Forecasting the Spread of SARS-CoV-2 is inherently Ambiguous given the Current State of Virus Research

Koenen, Melissa; Balvert, Marleen; Brekelmans, Ruud; Stienen, Valentijn; Wagenaar, Joris

Publication date:
2020

Document Version
Early version, also known as pre-print

CentER

Discussion Paper

No. 2020-026

FORECASTING THE SPREAD OF SARS-CoV-2 IS INHERENTLY AMBIGUOUS GIVEN THE CURRENT STATE OF VIRUS RESEARCH

By

M.F. Koenen, M. Balvert, R. Brekelmans,
H.A. Fleuren, V.F. Stienen, J.C. Wagenaar

29 September 2020

ISSN 0924-7815
ISSN 2213-9532

TILBURG UNIVERSITY

Forecasting the spread of SARS-CoV-2 is inherently ambiguous given the current state of virus research

M.F. Koenen[1], M. Balvert[1], R. Brekelmans[1],
H.A. Fleuren[1], V.F. Stienen[1], J.C. Wagenaar*[1]

[1]Zero Hunger Lab,
Tilburg School of Economics and Management, Tilburg University
Department of Econometrics and Operations Research

* E-mail: j.c.wagenaar_1@tilburguniversity.edu

Abstract

Since the onset of the COVID-19 pandemic many researchers and health advisory institutions have focused on virus spread prediction through epidemiological models. Such models rely on virus- and disease characteristics of which most are uncertain or even unknown. This study addressed the validity of various assumptions using an epidemiological simulation model. We showed that multiple scenarios all lead to realistic numbers of deaths and ICU admissions, two observable and verifiable metrics, but gave different estimates for the number of infected and immune individuals. As these metrics are particularly important for policy makers, further research on virus and disease progression characteristics is essential. Until that time, epidemiological modeling studies cannot give conclusive results and should come with careful analysis of several scenarios on virus- and disease characteristics.

1. Introduction

The COVID-19 pandemic has disrupted society all across the world. At the time of the SARS-CoV-2 virus outbreak in Wuhan province, China went into lockdown. Many countries across the world followed when the virus reached them a few weeks or months later. Since then many researchers and national health institutions have focused on predicting the course of the epidemic, assessing the effects of non-medical interventions in the form of social distancing, and evaluating the possibilities of an exit strategy [e.g. 7, 20, 11]. The epidemiological models underlying these studies heavily rely on virus and disease characteristics such as the case fatality ratio (CFR). Within just a few months researchers made great progress in estimating these characteristics [e.g. 17, 15, 30, 29], and a plethora of data sources and scientific studies rapidly became available [2, 4, 18]. These sources however report various estimates on parameters that describe the virus behavior

and disease progression. As a result, many aspects of the SARS-CoV-2 virus' behavior that forecasting models rely on, including CFR, still remain uncertain or unknown.

The aim of this paper was to test the validity of several assumptions on four model parameters: the probability of developing symptoms, case fatality ratios, when people develop immunity, and the probability of virus transmission between an infected and a non-infected individual. We did so by simulating the spread of SARS-CoV-2 under a variety of assumptions on these four parameters using data from the Netherlands. Combinations of assumptions that lead to a predicted number of ICU admissions and death toll that resembled reality were considered plausible, while scenarios that lead to predicted ICU admissions and death toll that substantially differed from reality were considered unrealistic.

Several combinations of assumptions yielded a realistic number of estimated daily ICU admissions and deaths and were hence plausible scenarios. However, they gave different predictions in terms of unobservable yet important characteristics such as the level of immunity among the population. This means that our current knowledge on virus- and disease characteristics is insufficient for an epidemiological modeling study to give conclusive answers concerning disease spread.

2. Methods and Materials

The simulation model that we used to validate assumptions on virus spread and disease progression holds the middle between a compartment SEIR model and an agent-based simulation. Classical SEIR models have been commonly used to model disease spread and form the basis of many of today's COVID-19 epidemiological models [14, 26]. As the classical SEIR models cannot capture all relevant aspects of COVID-19 such as differentiation between age groups and geographic location, various adaptations have been proposed [21, 22, 12, 41, 7]. An often used alternative is agent-based simulations [11, 40, 19], which allow for modeling at the individual level rather than aggregating over the entire population. This is important when modeling COVID-19 [5, 24], as it allows for including social patterns that depend on age group and location.

We used a coarse-grained agent-based simulation model that differentiates between age groups and geographic regions. The model is akin to agent-based simulations [11, 40, 19] in that distinctive individuals are simulated who commute between their region of residence and region of employment. The main difference is that we did not include social interactions at an individual level but aggregated over groups of people with the same age, region of residence and work region. This allowed us to simulate on a large scale, i.e. to simulate all inhabitants of a country or state, while including individual characteristics such as age, region of residence and commute patterns that are highly relevant when modeling the spread of SARS-CoV-2 [8, 29].

The disease progression as modeled in classical SEIR models, assuming that a population consists of susceptible (`S`), exposed (`E`), infectious (`I`) and recovered (`R`) individuals,

is insufficient to reflect COVID-19 [14, 17]. We extended the disease progression with several disease stages (Figure 1). First, while classical SEIR models assume that infectious individuals are symptomatic and hence observable, for COVID-19 asymptomatic individuals can be infectious as well [16, 42, 13, 28]. We therefore split (`I`) into two subgroups: asymptomatic (`I-a`) and symptomatic (`I-s`). Second, it is unclear whether all infections lead to immunity [6, 14]. We tested the effects of assuming some infected individuals to not develop immunity but return to the susceptible group instead (dotted lines in Figure 1). In order to explicitly model recovered patients that obtained immunity, recovered (`R`) was replaced by two states: immune (`IM`) and deceased (`D`). Third, since we tested the validity of our model outcomes based on among others the number of daily ICU admissions, we included a stage "ICU admission" (`ICU-a`) [14, 11]. In the Netherlands, only patients with severe symptoms who have a chance of survival are admitted to the ICU [3]. For patients with severe symptoms but a low chance of survival we used the stage "ICU refusal" (`ICU-r`). Finally, we used the classical compartments susceptible (`S`) and exposed (`E`).

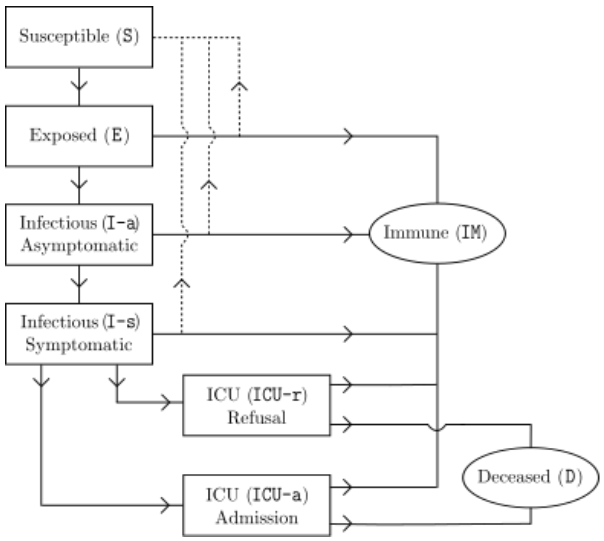

Figure 1: *Possible virus progression of disease stages in our simulation model.*

The remainder of this section is organized as follows. First an overview of the tested assumptions is given in Section 2.1. The commute and contact patterns used in this study are explained in Section 2.2, and the computation of the transmission probability, i.e. the probability that a non-infected individual gets infected when they meet an infectious individual, is explained in Section 2.3. Once infected, an agent goes through several health stages, which is explained in Section 2.4. Finally the initialization of the simulation model

is discussed in Section 2.5.

2.1. Scenarios

The aim of this study was to test the validity of four important disease characteristics and assumptions. First, we estimated the probability of developing symptoms after infection. A major fraction of infections is asymptomatic and often goes by unnoticed, thus this probability is largely unknown [16]. For this, we have tested four scenarios where the probability of developing symptoms is 0.375, 0.5, 0.625 or 0.75, representing a wide range of possibilities.

Second, we assessed three scenarios for case fatality ratios. In the first scenario, we estimated the probability that a symptomatic individual dies, $P(\texttt{D}|\texttt{I-s})$, as the ratio between the death toll and the number of symptomatic individuals as estimated for the Netherlands. We used the death toll reported by the Dutch National Institute for Public Health and Environment (RIVM) [25] combined with the excess death rates reported by the Dutch Statistics Bureau [32]. The number of symptomatic individuals was estimated using data reported by Sanquin, the Dutch blood bank, that tested donated blood for antibodies to estimate the fraction of the population that had been infected [37]. For details see the supplementary information in Appendix A. The second and third scenario were based on case fatality ratios per age group obtained from a study in China with approximately 72,000 cases [35]. The case counts in this study may either include all infections, or only symptomatic cases since these are observable. Therefore, in the second scenario, we assumed that this CFR reflects the death rate among all infected individuals, $P(\texttt{D}|\texttt{E})$, and in the third scenario we assumed that CFR reflects the fatality ratio among all symptomatic individuals, $P(\texttt{D}|\texttt{I-s})$.

Third, it is yet unknown which fraction of the non-lethal infections leads to immunity. For this we considered four scenarios. In the first scenario, all infections lead to immunity. The second scenario assumed that only those individuals who develop symptoms (`I-s`) become immune, while others return to the group of susceptible individuals (`S`). In the third and fourth scenario, we assumed that having symptoms leads to immunity in only 50% and 25% of the cases, respectively, while the remaining individuals return to the susceptible group. The latter three scenarios are indicated by the dotted lines in Figure 1.

Fourth, the virus transmission probability of an infectious individual encountering a susceptible individual is unknown. In total nine different probabilities were evaluated: 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50 and 0.55. Combined with the 48 scenarios for symptom development, CFR and developing immunity, this leads to 432 scenarios in total (see Table 1).

