PLOS ONE

Modeling the spread of COVID-19 in Germany: Early assessment and possible scenarios

Maria Vittoria Barbarossa[1,3]*, Jan Fuhrmann[1,2], Jan H. Meinke[2], Stefan Krieg[2], Hridya Vinod Varma[3], Noemi Castelletti[4,5], Thomas Lippert[1,2,6]

1 Frankfurt Institute of Advanced Studies, Frankfurt, Germany, 2 Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany, 3 Interdisciplinary Center for Scientific Computing, University of Heidelberg, Heidelberg, Germany, 4 Institute of Radiation Medicine, Helmholtz Zentrum München, Neuherberg, Germany, 5 Statistical Consulting Unit StaBLab, Ludwig Maximilian University, München, Germany, 6 Department of Mathematics and Computer Science, Goethe University Frankfurt, Frankfurt, Germany

* barbarossa@fias.uni-frankfurt.de

Abstract

The novel coronavirus (SARS-CoV-2), identified in China at the end of December 2019 and causing the disease COVID-19, has meanwhile led to outbreaks all over the globe with about 2.2 million confirmed cases and more than 150,000 deaths as of April 17, 2020. In this work, mathematical models are used to reproduce data of the early evolution of the COVID-19 outbreak in Germany, taking into account the effect of actual and hypothetical non-pharmaceutical interventions. Systems of differential equations of SEIR type are extended to account for undetected infections, stages of infection, and age groups. The models are calibrated on data until April 5. Data from April 6 to 14 are used for model validation. We simulate different possible strategies for the mitigation of the current outbreak, slowing down the spread of the virus and thus reducing the peak in daily diagnosed cases, the demand for hospitalization or intensive care units admissions, and eventually the number of fatalities. Our results suggest that a partial (and gradual) lifting of introduced control measures could soon be possible if accompanied by further increased testing activity, strict isolation of detected cases, and reduced contact to risk groups.

Introduction

In late December 2019, several cases of acute respiratory syndrome were first reported in Wuhan City (Hubei region, China) by Chinese public health authorities. A novel coronavirus was soon found as the main causative agent. It is now known as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The disease caused by SARS-CoV-2, which rapidly spread first through China and then to other countries, is now referred to as coronavirus disease 2019 (COVID-19). The World Health Organization declare COVID-19 a global pandemic on March 11, 2020 [1, 2]. As of April 17, 2020, about 2.2 million cases and more than 150,000

deaths have been reported worldwide [2, 3]. First cases in Germany were reported at the end of January 2020. As of April 17, 2020, the Robert Koch-Institute (short: RKI) counts over 130,000 confirmed infections and about 4,000 deaths in Germany.

Due to the novelty of the virus, it is reasonable to assume that nobody has prior immunity, that is, the entire human population is potentially susceptible to SARS-CoV-2 infection [1]. Droplet transmission, occurring when a susceptible person comes in close contact (droplets are generated by coughing, sneezing or talking) with an infective, seems to be the main infection route of SARS-CoV-2 in the population. Contact transmission, via surfaces in the immediate environment or objects used by an infectious person may also be possible [4, 5].

After an incubation period which varies from 2 to 14 days, first unspecific symptoms, such as fever, cough, sore throat, or muscular pains, might appear. The illness can then become more acute and lead to difficulties in breathing or progress into severe pneumonia. In critical cases, multi-organ failure can follow and lead to death [5].

Chinese data on the first two months of the outbreak report that about 80% of observed cases are mild to moderate infections [6]. According to the distribution of reported cases in early April in Italy [7], about one third are either completely asymptomatic ($\sim$7%), pauci-symptomatic ($\sim$15%), or non-specific infections ($\sim$13%), about 41% are mild respiratory infections, which do not require hospitalization, whereas about one fifth showed severe symptoms requiring hospitalization ($\sim$20%), and about 3% of the reported cases required intensive care. Severe illness and death are more common among the elderly or patients with other chronic underlying conditions [8]. Nevertheless, severe illness and death can also occur in young and healthy persons [9].

