How to coordinate vaccination and social distancing to mitigate SARS-CoV-2 outbreaks

Sara Grundel†  Stefan Heyder‡  Thomas Hotz‡  Tobias K. S. Ritschel†  Philipp Sauerteig‡  Karl Worthmann‡

December 22, 2020

Abstract

The world is waiting for a vaccine to mitigate the spread of SARS-CoV-2. However, once it becomes available, there will not be enough to vaccinate everybody at once. Therefore, vaccination and social distancing has to be coordinated. In this paper, we provide some insight on this topic using optimization-based control on an age-differentiated compartmental model. For real-life decision making we investigate the impact of the planning horizon on the optimal vaccination/social distancing strategy. We find that in order to reduce social distancing in the long run without overburdening the intensive care units it is essential to vaccinate the people with the highest contact rates first. However, for short-term planning it is optimal to focus on the high-risk group. Furthermore, large amounts of a vaccine with a lower success rate allows for more reduction of the social distancing than smaller amounts of a vaccine with higher success rate.

1 Introduction

In early 2020, the outbreak of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was declared a pandemic by the World Health Organization [50]. As the virus causes the respiratory illness coronavirus disease 2019 (COVID-19), many countries have been enforcing nonpharmaceutical countermeasures such as social distancing (also called contact restrictions) and travel restrictions [28, 29]. At the time of submitting this manuscript, strict contact restrictions are still in effect. Such countermeasures have a severe negative impact on national and international economies [40] and the general quality of life (in particular, mental health). Consequently, significant effort has been made to develop and deploy vaccines, and vaccination has already started in several countries, e.g. in the US.
and UK. Although manufacturers expect to supply substantial amounts of vaccines during 2021 [36], experts warn that nonpharmaceutical measures will remain necessary [42]. Therefore, there is still a need for identifying strategies for safely relaxing these nonpharmaceutical measures.

A variety of approaches has been proposed for studying the impact of nonpharmaceutical measures on the spread of SARS-CoV-2. They range from network-based over game-theoretical to data-driven methods [41, 35, 10]. One of the most popular methodologies are so-called compartmental models consisting of difference or differential equations [26]. Optimal control of such models is an active research topic, and both generalized interaction models [4] time-optimal policies [8, 25], and stochastic compartmental models [48] have been considered. Additionally, optimal mitigation policies have been proposed for several diseases, including dengue fever [18] and malaria [39]. Throughout 2020, many researchers have proposed compartmental models for predicting the impact of countermeasures on the spread of SARS-CoV-2 [20, 49]. Furthermore, they can be used in computing time-varying mitigation policies by solving optimal control problems (OCPs). Additionally, due to the significant uncertainty surrounding SARS-CoV-2 and COVID-19, it is necessary to repeatedly update these mitigation policies based on newly available information on the current state of the pandemic. This includes newly available parameter values, which may change as the pandemic evolves. Specifically, the initial condition in the OCP is updated and the prediction horizon is shifted, i.e. the horizon recedes as new information becomes available. The updated OCP is solved, and the first part of the solution is implemented. This is referred to as model predictive control (MPC) [24, 45], see also [14] for continuous-time systems and [53] for the relation between continuous- and discrete-time systems. MPC is a well-established control methodology which has been applied successfully in several fields [19, 34, 44]. For applications of MPC in, e.g. power electronics, we refer to [47, 15], in robotics to [52] and the references therein, and in energy to [9, 30]. A suitable choice of the prediction horizon is essential, see e.g. [51, 17].