We used data from the Netherlands to simulate the period between February 27, 2020, when the first case was identified, and March 25, 2020, when the effects of the lockdown became visible in the number of ICU admissions and the death toll. Using only the initial

Table 1: Overview of tested scenarios. All 432 combinations are tested. E = Exposed, I-s = Infectious Symptomatic. Note that for the probability of developing immunity, we only considered infections that do not lead to death.

Input parameter	Scenario	Description
Probability of developing symptoms	0.375	37.5% go from `E` to `I-s`
	0.5	50% go from `E` to `I-s`
	0.625	62.5% go from `E` to `I-s`
	0.75	75% go from `E` to `I-s`
Case fatality ratio	Data from NL	$P(\texttt{D}\|\texttt{E})$ based on case counts and death toll in the Netherlands
	Literature-E	$P(\texttt{D}\|\texttt{E})$ based on [35]
	Literature-I-s	$P(\texttt{D}\|\texttt{I-s})$ based on [35]
Developing immunity	All	All infections lead to immune
	High	All symptomatic infections lead to immune
	Medium	50% of symptomatic infections lead to immune
	Low	25% of symptomatic infections lead to immune
Virus transmission probability		$\{0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55\}$

period of the outbreak gives the most accurate view of virus parameters (thus excluding the effect of protection measures).

2.2. Daily commute and contact patterns

The Netherlands is divided into 40 regions termed corops, following a statistical division of the Netherlands for research institutions to present their data ([1]). All approximately 17 million inhabitants of the Netherlands, termed "agents" in the simulation, have a known corop of residency and corop of employment. The population per corop per age group [33] and commute data between corop regions [31] were obtained from Statistics Netherlands. As the Netherlands is a small and densely populated country with many commuters, these corops are vastly interconnected. We assumed that agents who are unemployed stay in their corop of residency during the day.

The simulation divided each day into two epochs: during the day epoch, most inhabitants are at their work corop, and during the night epoch all agents are in their corop of residency. Each epoch an agent may meet other agents that reside in the same corop during that epoch, potentially leading to a new SARS-CoV-2 infection.

The daily contact pattern of an agent is determined by their age group and was obtained from [9] (Table 2). This paper reports the total number of daily contacts an individual of a certain age group has, whereas the age distribution of the people someone has contact with differs per age group. Consequently, the total number of contacts had to be divided over the different age groups. For this we used the percentage of contacts each age group has with another age group obtained from [38], by converting the amount

of contacts found in that study into percentages. Combining these percentages with the total number of contacts per age group gave the contact pattern for each age group as shown in Table 3.

Table 2: Daily contact data obtained from [10].

Age group	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-150
# of daily contacts	12.5	16.1	21.2	21.8	22.1	20.9	15.4	10	9.5

Table 3: Social patterns obtained by combining [10] and [38]. The table shows the daily number of contacts an individual whose age group is found on one of the rows has with people from an age group found on one of the columns. For example, someone in the age group 30-39 has on average contact with 2.99 people from the age group 20-29 per day.

Age group	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-150
0-9	4.75	2.25	1.36	1.36	0.70	0.70	0.45	0.45	0.45
10-19	1.93	6.04	1.87	1.87	1.34	1.34	0.56	0.56	0.56
20-29	1.97	1.99	2.90	4.34	2.84	2.84	1.44	1.44	1.44
30-39	2.03	2.05	2.99	4.40	2.92	2.92	1.48	1.48	1.48
40-49	1.02	1.33	2.41	3.16	3.93	3.93	2.12	2.12	2.12
50-59	0.96	1.25	2.28	2.99	3.72	3.72	2.01	2.01	2.01
60-69	0.60	0.60	0.79	1.32	1.74	1.74	2.86	2.86	2.86
70-79	0.39	0.39	0.51	0.86	1.13	1.13	1.86	1.86	1.86
80-150	0.37	0.37	0.48	0.817	1.07	1.07	1.77	1.77	1.77

2.3. Transmission probability

At every time epoch the model determined for each susceptible agent whether they got infected based on their location and age group. The probability to get infected depends on the location of the agent, the number of (infected) other agents present in that location, the contact pattern of the agent, and the probability of transmission in case the agents interacts with an infectious agent.

In mathematical terms, we defined the infection probability $p_{a,c,t}$ as the probability that a susceptible agent from age group $a \in A$ gets infected while being in corop $c \in C$ at epoch $t = 1, \cdots, T$. This infection probability can be determined as follows:

$$p_{a,c,t} = 1 - \prod_{a' \in A} (1 - p_{a,a',c,t})^{[\#\text{EpochContacts}]_{a,a'}},$$

where $[\#\text{EpochContacts}]_{a,a'}$ is the number of contacts an agent of age group a has with agents of age group a' during an epoch (see Section 2.2) and $p_{a,a',c,t}$ represents the probability that a susceptible individual from age group $a \in \mathcal{A}$ gets infected through an encounter with an agent from age group $a' \in \mathcal{A}$, in corop $c \in \mathcal{C}$, at epoch $t = 1, \ldots, T$.

The definition of the infection probability was chosen such that it allows for region dependent probabilities as well as social patterns that depend on age groups. This makes the model flexible and realistic: suppose that in a specific corop many agents of a given age group are infected, then an agent who resides in that specific corop and meets many individuals of that age group has a high risk of getting infected.

The probability that a susceptible individual from age group a gets infected through an encounter with an individual of age group a', in corop c at epoch t ($p_{a,a',c,t}$) consisted of three components:

$$p_{a,a',c,t} = \mathbb{P}\{E_{aa'}\}\ \mathbb{P}\{I_{a',c,t}\}\ \mathbb{P}\{T\},$$

We assumed independence between these components.

- $\mathbb{P}\{E_{aa'}\}$ represented the probability that an individual from age group a encounters an individual from age group a'. See Section 2.2 on how to obtain this probability from the contact patterns.

- $\mathbb{P}\{I_{a',c,t}\}$ described the fraction of individuals in age group a' in corop c that was contagious at time t. This fraction was determined for each epoch separately, and was computed by dividing the number of contagious agents in age group a' in corop c at time t by the total number of agents of age group a' in corop c at time t. For the infectious agents we only included those that were in health stages `I-a` or `I-s`, as individuals with symptoms so severe that ICU admission is necessary (either `ICU-a` or `ICU-r`) were assumed to be too severely ill to (be allowed to) meet others.

- $\mathbb{P}\{T\}$ denoted the probability of virus transmission when a susceptible individual encounters an infectious individual. Since this is an unknown parameter, we tested several values for $\mathbb{P}\{T\}$, as shown in Table 1.

2.4. Disease progression

The progression of an exposed agent's health stage from one epoch to the next was assumed to depend only on the agent's current health stage and his or her age group. Hence, after being infected, the agent's health stage over time can be interpreted as a discrete time Markov chain where transitions can occur according to Figure 1. This Markov chain is represented by a transition matrix containing the transition probabilities of an agent moving from one health stage in a certain epoch to another health stage in the next epoch. Well-known results from Markov chain analysis [e.g. 27] were used to compute these probabilities such that certain pre-imposed properties were satisfied.

Two types of properties were used to compute the transition matrices. First, we used the probability that an agent will eventually reach stage j at some point in time, given that the agent's current state is i. For example, for $i =$ `E` and $j =$ `D`, $P(\texttt{D}|\texttt{E})$ is the probability that an exposed agent will eventually decease. This corresponds to the CFR, for which we tested three scenarios as discussed in Section 2.1. We also used assumptions on $P(\texttt{D}|\texttt{ICU-a})$ and $P(\texttt{ICU-a}|\texttt{I-s})$. Data regarding ICU admissions and number of deaths

at the ICU [34] was used to construct $P(\texttt{D}|\texttt{ICU-a})$. Furthermore, data from Sanquin on the percentage of their blood donors that had COVID-19 antibodies [37] (see Appendix A) was used to obtain an estimation of the number of infections in each age group. Combining the Sanquin data with the number of ICU admissions we estimated $P(\texttt{ICU-a}|\texttt{I-s})$. The resulting probabilities that were used for estimating the transition matrices are listed in Table 4.

Age group	$P(\texttt{D}\|\texttt{ICU-a})$	$P(\texttt{ICU-a}\|\texttt{I-s})$
0–9	0.32×10^{-2}	0.00
10–19	2.88×10^{-2}	0.66×10^{-4}
20–29	7.99×10^{-2}	1.47×10^{-4}
30–39	15.67×10^{-2}	4.34×10^{-4}
40–49	25.90×10^{-2}	11.22×10^{-4}
50–59	38.60×10^{-2}	26.85×10^{-4}
60–69	54.03×10^{-2}	77.03×10^{-4}
70–79	71.93×10^{-2}	187.42×10^{-4}
80+	92.39×10^{-2}	41.86×10^{-4}

Table 4: Age-dependent probability properties common to all scenarios.

The second property type considered was duration, i.e. the expected time an agent spends in a health stage or a collection of health stages, given the current health stage. The values used are listed in Table 5. For instance, a well-known quantity from literature is the incubation time, i.e., the time from infection until developing symptoms. This incubation time corresponds to the expected time spent in health stages `E` and `I-a` combined, given that an agent's current health stage is `E` (represented in the table by $E(\texttt{E} \rightarrow \texttt{I-s})$). The average incubation time was estimated using an aggregated dataset of clinical outcomes [4] by taking the average of the reported values from the 20 studies with the largest population sizes. The average duration from symptoms to ICU admission ($E(\texttt{I-s} \rightarrow \texttt{ICU-a}), E(\texttt{I-s} \rightarrow \texttt{ICU-r})$) was estimated based on an average from multiple papers in literature, among which [4] and [39].

Recall that we distinguished between age groups, as many properties of COVID-19, such as the probabilities mentioned above, depend on the patient's age. Therefore, a transition matrix was estimated for each age group. The Markov properties can be used to fit a transition matrix to empirical values by minimizing a measure of goodness of fit on these properties. We used the sum of squared errors for this purpose after weighting the probability properties by a factor 100 to ensure balanced scaling between probability and duration properties.

