common to numerous coronaviruses. For example, if the age classes are considered, young individuals are those whose cross-immunity is still active, causing a strong response to vaccination with possible systemic undesirable effects. Consequently, it would be interesting in the future to develop improvements of the model in order to refine the number of doses per target population at risk and thus ensure, for a smaller quantity of vaccinated, the same efficacy in the immunization of the general population.

Our model could be extended in several other directions. Here we did not distinguish between immunized, dead, and vaccinated individuals. These could be added to the model, and other phenomena could be included as well, like a different fading rate of immunity coming from the disease and the vaccine, provided the associated parameters (death rate, etc.) are known. By including age classes, we could also distinguish the strength of the immune response according to age and better measure the benefit-risk ratio of vaccination with respect to age. We could also implement a different rate of loss of immunity by age class.

Appendix

A Transformation of the system into a system into a standard epidemic model

The goal of this section is to connect the model with vaccination to the model without vaccination used previously in [11]. Our goal is to apply to the transformed system some of the results obtained in [11]. Set

$$W(t) := \exp\left(e\int_{t_0}^{t} \frac{-\mathrm{CV}_{\mathrm{Data}}(\sigma)'}{N-\mathrm{CV}_{\mathrm{Data}}(\sigma)}d\sigma\right).$$

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Then

$$W(t) = \exp\left(e\,\{\ln\left[N - \mathrm{CV}_{\mathrm{Data}}(t)\right] - \ln\left[N - \mathrm{CV}_{\mathrm{Data}}(t_0)\right]\}\right)$$

therefore

$$W(t) = \left(\frac{N - \mathrm{CV}_{\mathrm{Data}}(t)}{N - \mathrm{CV}_{\mathrm{Data}}(t_0)}\right)^e.$$

By integrating the S-equation of system (2.4) we obtain

$$S(t) = S_0 \exp\left(-\int_{t_0}^{t} \tau(\sigma)\left[I(\sigma) + \kappa\, U(\sigma)\right] d\sigma + e\int_{t_0}^{t} \frac{-\mathrm{CV}_{\mathrm{Data}}(\sigma)'}{N - \mathrm{CV}_{\mathrm{Data}}(\sigma)} d\sigma\right).$$

Hence we have

$$S(t) = S_0 \exp\left(-\int_{t_0}^{t} \tau(\sigma)\left[I(\sigma) + \kappa\, U(\sigma)\right] d\sigma\right) W(t). \tag{A.1}$$

Define

$$\widehat{S}(t) := \frac{S(t)}{W(t)}.$$

It follows that from (A.1) that

$$\widehat{S}'(t) = -\tau(t)\left[I(t) + \kappa\, U(t)\right]\widehat{S}(t).$$

By setting

$$\widehat{I}(t) := \frac{I(t)}{W(t)}., \widehat{U}(t) := \frac{U(t)}{W(t)}, \text{ and } \widehat{\tau}(t) = \tau(t) \times W(t),$$

we obtain

$$\widehat{S}'(t) = -\widehat{\tau}(t)\left[\widehat{I}(t) + \kappa\, \widehat{U}(t)\right]\widehat{S}(t).$$

By replacing (A.1) in the E-equation of system (2.4), we obtain

$$\begin{aligned} E'(t) = \;& \tau(t)\left[I(t) + \kappa\, U(t)\right] S_0 \exp\left(-\textstyle\int_{t_0}^{t} \tau(\sigma)\left[I(\sigma) + \kappa\, U(\sigma)\right] d\sigma\right) W(t) \\ & -\alpha\, E(t) - e\,\frac{\mathrm{CV}_{\mathrm{Data}}(t)'}{N - \mathrm{CV}_{\mathrm{Data}}(t)} E(t). \end{aligned}$$

We observe that

$$\frac{1}{W(t)} = \exp\left(\int_{t_0}^{t} e\,\frac{\mathrm{CV}_{\mathrm{Data}}(\sigma)'}{N - \mathrm{CV}_{\mathrm{Data}}(\sigma)} d\sigma\right) = \left(\frac{N - \mathrm{CV}_{\mathrm{Data}}(t)}{N - \mathrm{CV}_{\mathrm{Data}}(t_0)}\right)^e,$$

and

$$\left(\frac{E(t)}{W(t)}\right)' = \frac{E'(t)}{W(t)} + e\,\frac{\mathrm{CV}_{\mathrm{Data}}(t)'}{N - \mathrm{CV}_{\mathrm{Data}}(t)}\frac{E(t)}{W(t)},$$

therefore, we obtain

$$\left(\frac{E(t)}{W(t)}\right)' = \tau(t)\left[I(t) + \kappa\, U(t)\right] S_0 \exp\left(-\int_{t_0}^{t} \tau(\sigma)\left[I(\sigma) + \kappa\, U(\sigma)\right] d\sigma\right) - \alpha\,\frac{E(t)}{W(t)},$$

and by setting

$$\widehat{E}(t) := \frac{E(t)}{W(t)}.$$

we obtain

$$\widehat{E}'(t) = \tau(t)\left[I(t) + \kappa\, U(t)\right]\widehat{S}(t) - \alpha\widehat{E}(t).$$