Many authors have proposed SARS-CoV-2 mitigation strategies based on optimal control of nonpharmaceutical countermeasures (in particular, social distancing) [12, 31, 46, 22]. Nonpharmaceutical measures were used to attain a stable equilibrium with low case numbers in [13]. In [6], the authors propose an on-off (also called bang-bang) social distancing policy for mitigating a second wave. The subject of optimal vaccination has also been considered [37]. Both time-invariant and time-varying optimal vaccination policies have been presented. The question of where to vaccinate has also been addressed [21], e.g. should some cities or regions be prioritized over others. However, MPC has, to the best of our knowledge, not been used to compute time-varying vaccination policies. Matrajt et al. [33] present age-targeted vaccination policies in the absence of social distancing. They either consider the vaccination to be completed instantaneously at the initial time or assume constant vaccination rates. When minimizing deaths, they observe that for low vaccine efficacy, it is optimal to vaccinate the elderly. For high vaccine efficacy, and when sufficiently many are vaccinated, it is optimal to vaccinate the younger age groups which account for the most transmissions. Similarly, Hogan et al. [27] consider optimal age-targeted vaccination policies where the entire age group is assumed to be vaccinated at a constant rate over the course of one month. Early work on optimal vaccination involved age-uniform policies, and on-off poli-
cies were found to be optimal [32]. Similarly, Acuna-Zegarra et al. [2] use optimal control to compute time-varying vaccination policies and find intense vaccination for a limited time period to be optimal. Buckner et al. [11] optimize time-varying age-targeted vaccination policies. In particular, they account for essential workers, e.g. health care professionals, which are unable to significantly reduce their social interaction. They find that, depending on the objective function, either 1) younger essential workers are prioritized in order to control the spread of SARS-CoV-2 or 2) senior essential workers are prioritized to control the mortality. Finally, Bertsimas et al. [5] present age- and region-differentiated vaccination policies tailored for a number of states in the US. They model the social distancing as a predefined function of time which is fixed in advance. However, to the best of our knowledge, simultaneous optimal vaccination and social distancing has not been considered previously.

In this work, we address four key questions related to coordinating social distancing and vaccination: 1) How important are the availability and the success rate of the vaccine? 2) Who should be vaccinated first? 3) How much can social distancing measures be relaxed once the vaccines are available? 4) What is the minimal prediction horizon in the optimization step that recovers the qualitative features of the long-term policy? In order to address these questions, we present a novel compartmental model, which extends a recently developed model [23] to account for vaccination. Then, we use the new model together with optimal control and MPC to compute simultaneous vaccination and social distancing policies. The model accounts for vaccination failure, different levels of symptom severity, and age-dependent characteristics of SARS-CoV-2 and COVID-19. In our case study, we choose parameters tailored to the COVID-19 outbreak in Germany. However, we expect that the conclusions carry over to other developed countries.

We observe that it is optimal to first vaccinate the middle-aged group which spreads the disease the most. Subsequently, the elderly (the high-risk group) are vaccinated. In this case, the contact restrictions can be lifted almost half a year earlier than without vaccination. They can be lifted even earlier by increasing the number of successful vaccinations, which depends on both the number of available vaccines and their success rate. Conversely, the maximal ICU occupancy, which is closely related to the number of fatalities, can be reduced by prolonging the contact restrictions instead (without making them more strict). These conclusions are based on a prediction horizon of 2 years in the optimization step. We demonstrate that the same conclusions hold if a prediction horizon of at least 8 weeks is used. For too short prediction horizons, the elderly are vaccinated first. However, this is short-sighted and requires more strict contact restrictions than if the middle-age group is vaccinated first.

The remainder of this paper is structured as follows. We present the compartmental model in Section 2 and describe the optimal control problem in Section 3. Section 4 is dedicated to the case study, and the paper is concluded in Section 5.
2 A Compartmental Model with Social Distancing and Vaccination

In this section, we extend the dynamical model tailored to COVID-19 proposed in [23]. The aim is to be able to evaluate the effect of vaccination. Different vaccines have different properties. For simplicity, we focus on active vaccination, i.e. the body is triggered in order to produce antibodies itself. As a consequence, this kind of vaccination yields immunity but only if the patient has not been infected at time of vaccination. Still, there might be patients whose bodies do not produce (a sufficient amount of) antibodies. We assume that everyone has the same probability of vaccination failure and that a second try would yield the same outcome. Therefore, we allow vaccination at most once per person.

These considerations motivate the following assumptions.

(A1) Everyone who is not known to have been infected can be vaccinated. Vaccination can only be successful for people who have not been infected.

(A2) No one is vaccinated twice.

(A3) Each vaccination (of a non-infected person) has the same probability of failure.

(A4) Successful vaccination yields immediate immunity, i.e. you cannot get infected any longer.