The properties combined with the transition structure from Figure 1 do not uniquely determine the transition matrix. We therefore made the additional assumption that apart from the target properties, the transition matrices of the age groups should be as similar as

Stage(s)	Description	Literature	Value
$E(\texttt{E} \rightarrow \texttt{I-s})$	Average incubation time	[4]	5.758
$E(\texttt{I-a} \rightarrow \texttt{I-s})$	Average duration asymptomatic infectious	-	1.5
$E(\texttt{I-s} \rightarrow \texttt{ICU-a})$ and	Average duration from symptoms until		
$E(\texttt{I-s} \rightarrow \texttt{ICU-r})$	ICU need	[4], [39]	7.166
$E(\texttt{ICU-a} \rightarrow \texttt{IM})$ and	Average time in ICU care		
$E(\texttt{ICU-a} \rightarrow \texttt{D})$		[34]	11.3
$E(\texttt{ICU-r} \rightarrow \texttt{D})$ and	Average time in `ICU-r`		
$E(\texttt{ICU-r} \rightarrow \texttt{IM})$		-	5

Table 5: Average duration properties of Markov chain used to fit the transition matrix.

possible. So, rather than fitting the transition matrices for each age group independently, we fitted all transition matrices simultaneously, and extended the goodness-of-fit measure with a penalty on the pair-wise distances between the transition matrices of all age groups. The distance between two transition matrices was measured by the Euclidean norm between the two vectors containing the nonzero probabilities in these transition matrices. The resulting transition matrices are available on our github page.

2.5. Model initialization

We estimated the number of people in each health stage for each corop on the starting date of the simulation, February 27, 2020, following the rationale explained by [23]. The duration from infection till death was estimated to be on average 24 days, based on the estimated incubation time, the time from symptoms to ICU admission, and the time from ICU until death. The number of people that got infected on day t can be estimated as the number of deaths on day $t + 24$ divided by the CFR. This means that a different start situation was constructed for all CFR scenarios. To construct the start situation at February 27, day s, we used the daily number of deaths from March 6, the day of the first reported death, up to and including March 20, which is 24 days after the start date.

The CFR for people under 60 is very low ($< 0.5\%$) leading to few deaths in these age groups. We therefore only estimated the number of infected people in the age groups 60-70, 70-80 and 80+. Assuming that the fraction of the population that got infected was the same over all age groups we estimated the number of infections under 60 per day by multiplying this fraction with the population size.

People who got infected before our simulation started, thus on day $t \in \{s - 24, s - 23, \cdots, s - 1\}$, may have progressed to other stages. We therefore determined in which stage the people who got infected on day $t \in \{s - 24, s - 23, \cdots, s - 1\}$ are at day s by applying the transition matrix to the infected group for all epochs between t and s.

The total number of infected individuals was spread over the corops following the distribution of deaths over the country as reported in [37]. All initializations constructed using this approach can be obtained from our github page.

The death toll was crucial for computing the initial health stage distribution of the population. Since the actual death toll is likely higher than the number of reported COVID-19 induced deaths, we computed a lower- and an upper bound on the death toll. Two initial distributions of the population over the health stages were computed, one based on the lower bound, and one based on the upper bound. We used the average of the two as the starting point for our simulation.

The death toll reported by the National Institute for Public Health and the Environment [24] was used as a lower bound. To obtain an upper bound, we used the excess number of deaths in 2020 compared to 2015-2019. This was determined using the weekly number of deaths in the Netherlands as reported by Statistics Netherlands [32]. The excess number of deaths was computed as the number of deaths in a certain week in 2020 minus the average number of deaths in that same week for 2015-2019. The resulting excess number of deaths is 442 for week 12 and 1164 for week 13. The deaths were then distributed over the days in that week following the trend of the reported COVID-19 deaths obtained from [25]. For example, if 10% of the COVID-19 confirmed deaths in week 12 occurred on Monday, than 10% of the 442 excess deaths were added to that number. The resulting daily number of deaths are available on our github page.

3. Results

3.1. Assessing the likelihood of scenarios based on number of deaths and ICU admissions

To assess the validity of each of the scenarios, we first compared the death toll of COVID-19 predicted by the simulation model with an estimate of the actual death toll in the Netherlands. The predicted death toll by the simulation is computed as an average over ten different runs to avoid outliers as the simulation has stochastic components. As the true number of COVID-19 induced deaths is highly uncertain (see e.g. [16]), we employed a reliable but safe lower and upper bound of 602 and 2139 (see Appendix Appendix B for details), respectively. All scenarios that led to a prediction of the total number of deaths between February 27 and March 25 within this range were considered to be plausible. In this way, we limited the set of realistic parameter combinations to 126. Figure 2 presents a heatmap showing the deviation from the lower and upper bound on the number of deaths for all parameter combinations. A value of 0 means that the number of deaths in the simulation for the specific parameter combination was within the lower and upper bound and were thus acceptable parameter combinations. Table C.5 in Appendix C shows the simulated total number of deaths for each of the parameter scenarios.

We further assessed the validity of the set of realistic scenarios based on the daily number of ICU admissions. Since this was a highly reliable parameter, we were able to use the

Scenarios			Transmission probability in %								
Symptoms	CFR	Immunity	15	20	25	30	35	40	45	50	55
37.5%	Data from NL	All	320	227	87	0	0	0	0	851	1897
37.5%	Data from NL	High	337	237	96	0	0	0	0	856	1884
37.5%	Data from NL	Medium	336	248	98	0	0	0	0	847	1924
37.5%	Data from NL	Low	338	235	101	0	0	0	0	804	1914
37.5%	Literature-E	All	302	198	27	0	0	0	324	1305	2520
37.5%	Literature-E	High	295	186	24	0	0	0	332	1333	2559
37.5%	Literature-E	Medium	294	185	10	0	0	0	403	1367	2585
37.5%	Literature-E	Low	297	185	26	0	0	0	384	1315	2557
37.5%	Literature-I-s	All	296	185	4	0	0	0	550	1723	3278
37.5%	Literature-I-s	High	291	174	12	0	0	0	546	1691	3264
37.5%	Literature-I-s	Medium	294	186	11	0	0	0	535	1664	3236
37.5%	Literature-I-s	Low	300	187	8	0	0	0	546	1707	3278
50.0%	Data from NL	All	265	122	0	0	0	299	1404	2861	4745
50.0%	Data from NL	High	263	125	0	0	0	351	1544	3082	5000
50.0%	Data from NL	Medium	265	114	0	0	0	550	1839	3572	5000
50.0%	Data from NL	Low	262	107	0	0	0	536	1765	3467	5000
50.0%	Literature-E	All	209	5	0	0	131	1290	2905	5000	5000
50.0%	Literature-E	High	196	13	0	0	181	1394	3114	5000	5000
50.0%	Literature-E	Medium	204	4	0	0	168	1392	3089	5000	5000
50.0%	Literature-E	Low	186	0	0	0	176	1376	3047	5000	5000
50.0%	Literature-I-s	All	203	0	0	0	307	1620	3523	5000	5000
50.0%	Literature-I-s	High	186	0	0	0	297	1643	3648	5000	5000
50.0%	Literature-I-s	Medium	193	0	0	0	338	1748	3808	5000	5000
50.0%	Literature-I-s	Low	189	0	0	0	329	1699	3733	5000	5000
62.5%	Data from NL	All	187	0	0	0	688	2256	4368	5000	5000
62.5%	Data from NL	High	173	0	0	0	722	2331	4522	5000	5000
62.5%	Data from NL	Medium	172	0	0	0	924	2694	5000	5000	5000
62.5%	Data from NL	Low	178	0	0	0	898	2621	4968	5000	5000
62.5%	Literature-E	All	75	0	0	521	2261	4720	5000	5000	5000
62.5%	Literature-E	High	67	0	0	563	2286	4895	5000	5000	5000
62.5%	Literature-E	Medium	65	0	0	553	2379	5000	5000	5000	5000
62.5%	Literature-E	Low	72	0	0	553	2338	4942	5000	5000	5000
62.5%	Literature-I-s	All	94	0	0	335	1953	4422	5000	5000	5000
62.5%	Literature-I-s	High	80	0	0	384	2060	4610	5000	5000	5000
62.5%	Literature-I-s	Medium	84	0	0	354	2016	4629	5000	5000	5000
62.5%	Literature-I-s	Low	88	0	0	350	2013	4585	5000	5000	5000
75.0%	Data from NL	All	90	0	0	613	2448	5000	5000	5000	5000
75.0%	Data from NL	High	86	0	0	607	2486	5000	5000	5000	5000
75.0%	Data from NL	Medium	74	0	0	840	2956	5000	5000	5000	5000
75.0%	Data from NL	Low	67	0	0	836	2868	5000	5000	5000	5000
75.0%	Literature-E	All	12	0	0	1058	3133	5000	5000	5000	5000
75.0%	Literature-E	High	8	0	0	1046	3256	5000	5000	5000	5000
75.0%	Literature-E	Medium	2	0	0	1118	3266	5000	5000	5000	5000
75.0%	Literature-E	Low	0	0	0	1097	3283	5000	5000	5000	5000
75.0%	Literature-I-s	All	0	0	68	1893	4767	5000	5000	5000	5000
75.0%	Literature-I-s	High	0	0	132	1922	4875	5000	5000	5000	5000
75.0%	Literature-I-s	Medium	0	0	97	1933	4824	5000	5000	5000	5000
75.0%	Literature-I-s	Low	0	0	102	1887	4809	5000	5000	5000	5000

Figure 2: *A heatmap showing by how much the simulated number of deaths up to and including March 25 lies outside the interval [602,2139]. Columns correspond to virus transmission probabilities, rows represent the various combinations of the probability of developing symptoms, the case fatality ratio and the possibility of developing immunity. "Literature-E" and "Literature-I-s" correspond to the scenarios where the $P(D|E)$ and $P(D|I\text{-}s)$, respectively, are based on CFR estimates from the literature.*

reported values to compute the mean squared error (MSE) between the simulated and the real daily ICU admissions in the Netherlands. Any scenario that yielded an MSE greater

than or equal to 225 was considered unreliable. The threshold of 225 was chosen based on visual inspection of plots showing simulated and real daily ICU admissions, which are available at https://covid-results.herokuapp.com. The plots allow for visual comparison of the simulations with each other and with reality, as well as a visual evaluation of the progression of these metrics over time. Furthermore, an MSE of 225 corresponds roughly to a difference between simulated and real daily ICU admissions of 15 on average. Figure 3 shows a heatmap of the MSE for only those scenarios that were considered realistic based on the predicted death toll. MSEs for all scenarios are shown in Figure C.6 in Appendix C.