Whether asymptomatic infections contribute to the spread of the virus is yet to be clarified [5]. One study on the first COVID-19 cases near Munich reports that the infection appeared to be transmitted during the incubation period of the patients, and suggests reconsidering the role of asymptomatic persons in the transmission dynamics of the current outbreak [10], as does another study of patients in Guangzhou Eighth People's Hospital [11].

There is to date no specific treatment nor vaccine against SARS-CoV-2 [1]. Nevertheless, first studies showed that people infected with the virus develop specific antibodies [12], and based on previous knowledge of other coronaviruses, like SARS or MERS, it is possible that also SARS-CoV-2-specific immunity will last for about three years [6].

Since January 2020, several mathematical models for understanding and controlling the spread of SARS-CoV-2 have been proposed, in particular, concerning the Chinese outbreak [13–17] and the risk for this to spread to other countries [18, 19]. In all affected regions, non-pharmaceutical interventions to reduce virus transmission were introduced. Mathematical methods were developed to study the impact of such interventions [20] and simulate the effects of relaxing control measures [21]. Possible scenarios for the spread of SARS-CoV-2 in Germany were considered in preliminary studies by an der Heiden and Buchholz [22] who projected the effect of various model parameters (both related to unknown properties of the pathogen and the disease and to the effect of control measures) on the number of cases, in particular of those requiring hospitalization and intensive care, as well as deaths.

Among the key parameters to be approximated by a mathematical approach is the so-called basic reproduction number ($\mathcal{R}_0$), a metric that indicates the average number of secondary infections generated in a fully susceptible population by one infectious individual over the course of the infection. The basic reproduction number is a reference parameter in mathematical epidemiology used to understand if, and in which proportion, a disease will spread among the population. In simple disease transmission models, $\mathcal{R}_0 > 1$ implies that the virus will spread in the population, and the larger the value of $\mathcal{R}_0$, the harder it is to control the epidemic. First studies on the initial outbreak in China estimated that the basic reproduction

number of COVID-19 could range between 2.2 [13, 14] and 4.71 [15], or could even be larger than 6 [16]. The *(effective) reproduction number* $\mathcal{R}$ for a given population is the corresponding time-dependent quantity which reflects the number of secondary infections generated by one infectious individual in the current population, and is affected by intervention measures aimed at controlling the spread of the disease. Among the first studies to assess the practical implications of public health interventions, Tang et al. [16] identified contact tracing followed by quarantine and isolation, as well as travel restrictions as the most effective measures to contain the epidemic. Meyer-Hermann and coauthors [23] recently predicted the evolution of the reproduction number for the spread in Germany, with detailed analysis for all federal states and observed that as of April 3, 2020 the reproduction number $\mathcal{R}$ was lowered to values near 1 in all federal states.

In this study, we predict the spread of COVID-19 among the German population by means of mathematical modeling, simulating the implementation or withdrawal of non-pharmaceutical interventions. First model results were presented in our preliminary work [24] and implicated in the advisory paper [25]. Here we present a follow-up of our study, including case and fatality data, and information about testing activity [26], as of beginning of April 2020. The proposed setting allows investigating how a specific intervention scenario affects the dynamics of the epidemic, with particular attention to interactions between individuals of the same or different age groups (children, adults, and people older than 60 years). We consider the following scenarios:

- **Minimal intervention**: The main factor slowing down disease transmission is people's increased awareness in response to initial recommendations coming from health institutions and local governments, and to media coverage (e.g., washing hands, proper coughing and sneezing, keeping distance from obviously sick persons, limited (self-)quarantine of known or suspected cases);