In [23], we proposed a SEITPHR model consisting of 11 compartments which are divided into \( n_g \) age groups, \( n_g \in \mathbb{N} \). The compartments account for susceptible (\( S \)), exposed (or latent) (\( E \)), infectious with three different courses of severity (\( I^S \), \( I^M \), and \( I^A \)), tested (\( T \)), hospitalized with and without requiring an intensive care unit (ICU) (\( P \) and \( H \)), and detected and undetected removed people (\( R^K \) and \( R^U \)). Here, the superscripts \( S \), \( M \), and \( A \) indicate whether a person has a severe course of infection, i.e. he or she will go to an ICU at some point in time, a mild course, i.e. he or she will show symptoms and be put into quarantine and therefore will be removed, or an asymptomatic one, i.e. he or she will not be detected at all. Furthermore, the term removed captures all people who neither infect others nor need an ICU in the future, i.e. recovered, deceased, and quarantined people without severe infection. In the following, we describe the extended model, which is also shown in Figure 1. In order to emphasize the effect of vaccination in combination with social distancing, we neglect mass testing in this paper and, thus, drop the \( T \) compartments. However, keeping Assumption (A2) in mind, we split the SEIPHR model into two parts: The first one describes the compartments of people who have not been vaccinated while the second accounts for vaccinated people only. For clarity the latter are marked by an additional superscript \( V \).

Furthermore, we collect all infectious people in age group \( i \in \{1, \ldots, n_g\} \) at time \( t \geq 0 \) via

\[
I_i(t) = I^S_i(t) + I^M_i(t) + I^A_i(t) + I^{S,V}_i(t) + I^{M,V}_i(t) + I^{A,V}_i(t).
\]
The non-vaccinated part of the dynamics reads as

\[
\dot{S}_i(t) = -\sum_{j=1}^{n_g} \delta(t) \beta_{ij} S_i(t) I_j(t) - \nu_i(t) S_i(t) \tag{1a}
\]

\[
\dot{E}_i(t) = \sum_{j=1}^{n_g} \delta(t) \beta_{ij} S_i(t) I_j(t) - (\gamma + \nu_i(t)) E_i(t) \tag{1b}
\]

\[
\dot{I}_i^S(t) = \pi_i^S \gamma E_i(t) - (\eta^S + \nu_i(t)) I_i^S(t) \tag{1c}
\]

\[
\dot{I}_i^M(t) = \pi_i^M \gamma E_i(t) - (\eta^M + \nu_i(t)) I_i^M(t) \tag{1d}
\]

\[
\dot{I}_i^A(t) = \pi_i^A \gamma E_i(t) - (\eta^A + \nu_i(t)) I_i^A(t) \tag{1e}
\]

\[
\dot{R}_i^U(t) = \eta^A I_i^A(t) - \nu_i R_i^U(t) \tag{1f}
\]

\[
\dot{P}_i(t) = \eta^S I_i^S(t) - \rho P_i(t) \tag{1g}
\]

\[
\dot{H}_i(t) = \rho P_i(t) - \sigma H_i(t) \tag{1h}
\]

\[
\dot{R}_i^K(t) = \eta^M I_i^M(t) + \sigma H_i(t) \tag{1i}
\]

where the controls \( \delta : [0, \infty) \to [0, 1] \) incorporate the average contact reduction as well as the transmission probability and \( \nu_i : [0, \infty) \to \mathbb{R} \) denotes the vaccination rate within age group \( i \). Following Assumption (A1) we do not allow to vaccinate people who have been detected. The parameter \( 1 - q \in [0, 1] \) describes the above mentioned probability of a vaccination failure, see Assumption (A3). From here on, we refer to \( q \) as the success rate. Furthermore, compartments \( H \) and \( P \) collect all people in ICUs and those who have been detected, but do not yet require intensive care, respectively. The parameters \( \beta = (\beta_{ij})_{i,j=1}^{n_g} \), \( \gamma = (\eta^S, \eta^M, \eta^A) \), \( \rho \), and \( \sigma \) describe the transmission rates from one compartment to another while \( \pi_i^S \), \( \pi_i^M \), and \( \pi_i^A \) denote the age-dependent probabilities of having a severe, mild, or asymptomatic course of infection. For a more detailed description of the parameters, we refer to [23].