After removing all scenarios with an MSE of more than 225, our simulation still indicated 22 combinations of virus- and disease characteristics to be plausible. Note that the MSE for the ICU admission differed among these 22 settings. In Appendix Appendix C in Figure 3 a heatmap showing the MSE for the set of realistic scenarios based on the deaths is shown and in the MSE for all scenarios is displayed.

3.2. The selected scenarios lead to a wide range of predictions on the number of infected and immune individuals

For the 22 remaining parameter combinations, the left colored panel of Figure 4 shows the percentage of infected individuals up to March 12 according to our simulation. At that time social distancing measures were installed in the Netherlands, leading to a strong reduction in virus spread. Since we were interested in virus behavior under stable circumstances, i.e., without changes in human behavior, we only considered the number of infected individuals until March 12. Note that the number of infected individuals was affected by social distancing immediately, while the ICU admissions and deaths were affected only after some time, which was why we used a different time horizon for the infections (until 12 March) than with ICU admissions and deaths (until 25 March). There was substantial variation among the 22 scenarios: the percentage of the population that was infected varies from 1.4% up to 4.3%.

An interesting metric for many policy makers is the fraction of the population that develops immunity. As before, we excluded the effects of social distancing by considering infections that happened no later than March 12. Only for the individuals who were infected no later than March 12, we simulated the disease progression until they ended up in one of the stages susceptible, immune or deceased. The right colored panel of Figure 4 displays the fraction of the population that develops immunity based on the infections up to March 12. The results show major differences among the 22 scenarios: the percentage of the population that was immune varies between 1.1% and 4.0%.

The percentages of infected individuals up to March 12 and percentages of immune individuals for all parameter combinations are shown in Tables C.7 and C.8, respectively, in Appendix C.

			Transmission probability in %						
Symptoms	CFR	Immunity	15	20	25	30	35	40	45
37.5%	Data from NL	All	*	*	*	808	238	159	1540
37.5%	Data from NL	High	*	*	*	820	225	170	1455
37.5%	Data from NL	Medium	*	*	*	754	205	161	1542
37.5%	Data from NL	Low	*	*	*	801	238	138	1467
37.5%	Literature-E	All	*	*	*	1233	680	214	*
37.5%	Literature-E	High	*	*	*	1153	678	204	*
37.5%	Literature-E	Medium	*	*	*	1185	665	193	*
37.5%	Literature-E	Low	*	*	*	1201	682	218	*
37.5%	Literature-Is	All	*	*	*	1732	1464	1066	*
37.5%	Literature-Is	High	*	*	*	1747	1499	1125	*
37.5%	Literature-Is	Medium	*	*	*	1732	1503	1108	*
37.5%	Literature-Is	Low	*	*	*	1768	1466	1090	*
50.0%	Data from NL	All	*	*	813	187	348	*	*
50.0%	Data from NL	High	*	*	778	184	550	*	*
50.0%	Data from NL	Medium	*	*	707	144	618	*	*
50.0%	Data from NL	Low	*	*	806	148	644	*	*
50.0%	Literature-E	All	*	*	1082	413	*	*	*
50.0%	Literature-E	High	*	*	1073	403	*	*	*
50.0%	Literature-E	Medium	*	*	1084	443	*	*	*
50.0%	Literature-E	Low	*	1570	1099	439	*	*	*
50.0%	Literature-I-s	All	*	1869	1557	1070	*	*	*
50.0%	Literature-I-s	High	*	1859	1549	1008	*	*	*
50.0%	Literature-I-s	Medium	*	1864	1533	1039	*	*	*
50.0%	Literature-I-s	Low	*	1867	1554	1080	*	*	*
62.5%	Data from NL	All	*	939	189	649	*	*	*
62.5%	Data from NL	High	*	941	196	800	*	*	*
62.5%	Data from NL	Medium	*	913	175	1007	*	*	*
62.5%	Data from NL	Low	*	919	158	995	*	*	*
62.5%	Literature-E	All	*	1144	364	*	*	*	*
62.5%	Literature-E	High	*	1096	347	*	*	*	*
62.5%	Literature-E	Medium	*	1141	345	*	*	*	*
62.5%	Literature-E	Low	*	1097	314	*	*	*	*
62.5%	Literature-I-s	All	*	1570	932	*	*	*	*
62.5%	Literature-I-s	High	*	1536	904	*	*	*	*
62.5%	Literature-I-s	Medium	*	1503	914	*	*	*	*
62.5%	Literature-I-s	Low	*	1518	889	*	*	*	*
75.0%	Data from NL	All	*	467	293	*	*	*	*
75.0%	Data from NL	High	*	395	395	*	*	*	*
75.0%	Data from NL	Medium	*	375	557	*	*	*	*
75.0%	Data from NL	Low	*	352	533	*	*	*	*
75.0%	Literature-E	All	*	916	165	*	*	*	*
75.0%	Literature-E	High	*	888	178	*	*	*	*
75.0%	Literature-E	Medium	*	894	157	*	*	*	*
75.0%	Literature-E	Low	1572	874	163	*	*	*	*
75.0%	Literature-I-s	All	1623	967	*	*	*	*	*
75.0%	Literature-I-s	High	1629	958	*	*	*	*	*
75.0%	Literature-I-s	Medium	1665	973	*	*	*	*	*
75.0%	Literature-I-s	Low	1638	958	*	*	*	*	*

Figure 3: *A heatmap representing the prediction quality of different combinations of COVID-19 characteristics with respect to ICU occupation. Quality is measured as the MSE between simulated and real daily ICU admissions up to and including March 25. A * indicates that the combination is not accurate in predicting the number of deaths. Columns correspond to virus transmission probabilities, rows represent the various combinations of the probability of developing symptoms, the case fatality ratio and the possibility of developing immunity. "Literature-E" and "Literature-I-s" correspond to the scenarios where the $P(D|E)$ and $P(D|I\text{-}s)$, respectively, are based on CFR estimates from the literature.*

Symptoms	CFR	Immunity	Transmission Probability	% infected	% immune
37.5%	Data from NL	High	35%	3	2.8
37.5%	Data from NL	Medium	35%	3	2.8
37.5%	Data from NL	All	40%	4.3	4
37.5%	Data from NL	High	40%	4.3	4
37.5%	Data from NL	Medium	40%	4.3	4
37.5%	Data from NL	Low	40%	4.3	4
37.5%	Literature-E	All	40%	2.8	2.7
37.5%	Literature-E	High	40%	2.8	2.6
37.5%	Literature-E	Medium	40%	2.8	2.6
37.5%	Literature-E	Low	40%	2.8	2.6
50.0%	Data from NL	All	30%	2.2	2.1
50.0%	Data from NL	High	30%	2.3	1.7
50.0%	Data from NL	Medium	30%	2.4	1.6
50.0%	Data from NL	Low	30%	2.4	1.6
62.5%	Data from NL	All	25%	1.8	1.7
62.5%	Data from NL	High	25%	1.8	1.4
62.5%	Data from NL	Medium	25%	1.9	1.4
62.5%	Data from NL	Low	25%	1.9	1.3
75.0%	Literature-E	All	25%	1.4	1.4
75.0%	Literature-E	High	25%	1.5	1.2
75.0%	Literature-E	Medium	25%	1.5	1.1
75.0%	Literature-E	Low	25%	1.5	1.1

Figure 4: *The percentage of infected people on March 12 (penultimate column) and the percentage of immunity among the population based on individuals who were infected no later than March 12 (last column). Results are shown only for scenarios that lead to reasonable predictions of the death toll and the daily ICU admissions.*

4. Discussion and conclusions

Using an epidemiological simulation model, we evaluated the likelihood of a variety of scenarios for virus spread and disease progression characteristics for the SARS-CoV-2 virus. We identified 22 sets of parameter values that all led to accurately simulated daily ICU admissions and number of deaths. In particular, our analysis indicated the following conclusions (Figures 2 and 3):

(i) All Literature-I-s scenarios, where we determined $P(\texttt{D}|\texttt{I-s})$ based on [35], do not seem realistic, as none of the corresponding scenarios yielded a realistic death toll and number of daily ICU admissions. Both Literature-E and Data from NL were realistic case fatality ratio scenarios, and further research is required.

(ii) In order for the simulation results to match reality, a high probability of symptom development required the transmission probability to be low and vice versa. According to our analysis, a low transmission probability combined with a high probability of developing symptoms was equally likely as a high transmission probability combined with a low probability of developing symptoms.

(iii) No conclusions regarding the probability of developing immunity could be drawn from this simulation study: the death toll and the number of daily ICU admissions differed only slightly between scenarios with different assumptions on immunity,

ceteris paribus. Further research is required to obtain better estimates on the probability of developing symptoms.

While our study provided some clear pointers towards assumptions that were likely to reflect actual virus behavior, other important questions remain unanswered. This has several implications. First and foremost, modeling the spread of SARS-CoV-2 does not give conclusive insights in the number of infections and the level of immunity among the population, and requires a thorough analysis of the results for several uncertain scenarios regarding virus- and disease characteristics.