By using similar argument for the remaining equations of (2.4) we obtain the following result.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Lemma 2 (Transformation of the system) *We define for $t \geq t_0$,*

$$\widehat{S}(t) = \frac{S(t)}{W(t)}, \widehat{E}(t) = \frac{E(t)}{W(t)}, \widehat{I}(t) = \frac{I(t)}{W(t)}, \widehat{U}(t) = \frac{U(t)}{W(t)}.$$

Then the system (2.4) becomes for each $t \geq t_0$,

$$\begin{cases} \widehat{S}'(t) = -\widehat{\tau}(t)\left[\widehat{I}(t) + \kappa\,\widehat{U}(t)\right]\widehat{S}(t), \\ \widehat{E}'(t) = \widehat{\tau}(t)\left[\widehat{I}(t) + \kappa\,\widehat{U}(t)\right]\widehat{S}(t) - \alpha\,\widehat{E}(t), \\ \widehat{I}'(t) = \alpha\,\widehat{E}(t) - \nu\,\widehat{I}(t), \\ \widehat{U}'(t) = \nu\,(1-f)\,\widehat{I}(t) - \eta\,\widehat{U}(t), \end{cases} \tag{A.2}$$

with initial data

$$\widehat{S}(t_0) = S_0,\ \widehat{E}(t_0) = E_0,\ \widehat{I}(t_0) = I_0, \widehat{R}(t_0) = R_0,\ \widehat{U}(t_0) = U_0,$$

and

$$\widehat{\tau}(t) = \tau(t) \times W(t). \tag{A.3}$$

Connection with the data: By using the equation (2.5) (connecting the model and the data), we obtain

$$\nu f \widehat{I}(t) = \frac{\mathrm{CR}'_{\mathrm{Data}}(t)}{W(t)}, \forall t \geq t_0.$$

We define for all $t \geq t_0$,

$$\widehat{\mathrm{CR}}_{\mathrm{Data}}(t) := \int_{t_0}^{t} \frac{\mathrm{CR}'_{\mathrm{Data}}(\sigma)}{W(\sigma)} d\sigma. \tag{A.4}$$

Then we obtain the same formula than (2.5) (with hat), namely

$$\widehat{\mathrm{CR}}'_{\mathrm{Data}}(t) = \nu f \widehat{I}(t), \forall t \geq t_0. \tag{A.5}$$

B Table of parameters

Period	Interpretation	Parameters value	Method
S_0	Number of susceptible individuals at time t_0	8.419×10^6	[27]
E_0	Number of exposed individuals at time t_0		Computed
I_0	Number of asymptomatic infectious individuals at time t_0		Computed
U_0	Number of unreported symptomatic infectious at time t_0	0	Fixed
$\tau(t)$	Transmission rate		Computed
f	Fraction of reported symptomatic infectious	0.8	Fixed
κ	Fraction of unreported symptomatic infectious capable to transmit the pathogen	1	Fixed
$1/\alpha$	Average duration of the exposed period	3 days	Fixed
$1/\nu$	Average duration of the asymptomatic infectious period	7 days	Fixed
$1/\eta$	Average duration of the symptomatic infectious period	7 days	Fixed
e	Vaccine efficacy	0, 0.25, 0.5, 0.75, 1	Fixed

Table 1: *In this table we list the values of the parameters of the epidemic model used for the simulations.*

Conflicts of Interest: Declare conflicts of interest or state "The authors declare no conflict of interest." The authors declare no conflict of interest.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

References

[1] Anderson, R. M., and May, R. M., *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, 1992.

[2] Bailey, N. T. J. , *The Mathematical Theory of Epidemics*, Hafner Publishing Co., New York, 1957.

[3] Brauer, F., and Castillo-Chavez, C. , *Mathematical Models in Population Biology and Epidemiology*, Springer, New York, $2^n d$ edition, 2012.