The vaccinated part is then given by

\[
\dot{S}_i^V(t) = (1 - q) \nu_i(t) S_i(t) - \sum_{j=1}^{n_g} \delta(t) \beta_{ij} S_i^V(t) I_j(t) \tag{2a}
\]

\[
\dot{E}_i^V(t) = \nu_i(t) E_i(t) + \sum_{j=1}^{n_g} \delta(t) \beta_{ij} S_i^V(t) I_j(t) - \gamma E_i^V(t) \tag{2b}
\]

\[
\dot{I}_i^{SV}(t) = \nu_i(t) I_i^S(t) + \pi_i^S \gamma E_i^V(t) - \eta^S I_i^{SV}(t) \tag{2c}
\]

\[
\dot{I}_i^{MV}(t) = \nu_i(t) I_i^M(t) + \pi_i^M \gamma E_i^V(t) - \eta^M I_i^{MV}(t) \tag{2d}
\]

\[
\dot{I}_i^{AV}(t) = \nu_i(t) I_i^A(t) + \pi_i^A \gamma E_i^V(t) - \eta^A I_i^{AV}(t) \tag{2e}
\]

\[
\dot{P}_i^V(t) = \eta^S I_i^{SV}(t) - \rho P_i^V(t) \tag{2f}
\]

\[
\dot{H}_i^V(t) = \rho P_i^V(t) - \sigma H_i^V(t) \tag{2g}
\]

\[
\dot{R}_i^K(t) = \nu_i(t) R_i^K(t) + q \nu_i(t) S_i(t) + \eta^A I_i^{AV}(t) + \eta^M I_i^{MV}(t) + \sigma H_i^V(t), \tag{2h}
\]

where the transmission is subject to Assumptions (A1) and (A4). Note that we combine \( R_i^V = R_i^{KV} + R_i^{UV} \) since we do not need to distinguish between known and unknown removed cases once they are vaccinated.
Figure 1: Flow of the SEIPHR model for the $i$-th age group. The controls associated with social distancing are indicated with dashed red edges, the controls associated with vaccination are depicted as dotted blue edges.
We emphasize that \( q \) does not represent the efficacy often mentioned in media, as e.g. \cite{42}. The latter considers two test groups – one which gets the vaccine and another which gets a placebo. Then, only the patients who show symptoms are tested and from these, one infers the efficacy \cite{43}. However, in our model, the parameter \( q \) describes the probability of the patient being immune after the vaccination.

For a concise notation, we collect all states in \( x(t) \in \mathbb{R}^n \), controls in \( u(t) \in \mathbb{R}^m \), and parameters in \( p \in \mathbb{R}^\ell \) and write
\[
\dot{x}(t) = f(x(t), u(t), p), \quad x(0) = x^0
\]
with initial value \( x^0 \in \mathbb{R}^n \). Furthermore, all compartments describe fractions of the total population, i.e. \( \sum_{i=1}^n x_i(t) = 1 \) for all \( t \geq 0 \), and the proportion of age group \( i \) is denoted by \( N_i, i \in \{1, \ldots, n_g\} \).

### 3 Optimal Vaccination Strategy

In this section, we formulate an OCP to determine a coordinated social distancing and vaccination strategy that reduces the required social distancing while maintaining an ICU cap. To this end, we assume an amount of \( V_{\text{max}} \), \( V_{\text{max}} \in \mathbb{N} \), units of the vaccine to become available each day, i.e. at time \( t \geq 0 \), the vaccine distribution is subject to
\[
n_{\text{pop}} \cdot \int_0^t \sum_{i=1}^{n_g} \nu_i(s)V_i(s) \, ds \leq V_{\text{max}} t, \tag{3}
\]
where
\[
V_i(s) = S_i(s) + E_i(s) + I_i^s(s) + I_i^m(s) + I_i^a(s) + R_i^U(s)
\]
collects all people in age group \( i \) available for vaccination at time instant \( s \), and \( n_{\text{pop}} \in \mathbb{N} \) denotes the total population. Furthermore, we penalize social distancing by minimizing the objective function
\[
J(\delta) = \int_0^t (1 - \delta(t))^2 \, dt.
\]
Figure 2 provides some intuition for \( J \).