Second, the probability of developing symptoms was highly uncertain, which has major implications for virus spread predictions. When the probability for symptom development was assumed to be low, our simulation model could only reach a realistic number of ICU admissions and death toll if the transmission probability was high. This would imply a high attack rate (the percentage of the population that contracts the disease) during the early phase of the pandemic, leading to a large fraction of the community to be infected and possibly have developed immunity. On the other hand, if the probability of symptom development is high and the transmission probability is low, this would mean that only few infections have taken place so far and pre-symptomatic infections are less likely. It is thus very important to gain more insight in the probability of developing COVID-19 symptoms after infection.

Third, the progression of the virus spread in the long run was difficult to predict. If many people have already been infected, and if many infections have led to immunity, the level of immunity among the population has grown rapidly. This means that after a relatively short amount of time the susceptible group will decline and the death toll and ICU burden will reduce rapidly. On the other hand, a low number of infections and a low probability of developing immunity holds the potential of many more infections and hence deaths to come. This is crucial for policy makers when choosing e.g. the proper level of social distancing measures and scaling up the ICU capacity.

Fourth, the haziness around immunity has major implications for policy makers. It is unclear when an infection leads to immunity and whether immunity is obtained for life [36]. Some people may even have a good immune response already at the first infection and can therefore be considered immune prior to infection. Immunity is of vital importance for political decision making when developing a vaccination policy or aiming for herd immunity. Hence, developing exit strategies is not possible without further research on immunity.

By varying only four input parameters we already obtained 22 scenarios that we considered to be realistic, but led to a variety of predictions in terms of virus spread. Epidemiological models such as ours are based on several other uncertain or unknown inputs, such as the incubation time, the time until an exposed individual becomes infectious, the probability of ending up in the ICU and the fraction of symptomatic patients that goes into self-quarantine. Varying these may lead to an even larger variation in model predictions.

In case further research resolves one or more of the uncertainties in virus and/or disease characteristics, then our or a similar study can be used to narrow down the range of possibilities for the other characteristics. Additionally, a better estimate of the death toll allows for further reducing the number of realistic scenarios. For example, if the lower bound on the death toll estimate were increased by 50% and the upper bound reduced by 33%, only 10 scenarios remain. The bandwidth of the predictions on infections and immunity is already much smaller than for the original 22 scenarios: the percentage of the population that was infected varied from 1.8% to 3%, and the level of immunity ranged from 1.3% to 2.8%.

The values for virus transmission probability reported in this study should be interpreted with caution. The daily number of infected individuals in our simulation was determined by multiplying this virus transmission probability with the number of infectious individuals and their number of daily contacts. Here, the definition of a contact is important: including only conversations as contacts results in fewer social contacts than including also e.g. passers-by. A restricted definition of contacts naturally corresponds to a higher transmission probability. This was reflected by the model as well: a more restricted definition leads to fewer contacts, hence a higher transmission probability was necessary to achieve the same number of infections, ICU admissions and deaths. We therefore did not consider our validated values for the virus transmission probabilities to be exact and universally applicable. Rather we showed that for a variety of assumptions on CFR, developing immunity and symptom development, and for a given definition of contacts, there exist transmission probabilities that lead to realistic simulation results.

In summary, this paper presents a comprehensive study to test the validity of a wide range of SARS-CoV-2 virus- and COVID-19 disease characteristics. A variety of assumptions yielded a realistic number of deaths and daily ICU admissions. However, these scenarios disagreed on the predicted number of infections and immune individuals, two unobservable but important metrics. From this we conclude that the currently available information on the behavior of the SARS-CoV-2 virus is insufficient to accurately model and predict the virus spread, evaluate the effects of social distancing measures in detail and develop social distancing policies or even exit strategies. Note that this does not imply that such studies are not informative: epidemiological forecasting models have helped us understand the severity of the pandemic already early on, and are well capable of forecasting the trend of the virus spread. When conducting a forecasting analysis that requires insights in the infections and immunity among a population we highly recommend to analyze the results for several scenarios on virus- and disease characteristics, report the range of results obtained with these scenarios and draw conclusions and recommendations based on the full set of scenario-specific outcomes.

Acknowledgement

We thank dr. Jean-Luc Murk of the Elisabeth-Tweesteden hospital Tilburg, the Netherlands, for his valuable input and feedback.

References

[1] Indeling van Nederland in 40 COROP-gebieden. `https://www.cbs.nl/-/media/_pdf/2019/04/2019ov12_kaart_40-coropgebieden.pdf`. Accessed: 2020-08-03.

[2] T. Alamo, D. G. Reina, M. Mammarella, and A. Abella. Covid-19: Open-data resources for monitoring, modeling, and forecasting the epidemic. Electronics, 9(5):827, 2020.

[3] J. Bakker, J. Damen, A. Van Zanten, and J. Hubben. Criteria voor opname en ontslag van intensive care afdelingen in nederland. Ned Tijdschr Geneesk, 147:110–115, 01 2003.

[4] D. Bertsimas, H. Bandi, L. Boussioux, R. Cory-Wright, A. Delarue, V. Digalakis, S. Gilmour, J. Graham, A. Kim, D. Lahlou Kitane, Z. Lin, G. Lukin, M. Li, L. Mingardi, L. Na, A. Orfanoudaki, T. Papalexopoulos, I. Paskov, J. Pauphilet, O. Skali Lami, M. Sobiesk, B. Stellato, K. Carballo, Y. Wang, H. Wiberg, and C. Zeng. An aggregated dataset of clinical outcomes for COVID-19 patients, 2020. `http://www.covidanalytics.io/dataset_documentation` accessed on April 10, 2020.

[5] J. F.-W. Chan, S. Yuan, K.-H. Kok, K. K.-W. To, H. Chu, J. Yang, F. Xing, J. Liu, C. C.-Y. Yip, R. W.-S. Poon, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. The Lancet, 395(10223):514–523, 2020.

[6] S. Cobey. Modeling infectious disease dynamics. Science, 368(6492):713–714, 2020.

[7] N. G. Davies, A. J. Kucharski, R. M. Eggo, A. Gimma, W. J. Edmunds, C. C.-. W. Group, et al. The effect of non-pharmaceutical interventions on covid-19 cases, deaths and demand for hospital services in the uk: a modelling study. MedRxiv, 2020.

[8] S. J. de Vlas and L. E. Coffeng. A phased lift of control: a practical strategy to achieve herd immunity against covid-19 at the country level. medRxiv, 2020.

[9] S. Y. Del Valle, J. M. Hyman, H. W. Hethcote, and S. G. Eubank. Mixing patterns between age groups in social networks. Social Networks, 29(4):539–554, 2007.

[10] S. Y. Del Valle, J. M. Hyman, H. W. Hethcote, and S. G. Eubank. Mixing patterns between age groups in social networks. Social Networks, 29(4):539–554, 2007.

[11] N. Ferguson, D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunuba Perez, G. Cuomo-Dannenburg, et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. 2020.

[12] C. Hou, J. Chen, Y. Zhou, L. Hua, J. Yuan, S. He, Y. Guo, S. Zhang, Q. Jia, C. Zhao, et al. The effectiveness of quarantine of Wuhan city against the Corona Virus Disease 2019 (COVID-19): A well-mixed SEIR model analysis. Journal of medical virology, 2020.

[13] Z. Hu, C. Song, C. Xu, G. Jin, Y. Chen, X. Xu, H. Ma, W. Chen, Y. Lin, Y. Zheng, et al. Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. Science China Life Sciences, 63(5):706–711, 2020.

[14] S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, and M. Lipsitch. Projecting the transmission dynamics of sars-cov-2 through the postpandemic period. Science, 368(6493):860–868, 2020.

[15] A. J. Kucharski, T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, R. M. Eggo, F. Sun, M. Jit, J. D. Munday, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. The lancet infectious diseases, 2020.

[16] R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, and J. Shaman. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (sars-cov-2). Science, 368(6490):489–493, 2020.

[17] Y. Liu, A. A. Gayle, A. Wilder-Smith, and J. Rocklöv. The reproductive number of covid-19 is higher compared to sars coronavirus. Journal of travel medicine, 2020.

[18] MIDAS Network, 2020. `https://midasnetwork.us/covid-19/`.

[19] S. M. Mniszewski, S. Y. Del Valle, P. D. Stroud, J. M. Riese, and S. J. Sydoriak. Episims simulation of a multi-component strategy for pandemic influenza. In Proceedings of the 2008 Spring simulation multiconference, pages 556–563. Society for Computer Simulation International, 2008.

[20] S. A. Muller, M. Balmer, A. Neumann, and K. Nagel. Mobility traces and spreading of covid-19. medRxiv, 2020.

[21] L. Peng, W. Yang, D. Zhang, C. Zhuge, and L. Hong. Epidemic analysis of covid-19 in china by dynamical modeling. arXiv preprint arXiv:2002.06563, 2020.

[22] K. Prem, Y. Liu, T. W. Russell, A. J. Kucharski, R. M. Eggo, N. Davies, S. Flasche, S. Clifford, C. A. Pearson, J. D. Munday, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. The Lancet Public Health, 2020.

[23] T. Pueyo, 2020. https://medium.com/@tomaspueyo/coronavirus-act-today-or-people-will-die-f4d3d9cd99ca.

[24] J. Riou and C. L. Althaus. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. Eurosurveillance, 25(4):2000058, 2020.

[25] RIVM, 2020. https://www.rivm.nl/en/novel-coronavirus-covid-19/current-information-about-novel-coronavirus-covid-19, accessed on April 10, 2020.

[26] W. C. Roda, M. B. Varughese, D. Han, and M. Y. Li. Why is it difficult to accurately predict the covid-19 epidemic? Infectious Disease Modelling, 2020.

[27] S. M. Ross. Introduction to Probability Models. Academic Press, 11th edition, 2014.

[28] C. Rothe, M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, T. Zimmer, V. Thiel, C. Janke, W. Guggemos, et al. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. New England Journal of Medicine, 382(10):970–971, 2020.

[29] H. Salje, C. T. Kiem, N. Lefrancq, N. Courtejoie, P. Bosetti, J. Paireau, A. Andronico, N. Hozé, J. Richet, C.-L. Dubost, et al. Estimating the burden of SARS-CoV-2 in France. Science, 2020.