[4] Brauer, F. , Castillo-Chavez, C. and Feng, Z. , *Mathematical Models in Epidemiology*, Springer, New York, 2019.

[5] Brauer, F. , van den Driessche, P., and Wu, J. , *Mathematical Epidemiology*, Springer, Berlin, Germany, 2008.

[6] Busenberg, S., and Cooke, K. , *Vertically Transmitted Diseases*, Springer-Verlag, Berlin, 1993.

[7] Demongeot, J. , Griette, Q., and Magal, P., SI epidemic model applied to COVID-19 data in mainland China. *Royal Society Open Science* 7:201878 (2020).

[8] Diekmann, O., Heesterbeek, H., and Britton, T., *Mathematical Tools for Understanding Infectious Disease Dynamics*, Princeton University Press, Princeton, NJ, 2013.

[9] Diekmann, O., Heesterbeek, J. A. P., and Metz, J. A. J., On the definition and the computation of the basic reproduction ratio r_0 in models for infectious diseases in heterogeneous populations, *J. Math. Biol.*, **28** (1990), 365–382. doi: 10.1007/BF00178324.

[10] Griette, Q. , Demongeot, J., and Magal, P., A robust phenomenological approach to investigate COVID-19 data for France. *Mathematics in Applied Sciences and Engineering*, **2(3)** (2021), 149-218.

[11] Griette, Q. , Demongeot, J., and Magal, P., What can we learn from COVID-19 data by using epidemic models with unidentified infectious cases? *Mathematical Biosciences and Engineering*, **19(1)** (2021), 537?594.

[12] Griette, Q., and Magal, P., Clarifying predictions for COVID-19 from testing data: The example of New York State, *Infectious Disease Modelling*, **6** (2021), 273-283.

[13] Iboi, E. A., Ngonghala, C. N. , and Gumel, A. B. , Will an imperfect vaccine curtail the COVID-19 pandemic in the US? *Infectious Disease Modelling*, **5** (2020), 510-524.

[14] Li, Q. , Tang, B. , Bragazzi, N. L., Xiao, Y., and Wu, J., Modeling the impact of mass influenza vaccination and public health interventions on COVID-19 epidemics with limited detection capability. *Mathematical biosciences*, **325**, 108378 (2020).

[15] Moore, S. , Hill, E., Dyson, L. , Tildesley, M., and Keeling, M. , Modelling optimal vaccination strategy for SARS-CoV-2 in the UK, *PLOS Computational Biology*, **17(5)** (2021).

[16] Moore, S., Hill, E., Tildesley, M. , Dyson, L., and Keeling, M., Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study, *The Lancet Infectious Diseases*, **21(6)** (2021), 793-802.

[17] Murray, J. D., *Mathematical Biology*, Springer-Verlag, Berlin, 1989.

medRxiv preprint doi: https://doi.org/10.1101/2021.12.18.21268024; this version posted December 18, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

[18] Perra N. , Non-pharmaceutical interventions during the COVID-19 pandemic: A review. *Physics Reports* (2021).

[19] Simula, E.R., et al., HCoV-NL63 and SARS-CoV-2 share recognized epitopes by the humoral response in sera of people collected pre-and during CoV-2 pandemic. *Microorganisms*, **8(12)** (2020), 1993.

[20] Tang, B., et al., Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions, *J. Clin. Med.*, **9** (2020), 462. doi: 10.3390/jcm9020462.

[21] Thieme, H.R., *Mathematics in Population Biology*, Princeton University Press, Princeton, NJ, 2003.

[22] Webb, G., A COVID-19 epidemic model predicting the effectiveness of vaccination, *Mathematics in Applied Sciences and Engineering*, (2021), 1-15.

[23] Webb, G., A COVID-19 epidemic model predicting the effectiveness of vaccination in the US, *Infectious Disease Reports*, **13(3)** (2021), 654-667.

[24] Wu, J. T., et al., Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China, *Nat. Med.*, **26** (2020), 506–510. doi: 10.1038/s41591-020-0822-7.

[25] Yahi, N., Chahinian, H., and Fantini J., Infection-enhancing anti-SARS-CoV-2 antibodies recognize both the original Wuhan/D614G strain and Delta variants. A potential risk for mass vaccination?. *Journal of Infection* (2021).

[26] New York City Department of Health and Mental Hygiene https://www1.nyc.gov/site/doh/covid/covid-19-data.page (accessed on 17 December 2021).

[27] United States census bureau https://www.census.gov/en.html (accessed on 17 December 2021).