The OCP is then given by
\[
\min_{(\delta, \nu)} \quad J(\delta) + \kappa \|\nu\|_2^2 \tag{4a}
\]
subject to
\[
n_{\text{pop}} \cdot \sum_{i=1}^{n_g} H_i(t) + H_i^V(t) \leq H_{\text{max}} \tag{4b}
\]
\[
\dot{x}(t) = f(x(t), u(t), p), \quad x(0) = x^0 \tag{4c}
\]
\[
\delta(t) \in [0, 1] \tag{4d}
\]
\[
n_{\text{pop}} \cdot \int_0^t \sum_{i=1}^{n_g} \nu_i(s)V_i(s) \, ds \leq V_{\text{max}} t, \quad \forall t \geq 0 \tag{4e}
\]
where constraint (4b) caps the total number of required ICU beds. The positive parameter \( \kappa \ll 1 \) weights the regularization term, which ensures that the optimal solution is smooth and vanishes once the ICU cap is maintained without social distancing and vaccination. We assume the controls to be constant over one week.
Figure 2: The three steps of evaluating the objective function. The dashed black line corresponds to no social distancing and the objective is to minimize the blue shaded area. We minimize the area under the squared deviation in order to discourage very strict contact restrictions.

4 Case Study

In this case study, we investigate

1) the importance of vaccine availability and efficacy,

2) whom to vaccinate first,

3) how much we can relax social distancing while distributing the vaccines,

4) the difference between short- and long-term planning

by analyzing numerical solutions to the OCP (4). We solve the OCPs using a direct single-shooting approach [7] combined with the standard sequential quadratic programming [38] algorithm implemented in fmincon in Matlab. Furthermore, we compute the gradients of the left-hand side of the nonlinear inequality constraint (3) using a continuous forward method (i.e. we numerically integrate the sensitivity equations forward in time). For convenience, we refer to contact reductions by up to 20% as light, between 20% and 60% as strict, and we consider reductions by over 60% a lockdown.

First, we consider long-term open-loop solutions (i.e. where the strategies are not repeatedly updated). Specifically, we compare solutions for different values of the vaccination success rate, $q$, the number of vaccines supplied each day, $V_{max}$, and the ICU capacity, $H_{max}$.

Note that all simulations come along with uncertainties, e.g. resulting from not modelled effects, inaccurate parameters, new developments. This model-plant mismatch particularly causes problems for long-term simulations. In the context of mitigation of COVID-19 it is essential to update model parameters based on newly acquired data in order to develop adequate interventions. To this end, we also analyse short- to medium-term closed-loop solutions (where the strategies are updated repeatedly on newly available measurements). This emulates the real-life decision process where mitigation strategies are continuously updated when new data becomes available.

Throughout the simulations, we use the (fixed) parameters shown in Table 1, see [23] for a more thorough description.
Table 1: Parameter values.

<table>
<thead>
<tr>
<th>Description</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total population</td>
<td>(n_{\text{pop}})</td>
<td>(8.3 \cdot 10^7)</td>
</tr>
<tr>
<td>Number of age groups</td>
<td>(n_{\text{g}})</td>
<td>3</td>
</tr>
<tr>
<td>Regularization parameter</td>
<td>(\kappa)</td>
<td>(10^{-3})</td>
</tr>
<tr>
<td>Removal rate (severe)</td>
<td>(\eta^S)</td>
<td>0.2500</td>
</tr>
<tr>
<td>Removal rate (mild)</td>
<td>(\eta^M)</td>
<td>0.2500</td>
</tr>
<tr>
<td>Removal rate (asymptomatic)</td>
<td>(\eta^A)</td>
<td>0.1667</td>
</tr>
<tr>
<td>Infection rate</td>
<td>(\gamma)</td>
<td>0.1923</td>
</tr>
<tr>
<td>ICU admittance rate</td>
<td>(\rho)</td>
<td>0.0910</td>
</tr>
<tr>
<td>ICU discharge rate</td>
<td>(\sigma)</td>
<td>0.0952</td>
</tr>
</tbody>
</table>