[30] G. Sebastiani, M. Massa, and E. Riboli. Covid-19 epidemic in Italy: evolution, projections and impact of government measures. European Journal of Epidemiology, page 1, 2020.

[31] Statistics Netherlands (CBS). Banen van werknemers naar woon- en werkregio, 2020. data accessed on March 20, 2020, https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83628NED/table.

[32] Statistics Netherlands (CBS). Overledenen; geslacht en leeftijd, per week, 2020. data retrieved on May 12, 2020, https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70895ned/table?ts=1591770380485.

[33] Statistics Netherlands (CBS). Regionale kerncijfers Nederland, 2020. data retrieved on March 20, 2020, https://opendata.cbs.nl/statline/#/CBS/nl/dataset/03759ned/table?ts=1591775235782.

[34] Stichting NICE, 2020. https://www.stichting-nice.nl/ accessed on April 28, 2020.

[35] The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team. The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) — China, 2020. China CDC Weekly, 2:113, 2020.

[36] K. K.-W. To, I. F.-N. Hung, J. D. Ip, A. W.-H. Chu, W.-M. Chan, A. R. Tam, C. H.-Y. Fong, S. Yuan, H.-W. Tsoi, A. C.-K. Ng, et al. Covid-19 re-infection by a phylogenetically distinct sars-coronavirus-2 strain confirmed by whole genome sequencing. Clinical Infectious Diseases, 2020.

[37] J. Van Dissel, 2020. Slides "Technische briefing Tweede Kamer, 22 april 2020". Available from https://www.tweedekamer.nl/debat_en_vergadering/commissievergaderingen/details?id=2020A01701.

[38] J. Wallinga, P. Teunis, and M. Kretzschmar. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. American journal of epidemiology, 164(10):936–944, 2006.

[39] D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in wuhan, china. Jama, 323(11):1061–1069, 2020.

[40] B. Wilder, M. Charpignon, J. A. Killian, H.-C. Ou, A. Mate, S. Jabbari, A. Perrault, A. Desai, M. Tambe, and M. S. Majumder. The role of age distribution and family structure on COVID-19 dynamics: A preliminary modeling assessment for hubei and lombardy. Available at SSRN 3564800, 2020.

[41] Z. Yang, Z. Zeng, K. Wang, S.-S. Wong, W. Liang, M. Zanin, P. Liu, X. Cao, Z. Gao, Z. Mai, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. Journal of Thoracic Disease, 12(3):165, 2020.

[42] W. Zhang, W. Cheng, L. Luo, Y. Ma, C. Xu, P. Qin, and Z. Zhang. Secondary transmission of coronavirus disease from presymptomatic persons, china. Emerging Infectious Diseases, 26(8), 2020.

Appendix A. Case fatality ratio

The case fatality ratio, i.e. the probability that an infection leads to death, was estimated following three different approaches. The case fatality ratios used in scenario "Data from NL" were defined as the estimated death toll in the Netherlands divided by the estimated number of infected individuals. We assumed that this provides the probability that someone who gets infected, dies from COVID-19 ($P(\texttt{D}|\texttt{E})$). In scenarios Literature-E and Literature-I-s we used case fatality ratios obtained from the literature [35], where we assumed that these ratios reflect $P(\texttt{D}|\texttt{E})$ for Literature-E and $P(\texttt{D}|\texttt{I-s})$ for Literature-I-s. The computed case fatality ratios for each scenario are provided in Table A.6.

Age group	Data from NL ($P(\text{D}\|\texttt{I-s})$)	Literature-E ($P(\texttt{D}\|\texttt{I-s})$)	Literature-I-s ($P(\texttt{D}\|\texttt{I-a})$)
0-9	0	0	0
10-19	3.32×10^{-5}	2.00×10^{-3}	2.00×10^{-3}
20-29	4.00×10^{-5}	2.00×10^{-3}	2.00×10^{-3}
30-39	8.13×10^{-5}	2.00×10^{-3}	2.00×10^{-3}
40-49	1.62×10^{-4}	4.00×10^{-3}	4.00×10^{-3}
50-59	1.03×10^{-3}	1.30×10^{-2}	1.30×10^{-2}
60-69	6.43×10^{-3}	3.60×10^{-2}	3.60×10^{-2}
70-79	4.82×10^{-2}	8.00×10^{-2}	8.00×10^{-2}
80+	2.43×10^{-1}	1.48×10^{-1}	1.48×10^{-1}

Table A.6: Case fatality ratios used for the three different CFR scenarios.

For scenario "Data from NL" we needed estimates on the number of deaths and the number of infected individuals. The death toll as reported by the Dutch national health institute, RIVM [25], is likely an underestimation: many deaths caused by COVID-19 may not be reported as such because COVID-19 was not confirmed or even suspected. Another estimate for the number of COVID-19 deaths can be obtained from the excess deaths in 2020 compared to 2015-2019 based on weekly death rates reported by Statistics Netherlands [32]. Since Sanquin reported test results on blood samples obtained up to and including week 15, we used the excess deaths up to and including week 16 for computing the CFR. The extra week was to account for the time from infection till death. CFRs were then computed based on the average between the death toll reported by RIVM and the excess deaths reported by Statistics Netherlands. Details on the number of deaths reported by RIVM, excess deaths and the average between the two are provided in Table A.7. Note that no excess deaths were observed for people under 70 years old.

The number of infected individuals was estimated based on research by Sanquin, the Dutch blood bank, who tested all blood donations of individuals between 20 and 60 years old for antibodies and found that on average 3% of donors has developed antibodies [37] up to 13 April. Estimates for age groups 0-9 and 10-19 were obtained from the Pienter investigation [37]. People over the age of 70 rarely donate blood, and no estimates of

the infected fractions were available. We therefore fit a linear relationship between the number of daily contacts and the infected fraction for each age group, and used this to estimate the fraction of infected individuals for all age groups above 20 years of age. Details on the final percentages used for computing the CFR are provided in Table A.7.

Table A.7: COVID-19 deaths as reported by RIVM, excess deaths as reported by Statistics Netherlands and the average of the two up to and including week 12. Excess deaths were only observed for the age groups 70-79 and 80+.

Age group	Death rates			% of the population infected	
	COVID-19	Excess deaths	Average	Sanquin/Pienter	Estimated CFR
0-9	0	-	0	1	1
10-19	1	-	1	1.5	1.5
20-29	3	-	3	3.6	3.4
30-30	6	-	6	3.4	3.5
40-49	13	-	13	3.5	3.5
50-59	87	-	87	3.1	3.4
60-69	332	-	332	2.8	2.5
70-79	1084	1274	1179		1.6
80-150	1636	4277	2957		1.5

Appendix B. Actual number of deaths

The death toll reported by RIVM is an underestimation on the real death toll. RIVM reported a total of 602 COVID-19 induced deaths up to March 25 and we used this as a lower bound in our verification of parameter combinations; if the number of deaths were not at least equal to 602, we knew that the parameter combination was not realistic. An upper bound can be obtained from the excess deaths reported by Statistics Netherlands in comparison with the years 2015-2019 [32]. However, we computed the excess deaths not only by comparing the number of deaths in a week to the number of deaths in the same week in 2015-2019, but also to the number of deaths occurring in previous weeks in 2020. Thus the upper bound on the number of deaths in each week of our study is computed by taking the maximum difference between the reported number of deaths by CBS in this week and the reported number of deaths in this particular week in the years 2015-2019 and the number of deaths reported in week 1 up to (but not including) this week in 2020. This gives us an upper bound of 2139 up to and including March 25, and in case the simulation has more deaths than 2139 the parameter combination used is considered unrealistic.

Appendix C. Additional results

Total number of deaths.
Figure C.5 shows the total number of deaths obtained for each parameter combination.