**Age-differentiated parameters**

<table>
<thead>
<tr>
<th>Age group</th>
<th>(i)</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age range (in years)</td>
<td>(N_i)</td>
<td>(&lt;15)</td>
<td>15 – 59</td>
<td>(&gt;60)</td>
</tr>
<tr>
<td>Relative age group size</td>
<td>(N_i)</td>
<td>0.1370</td>
<td>0.5776</td>
<td>0.2854</td>
</tr>
<tr>
<td>Probability of severe symptoms</td>
<td>(\pi^S_i)</td>
<td>0.0053</td>
<td>0.0031</td>
<td>0.0302</td>
</tr>
<tr>
<td>Probability of mild symptoms</td>
<td>(\pi^M_i)</td>
<td>0.1211</td>
<td>0.2201</td>
<td>0.2512</td>
</tr>
<tr>
<td>Probability of no symptoms</td>
<td>(\pi^A_i)</td>
<td>0.8737</td>
<td>0.7768</td>
<td>0.7186</td>
</tr>
<tr>
<td>Transmission rate (age group 1)</td>
<td>(\beta_{1i})</td>
<td>0.4612</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transmission rate (age group 2)</td>
<td>(\beta_{2i})</td>
<td>0.4819</td>
<td>0.6304</td>
<td></td>
</tr>
<tr>
<td>Transmission rate (age group 3)</td>
<td>(\beta_{3i})</td>
<td>0.1243</td>
<td>0.2944</td>
<td>0.1802</td>
</tr>
</tbody>
</table>

### 4.1 Long-Term Simulations

Results for the OCP (4) can be found in Figures 3 and 4. Here, we chose \(q = 0.9\) since the general expectation is that the success rate is quite high. Furthermore, \(V^{\text{max}} = 100,000\) and \(H^{\text{max}} = 10,000\) are based on [1, 16]. The vaccination constraint (3) is active until the contact restrictions are lifted. At that point, the pandemic is contained without further interventions, i.e. after approximately 40 weeks.

[Figure 3: Optimal vaccination strategy \(\nu^* = (\nu^*_1, \nu^*_2, \nu^*_3)\) with vaccination success rate \(q = 0.9\) and daily available units \(V^{\text{max}} = 10^5\); the dotted black line depicts \(V^{\text{max}}t\). The vaccination process is stopped once all contact restrictions are lifted (Figure 4) and the pandemic is contained without further interventions, i.e. after approximately 40 weeks.]

The objective of this is to minimize social distancing. Specifically, it allows us to relax the social distancing measures earlier while still maintaining the hard infection cap modelled by the upper bound on the ICU capacity. Hence, the elderly (high-
Figure 4: Optimal social distancing policy $\delta$ (left) and ICU occupancy $H = \sum_{i=1}^{n_g} H_i + H_i^V$ (right) with and without vaccine. The dotted black lines mark strict social distancing, i.e. $\delta \in [0.4, 0.8]$ (left) and $H_{\text{max}}$ (right).

Risk group) get vaccinated once the social distancing is significantly reduced, i.e. around week 36. In contrast, an actual lockdown is advisable when no vaccine is available. Without vaccination, strict social distancing is required for 43 weeks. With vaccination, the strict social distancing can be lifted approximately 20 weeks earlier, i.e. a reduction by almost 50%. The same conclusion applies to the ICU occupancies.

Figure 5 (left) shows the optimal social distancing strategies for different ICU capacities, $H_{\text{max}}$. Interestingly, the maximum amount of contact restrictions, $\delta = 0.435$, is (almost) independent of $H_{\text{max}}$. Consequently, it is possible to reduce the total number of people who become admitted to ICUs at the cost of longer but not stricter social distancing measures. This is particularly important as the number of people in ICUs is closely related to the number of fatalities. Figure 5 (right) visualizes the impact of the number of available ICUs on the number of weeks where contact restrictions have to be enforced.