Scenarios			Transmission probability								
Symptoms	CFR	Immunity	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55
37.5%	Data from NL	All	281	374	515	741	1046	1493	2124	2990	4035
37.5%	Data from NL	High	264	365	506	703	1045	1480	2118	2994	4023
37.5%	Data from NL	Medium	266	354	504	711	1037	1491	2139	2985	4062
37.5%	Data from NL	Low	264	366	501	707	1018	1480	2098	2942	4053
37.5%	Literature-E	All	300	404	574	846	1230	1721	2463	3443	4658
37.5%	Literature-E	High	307	415	577	853	1214	1750	2471	3471	4697
37.5%	Literature-E	Medium	307	416	592	852	1227	1757	2541	3505	4723
37.5%	Literature-E	Low	304	417	576	838	1217	1745	2523	3454	4696
37.5%	Literature-I-s	All	305	417	598	878	1259	1878	2689	3861	5416
37.5%	Literature-I-s	High	311	428	590	860	1267	1847	2685	3829	5402
37.5%	Literature-I-s	Medium	307	415	591	865	1257	1830	2674	3803	5375
37.5%	Literature-I-s	Low	302	415	594	859	1272	1831	2685	3846	5416
50.0%	Data from NL	All	337	479	705	1072	1632	2438	3543	5000	6883
50.0%	Data from NL	High	338	477	713	1086	1633	2490	3682	5220	7215
50.0%	Data from NL	Medium	337	487	734	1140	1762	2689	3977	5710	7853
50.0%	Data from NL	Low	339	495	740	1143	1718	2675	3903	5606	7797
50.0%	Literature-E	All	392	596	934	1455	2270	3429	5044	7177	9798
50.0%	Literature-E	High	406	588	934	1484	2319	3533	5252	7514	10379
50.0%	Literature-E	Medium	398	597	931	1465	2307	3531	5227	7540	10420
50.0%	Literature-E	Low	415	605	931	1472	2315	3515	5186	7515	10329
50.0%	Literature-I-s	All	399	608	968	1537	2445	3758	5662	8384	11961
50.0%	Literature-I-s	High	415	619	978	1534	2436	3782	5787	8577	12220
50.0%	Literature-I-s	Medium	408	619	975	1554	2477	3887	5946	8776	12561
50.0%	Literature-I-s	Low	412	630	993	1566	2468	3837	5871	8665	12394
62.5%	Data from NL	All	414	657	1073	1744	2826	4394	6507	9197	12343
62.5%	Data from NL	High	429	657	1066	1772	2861	4469	6661	9479	12854
62.5%	Data from NL	Medium	429	683	1126	1894	3063	4833	7230	10297	14001
62.5%	Data from NL	Low	424	674	1119	1833	3036	4760	7107	10273	14011
62.5%	Literature-E	All	527	887	1562	2660	4400	6858	10241	14429	19377
62.5%	Literature-E	High	535	899	1575	2702	4425	7034	10531	15006	20243
62.5%	Literature-E	Medium	536	903	1580	2692	4518	7159	10774	15285	20599
62.5%	Literature-E	Low	530	899	1582	2691	4476	7080	10666	15146	20640
62.5%	Literature-I-s	All	507	851	1445	2474	4091	6561	10133	14741	20423
62.5%	Literature-I-s	High	522	871	1485	2523	4199	6749	10359	15230	21483
62.5%	Literature-I-s	Medium	518	875	1472	2493	4154	6767	10385	15304	21449
62.5%	Literature-I-s	Low	513	856	1450	2489	4152	6723	10327	15126	21327
75.0%	Data from NL	All	512	876	1547	2751	4587	7228	10688	14873	19782
75.0%	Data from NL	High	515	892	1569	2746	4625	7373	10932	15357	20420
75.0%	Data from NL	Medium	528	931	1698	2979	5095	8015	11965	16819	22373
75.0%	Data from NL	Low	534	934	1693	2975	5007	8014	11943	16775	22361
75.0%	Literature-E	All	590	1048	1836	3196	5272	8320	12326	17297	23084
75.0%	Literature-E	High	594	1067	1862	3184	5394	8452	12662	17803	24036
75.0%	Literature-E	Medium	600	1058	1870	3256	5404	8687	12900	18324	24525
75.0%	Literature-E	Low	613	1071	1884	3235	5422	8582	12866	18233	24554
75.0%	Literature-I-s	All	661	1206	2206	4032	6906	11182	16975	24255	32471
75.0%	Literature-I-s	High	658	1211	2271	4061	7013	11452	17460	24940	33770
75.0%	Literature-I-s	Medium	653	1204	2235	4072	6963	11377	17412	24811	33663
75.0%	Literature-I-s	Low	654	1205	2241	4025	6948	11411	17298	24716	33630

Figure C.5: *The total number of deaths as simulation output for various parameter combinations. The columns denote the different virus transmission probabilities and every row a certain combination of the probability of developing symptoms, the case fatality ratio and the possibility of developing immunity.*

ICU.

Then, Figure C.6 shows the MSE on the ICU for all possible parameter combinations, not only for the realistic ones.

Scenarios			Transmission probability								
Symptoms	CFR	Immunity	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55
37.5%	Data from NL	All	1936	1683	1337	808	238	159	1540	5690	13395
37.5%	Data from NL	High	1940	1715	1348	820	225	170	1455	5444	13777
37.5%	Data from NL	Medium	1930	1705	1350	754	205	161	1542	5430	13626
37.5%	Data from NL	Low	1965	1710	1325	801	238	138	1467	5442	13790
37.5%	Literature-E	All	2011	1862	1554	1233	680	214	192	1276	4314
37.5%	Literature-E	High	2012	1839	1577	1153	678	204	149	1337	4551
37.5%	Literature-E	Medium	1993	1838	1569	1185	665	193	162	1312	4760
37.5%	Literature-E	Low	1992	1829	1605	1201	682	218	163	1322	4664
37.5%	Literature-Is	All	2105	2046	1937	1732	1464	1066	670	271	270
37.5%	Literature-Is	High	2124	2063	1942	1747	1499	1125	624	267	265
37.5%	Literature-Is	Medium	2099	2048	1948	1732	1503	1108	638	262	274
37.5%	Literature-Is	Low	2108	2060	1943	1768	1466	1090	633	273	266
50.0%	Data from NL	All	1806	1431	813	187	348	3225	10000	10000	10000
50.0%	Data from NL	High	1771	1387	778	184	550	3754	10000	10000	10000
50.0%	Data from NL	Medium	1786	1389	707	144	618	4962	10000	10000	10000
50.0%	Data from NL	Low	1812	1356	806	148	644	4749	10000	10000	10000
50.0%	Literature-E	All	1881	1592	1082	413	119	1393	6750	10000	10000
50.0%	Literature-E	High	1905	1571	1073	403	138	1757	8256	10000	5181
50.0%	Literature-E	Medium	1886	1584	1084	443	146	1630	7938	10000	10000
50.0%	Literature-E	Low	1889	1570	1099	439	145	1673	7916	10000	10000
50.0%	Literature-Is	All	2044	1869	1557	1070	470	172	1154	5406	10000
50.0%	Literature-Is	High	2018	1859	1549	1008	470	179	1316	6246	10000
50.0%	Literature-Is	Medium	2036	1864	1533	1039	488	197	1358	6095	10000
50.0%	Literature-Is	Low	2041	1867	1554	1080	458	185	1422	6550	10000
62.5%	Data from NL	All	1592	939	189	649	6273	10000	10000	10000	10000
62.5%	Data from NL	High	1599	941	196	800	6852	2615	10000	10000	10000
62.5%	Data from NL	Medium	1588	913	175	1007	8399	10000	10000	10000	10000
62.5%	Data from NL	Low	1569	919	158	995	8133	3173	10000	10000	10000
62.5%	Literature-E	All	1707	1144	364	325	4083	10000	10000	10000	10000
62.5%	Literature-E	High	1714	1096	347	343	4893	10000	10000	10000	10000
62.5%	Literature-E	Medium	1675	1141	345	358	5016	10000	5903	10000	10000
62.5%	Literature-E	Low	1696	1097	314	462	4771	10000	10000	10000	10000
62.5%	Literature-Is	All	1885	1570	932	248	563	4690	10000	10000	10000
62.5%	Literature-Is	High	1880	1536	904	225	690	5531	10000	10000	10000
62.5%	Literature-Is	Medium	1898	1503	914	242	631	4864	10000	10000	10000
62.5%	Literature-Is	Low	1904	1518	889	240	671	5088	10000	10000	10000
75.0%	Data from NL	All	1311	467	293	6037	10000	10000	10000	10000	10000
75.0%	Data from NL	High	1324	395	395	6145	10000	10000	10000	10000	10000
75.0%	Data from NL	Medium	1268	375	557	8128	10000	10000	10000	10000	10000
75.0%	Data from NL	Low	1267	352	533	8404	10000	10000	10000	10000	10000
75.0%	Literature-E	All	1584	916	165	1184	9201	10000	8089	10000	10000
75.0%	Literature-E	High	1594	888	178	1205	9910	10000	10000	10000	10000
75.0%	Literature-E	Medium	1565	894	157	1318	10000	10000	10000	10000	10000
75.0%	Literature-E	Low	1572	874	163	1340	10000	10000	9874	10000	10000
75.0%	Literature-Is	All	1623	967	213	1267	10000	10000	10000	10000	10000
75.0%	Literature-Is	High	1629	958	200	1324	10000	10000	10000	10000	10000
75.0%	Literature-Is	Medium	1665	973	195	1321	10000	10000	10000	10000	10000
75.0%	Literature-Is	Low	1638	958	193	1249	10000	10000	10000	10000	10000

Figure C.6: *A heatmap representing the quality of prediction of various combinations of COVID-19 characteristics with respect to ICU occupation. The columns denote the different virus transmission probabilities and every row a certain combination of the probability of developing symptoms, the case fatality ratio and the possibility of developing immunity.*

Percentage Infected and Immune individuals.
Figures C.7 and C.8 show the percentages of infected and immune individuals for all possible parameter combinations.