Moreover, the solutions with vaccination all enforce contact restrictions from the beginning whereas the solution without vaccination lets the pandemic evolve for some weeks before implementing a strict lockdown. We explain this using Figure 6, which shows the optimal social distancing strategies with (◦) and without (+) vaccination (also shown in Figure 4). We compare these two strategies to 1) enforcing contact restrictions in the beginning ($\delta = 0.8$) without any vaccination ($\triangle$) and 2) prohibiting them in the beginning ($\delta = 1$) and allowing vaccination (⋆). The key observation is that without social distancing in the beginning, a hard lockdown is necessary regardless of whether a vaccine is available or not. However, when a vaccine is available, some social distancing in the beginning avoids a harder lockdown later because a significant amount of people are already vaccinated at this point. Furthermore, if no vaccine is available, it is not beneficial to enforce...
Figure 6: Impact of enforcing/prohibiting social distancing during the initial phase with $H^\text{max} = 10^4$ on the optimal social distancing strategy $\delta$ (left) and total amount, $J(\delta)$, of contact restrictions (right).

Social distancing early on because it only slows down the natural vaccination (i.e. the immunity following an infection).

According to [1], a realistic number of daily vaccines is $V^\text{max} = 100,000$. However, at the time of submitting this manuscript, no actual numbers are available. For this reason, we investigate the impact of varying both the available number of vaccines as well as the success rate on the social distancing and on the end time of the contact restrictions. The results depicted in Figure 7 show that it is more important to increase the available amount of the vaccines than the success rate. Note that when increasing $V^\text{max}$ (right), there is a threshold where the required amount of strict social distancing increases before it decreases again. The reason can be seen in Figure 8. For $V^\text{max} = 200,000$, the contact reductions become classified as strict in the beginning whereas for lower and higher values, the contact reductions are still classified as light.

Figure 7: Impact of vaccination success rate $q$ (left) and daily available units of vaccine $V^\text{max}$ (right) on the total time (severe) contact restrictions have to be implemented with ICU cap $H^\text{max} = 10^4$.

Figure 8: Impact of the amount of daily available units of vaccine $V^\text{max}$ on the optimal social distancing strategy $\delta$ with vaccination success rate $q = 0.9$ and ICU cap $H^\text{max} = 10^4$.

Furthermore, Figure 9 provides some insight on how to improve the impact of
a vaccine. For instance, if the supply rate is small, e.g. $V^{\max} = 12,500$, improving the success rate $q$ does not help reducing the contact restrictions. Once sufficiently many units of vaccine can be produced, it is possible to reduce the social distancing by increasing the success rate.

Based on the results shown in Figures 7 and 9 we conclude that increasing the number of successful vaccinations has the biggest impact on the social distancing policy. For instance, if $q = 0.5$ and $V^{\max} = 100,000$, doubling $V^{\max}$ has a bigger impact than increasing $q$ to 0.9. Doubling $V^{\max}$ adds another 50,000 successful vaccinations whereas increasing $q$ only adds another 40,000 successful vaccinations.

4.2 Consecutive Short-Term Simulations

In order to simulate real-life decision-making processes and to account for uncertainties as mentioned in the beginning of Section 4 we use MPC [14]. The main idea of MPC is to solve a sequence of OCPs of the form

$$\begin{align*}
\min_{(\delta, \nu)} & \int_{k\Delta t}^{(k+K)\Delta t} (1 - \delta(t))^2 \, dt + \kappa \| \nu \|^2_2 \\
\text{subject to} & \quad n_{\text{pop}} \cdot \sum_{i=1}^{n_{\text{pop}}} H_i(t) + H^V_i(t) \leq H^{\max} \\
& \quad \dot{x}(t) = f(x(t), u(t), p), \quad x(0) = x^0 \\
& \quad \delta(t) \in [0, 1] \\
& \quad n_{\text{pop}} \cdot \int_{k\Delta t}^{t} \sum_{i=1}^{n_{\text{pop}}} \nu_i(s)V_i(s) \, ds \leq V^{\max} \cdot (t - k\Delta t) + V^k \\
& \quad \forall t \in [k\Delta t, (k + K)\Delta t]
\end{align*}$$

over a moving time window of length $K\Delta t$, where $K \in \mathbb{N}_{\geq 2}$ denotes the number of time steps of length $\Delta t > 0$. Here, the parameter $V^k$ accounts for the units of vaccine that have been saved in previous MPC steps. This scheme can be summarized as follows.
for $k = 0$ to $K - 1$

1. Measure/estimate current state $x(k\Delta t) = x^k$ at the $k$-th time instant.

2. Solve the OCP (5) on the time window $[k\Delta t, (k+K)\Delta t]$ to get an optimal control $u^* : [k\Delta t, (k + K)\Delta t] \rightarrow \mathbb{R}^m$.

3. Implement first portion of solution $u^k|_{[k\Delta t, (k+1)\Delta t]} = u^*|_{[k\Delta t, (k+1)\Delta t]}$, increment $k \leftarrow k + 1$.

end

Here, $K = t_f/\Delta t$ denotes the number of MPC steps, i.e. the number of OCPs of the form (5) that have to be solved to arrive at a solution of (4). In our simulations, we set $\Delta t$ to one week.

We study the impact of different prediction horizon lengths $K$ on the closed-loop solution. This corresponds to a simulation of the whole pandemic, while the decision for the next week is always made based on a forecast horizon of only $K$ weeks. It is essential to choose the prediction horizon sufficiently long; otherwise, the ICU caps might be violated due to time delays within the model, e.g. caused by the incubation time. In our simulations, the smallest possible integer value for the prediction horizon $K$ that allows for maintaining the ICU cap is 3 weeks. The impact of the choice of $K \geq 4$ on the objective function value is visualized in Figure 10. The longer the prediction, the closer the objective function value gets to the one corresponding to the open-loop solution shown in the previous subsection. Moreover, the marginal gain of horizon lengths larger than eight weeks is negligible. Therefore, if a prediction horizon of eight weeks is used, the strategy does not suffer from the short-sightedness of operating on a limited time window (but with significantly reduced uncertainty) while being able to adapt to newly acquired data.

Optimal vaccination and social distancing strategies depending on the prediction horizon length are depicted in Figures 11 and 12 respectively. If the prediction horizon is small, it is optimal to vaccinate the high-risk group only (Figure 11), since they directly affect the number of required ICUs. Furthermore, the contact restrictions are severe, but are lifted comparatively early (Figure 12). For longer prediction horizons, the people with highest contact rates are vaccinated first, and the restrictions are not as severe. In particular, the vaccination strategies for $K = 8$ and $K = 12$ coincide. Furthermore, for these prediction horizons,
Figure 11: Impact of prediction horizon length $K$ on vaccination strategy. For short prediction horizons the high-risk groups are vaccinated first (immediate impact), for long prediction horizons the people with most contacts are vaccinated first in order to relax the social distancing in the long run.

Figure 12: Impact of prediction horizon length $K$ on social distancing and ICU capacities. Short prediction horizons might yield necessity of a (short) lockdown. The contact restrictions are not as strict as in the open-loop policy, they develop more smoothly, and they can be lifted monotonically. Note that with prediction horizon length $K = 12$ weeks, the social distancing around week 13 is quite relaxed ($\delta \geq 0.6$). As a result more restrictions have to be enforced around week 22.

5 Conclusions and outlook

In this paper, we address the four questions related to simultaneous vaccination and social distancing set out in the introduction. We address them by extending a previous compartmental model to include vaccination. Based on this model, we use optimal control and MPC to compute time-varying vaccination and social distancing policies. Our simulations show that contact restrictions can be lifted almost half a year earlier as compared to a scenario without vaccination. This is achieved by first vaccinating the middle-aged group which is most responsible for spreading the virus. Thereafter, the elderly, which are most vulnerable to COVID-19, are vaccinated. Furthermore, we find that the contact restrictions can be lifted even earlier by increasing the number of successful vaccinations, which depends on both the number of available vaccines and the efficacy. We also observe that the maximal ICU occupancy can be reduced by extending the contact restrictions and that it is not necessary to make them more strict. The above conclusions assume that long-term planning is possible, e.g. over a 2 year period. This is not possible in practice, but we demonstrate that optimal closed-loop policies obtained with an 8 week prediction horizon are qualitatively similar to those obtained with a 2 year horizon. However, if too short horizons are used, only the elderly are vaccinated, and more social distancing is necessary.

In future work, we will account for the uncertainty of the model parameters and use uncertainty quantification techniques to assess their impact on the policies.
presented in this paper.

Acknowledgments

We thank Martin Weissleder (BWK Berlin) for some insights on the medical process of vaccination.

References