Scenarios			Transmission probability								
Symptoms	CFR	Immunity	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55
37.5%	Data from NL	All	0.4	0.8	1.2	2	2.9	4.3	6	8.3	11.1
37.5%	Data from NL	High	0.4	0.8	1.2	2	3	4.3	6	8.3	11.1
37.5%	Data from NL	Medium	0.4	0.8	1.3	2	3	4.3	6.1	8.4	11.2
37.5%	Data from NL	Low	0.4	0.8	1.3	2	3	4.3	6	8.3	11.1
37.5%	Literature-E	All	0.3	0.5	0.9	1.3	2	2.8	3.9	5.4	7.2
37.5%	Literature-E	High	0.3	0.5	0.9	1.3	1.9	2.8	3.9	5.4	7.2
37.5%	Literature-E	Medium	0.3	0.5	0.9	1.3	2	2.8	4	5.4	7.2
37.5%	Literature-E	Low	0.3	0.5	0.9	1.3	2	2.8	3.9	5.4	7.2
37.5%	Literature-I_s	All	0.1	0.2	0.3	0.5	0.8	1.1	1.5	2.1	2.8
37.5%	Literature-I_s	High	0.1	0.2	0.3	0.5	0.7	1.1	1.5	2.1	2.8
37.5%	Literature-I_s	Medium	0.1	0.2	0.3	0.5	0.7	1.1	1.5	2.1	2.8
37.5%	Literature-I_s	Low	0.1	0.2	0.3	0.5	0.7	1.1	1.5	2.1	2.8
50.0%	Data from NL	All	0.5	0.8	1.4	2.2	3.4	5	7.2	9.9	13.2
50.0%	Data from NL	High	0.5	0.8	1.4	2.3	3.5	5.1	7.3	10.1	13.6
50.0%	Data from NL	Medium	0.5	0.9	1.5	2.4	3.7	5.5	7.9	11	14.9
50.0%	Data from NL	Low	0.5	0.9	1.5	2.4	3.7	5.5	7.9	11	14.8
50.0%	Literature-E	All	0.4	0.6	1.1	1.7	2.7	3.9	5.6	7.8	10.6
50.0%	Literature-E	High	0.4	0.7	1.1	1.8	2.7	4	5.7	8	10.9
50.0%	Literature-E	Medium	0.4	0.7	1.1	1.8	2.7	4	5.7	8	10.8
50.0%	Literature-E	Low	0.4	0.6	1.1	1.7	2.7	4	5.7	7.9	10.8
50.0%	Literature-I_s	All	0.2	0.3	0.6	0.9	1.4	2	3	4.2	5.9
50.0%	Literature-I_s	High	0.2	0.3	0.6	0.9	1.4	2.1	3	4.2	5.9
50.0%	Literature-I_s	Medium	0.2	0.3	0.6	0.9	1.4	2.1	3.1	4.3	6
50.0%	Literature-I_s	Low	0.2	0.3	0.6	0.9	1.4	2.1	3	4.3	6
62.5%	Data from NL	All	0.5	1	1.8	2.9	4.6	6.8	9.9	13.8	18.5
62.5%	Data from NL	High	0.5	1	1.8	2.9	4.6	6.9	10	14	18.8
62.5%	Data from NL	Medium	0.5	1.1	1.9	3.1	4.9	7.4	10.8	15.1	20.4
62.5%	Data from NL	Low	0.5	1.1	1.9	3.1	4.9	7.4	10.8	15.2	20.5
62.5%	Literature-E	All	0.4	0.8	1.5	2.5	3.9	5.9	8.7	12.3	16.7
62.5%	Literature-E	High	0.4	0.8	1.5	2.5	3.9	6	8.8	12.4	17
62.5%	Literature-E	Medium	0.4	0.8	1.5	2.5	3.9	6	8.9	12.6	17.2
62.5%	Literature-E	Low	0.4	0.8	1.5	2.5	3.9	6	8.8	12.5	17.2
62.5%	Literature-I_s	All	0.3	0.5	0.8	1.4	2.2	3.4	5	7.2	9.9
62.5%	Literature-I_s	High	0.3	0.5	0.9	1.5	2.2	3.4	5.1	7.2	10.1
62.5%	Literature-I_s	Medium	0.3	0.5	0.9	1.4	2.2	3.4	5	7.2	10
62.5%	Literature-I_s	Low	0.3	0.5	0.8	1.4	2.2	3.4	5	7.2	10
75.0%	Data from NL	All	0.6	1.2	2.2	3.7	5.9	8.9	13	18.1	24.1
75.0%	Data from NL	High	0.6	1.2	2.2	3.7	5.9	9	13.1	18.4	24.6
75.0%	Data from NL	Medium	0.6	1.3	2.3	4	6.4	9.9	14.4	20.1	27
75.0%	Data from NL	Low	0.6	1.3	2.3	4	6.4	9.9	14.5	20.2	27.1
75.0%	Literature-E	All	0.4	0.8	1.4	2.4	3.8	5.8	8.4	11.8	16.1
75.0%	Literature-E	High	0.4	0.8	1.5	2.4	3.8	5.9	8.5	12	16.4
75.0%	Literature-E	Medium	0.4	0.8	1.5	2.4	3.9	5.9	8.5	12.1	16.4
75.0%	Literature-E	Low	0.4	0.8	1.5	2.5	3.9	5.9	8.6	12.1	16.5
75.0%	Literature-I_s	All	0.4	0.7	1.3	2.2	3.6	5.5	8.2	11.8	16.2
75.0%	Literature-I_s	High	0.4	0.7	1.3	2.2	3.6	5.6	8.3	11.9	16.4
75.0%	Literature-I_s	Medium	0.4	0.7	1.3	2.2	3.6	5.5	8.2	11.8	16.3
75.0%	Literature-I_s	Low	0.4	0.7	1.3	2.2	3.6	5.6	8.3	11.8	16.4

Figure C.7: *The percentage of the population that was infected up to and including March 12 for various parameter combinations. The columns denote the different virus transmission probabilities and every row a certain combination of the probability of developing symptoms, the case fatality ratio and the possibility of developing immunity.*

Scenarios			Transmission probability								
Symptoms	CFR	Immunity	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55
37.5%	Data from NL	All	0.6	0.9	1.3	2	2.8	4	5.5	7.5	9.9
37.5%	Data from NL	High	0.5	0.8	1.3	1.9	2.8	4	5.5	7.4	9.9
37.5%	Data from NL	Medium	0.5	0.8	1.3	1.9	2.8	4	5.5	7.5	10
37.5%	Data from NL	Low	0.5	0.8	1.3	1.9	2.8	4	5.5	7.4	9.9
37.5%	Literature-E	All	0.4	0.6	0.9	1.3	1.9	2.7	3.6	4.9	6.4
37.5%	Literature-E	High	0.4	0.6	0.9	1.3	1.9	2.6	3.6	4.8	6.4
37.5%	Literature-E	Medium	0.4	0.6	0.9	1.3	1.9	2.6	3.6	4.9	6.4
37.5%	Literature-E	Low	0.4	0.6	0.9	1.3	1.9	2.6	3.6	4.8	6.4
37.5%	Literature-I_s	All	0.2	0.2	0.3	0.5	0.7	1	1.4	1.9	2.5
37.5%	Literature-I_s	High	0.1	0.2	0.3	0.5	0.7	1	1.4	1.8	2.5
37.5%	Literature-I_s	Medium	0.1	0.2	0.3	0.5	0.7	1	1.4	1.8	2.5
37.5%	Literature-I_s	Low	0.1	0.2	0.3	0.5	0.7	1	1.4	1.9	2.5
50.0%	Data from NL	All	0.6	0.9	1.4	2.1	3.2	4.5	6.4	8.7	11.6
50.0%	Data from NL	High	0.4	0.6	1.1	1.7	2.5	3.7	5.2	7.3	9.8
50.0%	Data from NL	Medium	0.3	0.6	1	1.6	2.6	3.8	5.5	7.7	10.5
50.0%	Data from NL	Low	0.3	0.6	1	1.6	2.5	3.7	5.4	7.6	10.3
50.0%	Literature-E	All	0.4	0.7	1.1	1.7	2.5	3.5	5	6.9	9.2
50.0%	Literature-E	High	0.3	0.5	0.8	1.3	1.9	2.9	4.1	5.7	7.8
50.0%	Literature-E	Medium	0.3	0.5	0.8	1.3	1.9	2.8	4.1	5.7	7.7
50.0%	Literature-E	Low	0.3	0.5	0.8	1.3	1.9	2.8	4.1	5.7	7.7
50.0%	Literature-I_s	All	0.2	0.3	0.6	0.8	1.3	1.8	2.6	3.7	5
50.0%	Literature-I_s	High	0.1	0.2	0.4	0.6	1	1.5	2.1	3	4.2
50.0%	Literature-I_s	Medium	0.1	0.2	0.4	0.6	1	1.4	2.1	3	4.1
50.0%	Literature-I_s	Low	0.1	0.2	0.4	0.6	0.9	1.4	2	2.9	4.1
62.5%	Data from NL	All	0.6	1	1.7	2.7	4.1	6.1	8.6	11.9	16
62.5%	Data from NL	High	0.5	0.8	1.4	2.3	3.5	5.3	7.6	10.6	14.3
62.5%	Data from NL	Medium	0.4	0.8	1.4	2.3	3.6	5.5	8	11.2	15.2
62.5%	Data from NL	Low	0.4	0.8	1.3	2.2	3.5	5.4	7.8	11.1	15.1
62.5%	Literature-E	All	0.5	0.8	1.4	2.3	3.5	5.2	7.5	10.6	14.3
62.5%	Literature-E	High	0.4	0.7	1.2	1.9	3	4.6	6.7	9.4	12.9
62.5%	Literature-E	Medium	0.3	0.6	1.1	1.8	2.9	4.5	6.5	9.3	12.7
62.5%	Literature-E	Low	0.3	0.6	1.1	1.8	2.8	4.4	6.4	9.1	12.6
62.5%	Literature-I_s	All	0.3	0.5	0.8	1.3	2	3	4.3	6.1	8.4
62.5%	Literature-I_s	High	0.2	0.4	0.7	1.1	1.7	2.6	3.8	5.4	7.5
62.5%	Literature-I_s	Medium	0.2	0.4	0.7	1.1	1.7	2.6	3.8	5.4	7.5
62.5%	Literature-I_s	Low	0.2	0.4	0.7	1.1	1.7	2.6	3.8	5.4	7.4
75.0%	Data from NL	All	0.7	1.2	2	3.3	5.2	7.8	11.2	15.5	20.7
75.0%	Data from NL	High	0.6	1	1.8	3	4.8	7.2	10.4	14.6	19.6
75.0%	Data from NL	Medium	0.5	1	1.8	3.1	5	7.6	11.2	15.6	21
75.0%	Data from NL	Low	0.5	1	1.8	3	4.9	7.5	11	15.5	20.9
75.0%	Literature-E	All	0.5	0.8	1.4	2.2	3.4	5	7.3	10.1	13.7
75.0%	Literature-E	High	0.4	0.7	1.2	1.9	3	4.6	6.6	9.3	12.6
75.0%	Literature-E	Medium	0.3	0.6	1.1	1.8	2.9	4.4	6.4	9	12.3
75.0%	Literature-E	Low	0.3	0.6	1.1	1.8	2.9	4.3	6.3	8.9	12.2
75.0%	Literature-I_s	All	0.4	0.7	1.2	2	3.1	4.7	7	9.9	13.6
75.0%	Literature-I_s	High	0.3	0.6	1.1	1.8	2.9	4.4	6.5	9.3	12.8
75.0%	Literature-I_s	Medium	0.3	0.6	1.1	1.8	2.8	4.3	6.4	9.2	12.7
75.0%	Literature-I_s	Low	0.3	0.6	1.1	1.8	2.8	4.4	6.5	9.2	12.7

Figure C.8: *The percentage of immunity among the population based on individuals who were infected no later than March 12 for various parameter combinations. The columns denote the different virus transmission probabilities and every row a certain combination of the probability of developing symptoms, the case fatality ratio and the possibility of developing immunity.*