- **Baseline scenario:** Hitherto adopted main control measures (closure of schools and universities, remote working policy, isolation of identified cases, contact restrictions, a partial economic shutdown, and levels of testing activity as of March 15, 2020) are assumed to be maintained throughout 2020. Model parameters which include such control measures were calibrated on reported cases in Germany as of April 5, 2020 and were used to project data until April 14, 2020;

- **High vigilance**: This scenario is obtained by enriching the *baseline* measures with significantly increased testing activity (not only suspected COVID-19 cases but also persons without symptoms or known close contacts to identified cases) and a strict isolation protocol of detected COVID-19 cases for about two weeks;

- **Educational/economic reopening**: Partial lifting of the restrictions imposed thus far, gradually reopening schools, universities, and resuming economic activities, though largely maintaining remote working policy and limiting use of public transport service and organized club activities. Significantly increased testing activity and strict isolation protocol of detected COVID-19 cases are maintained;

- **Phase-out**: Gradual lifting of most control measures applied so far (reopening schools, resuming work, regular public transport service, resuming most social and economic activities). Increased testing activity and strict isolation protocol of detected COVID-19 cases are maintained, and the population is assumed to uphold minimal awareness measures;

- **Cautious phase-out**: As the *phase-out* scenario, but with slower rollback to educational, social and economical activities, and with improved measures to protect elderly and people at risk.

Our model predicts that the current control measures are necessary to slow down or even suppress the spread of the epidemic and that the removal of restrictions in favor of social and economic activities will accelerate the growth of case numbers unless it is accompanied by a significantly strengthened testing and case isolation policy. However, under such increased vigilance, combined with particular care regarding patients at high risk, we project that a gradual phase out of the most severe measures starting around May 5 will lead to a progression of the epidemic that is sufficiently slow to be handled by the health care system.

Mathematical models and methods

The mathematical models adopted for this study are based on systems of differential equations that describe interactions between different groups of individuals in the population. The proposed approach extends the known S-E-I-R (susceptibles-exposed-infected-recovered) model for disease dynamics [27]. Individuals are classified according to their status with respect to the virus spread in the community. In particular we distinguish between individuals who have been exposed to the virus but are not yet infectious, asymptomatic infectives, infectives with mild or influenza-like symptoms (not reported as SARS-CoV-2 infections), and reported SARS-CoV-2 infectives. We assume that infected patients without SARS-CoV-2 diagnosis are unlikely to die of the virus-induced disease. Undetected infections lead to undetected recoveries in the population, which cannot be reported unless testing for ongoing (virus detection) or previous (antibody detection) SARS-CoV-2 infections is performed.

The core model—Homogeneous population

The most basic approach that we adopted to understand the evolution of the epidemic in time is based on the assumption that the population is homogeneous (in particular with respect to space, as well as to age and social habits of individuals). Of course, this simplistic assumption does not reflect the multidimensional complexity of the ongoing situation, but it can help in determining major factors affecting disease spread. Individuals are classified as follows. Susceptible individuals (S) are those who can be infected with SARS-CoV-2. Exposed individuals (E) have been infected with SARS-CoV-2, though they are not yet infectious, nor symptomatic. After a latent phase (on average about 5.5 days from infection [6]) the first COVID-19 symptoms occur (with probability ρ_0) and the exposed individual becomes infectious. We distinguish between asymptomatic undetected (U) COVID-19 patients, symptomatic but not yet detected infectives (I) and reported COVID-19 cases (H). For individuals developing symptoms, the probability of the disease being detected immediately is η_0, the probability of detection at a later stage is denoted by η_1. Individuals who recovered from a detected (R) or an undetected (R_u) infection, as well as patients who died from the infection (D), are removed from the chain of transmission. The transmission diagram of the model is depicted in Fig 1. Susceptible individuals can be infected via contacts with asymptomatic (transmission rate β_U), symptomatic undetected (β_I) and reported cases (β_H). We assume that asymptomatic infectives do not restrict their contacts to others, and therefore have higher transmission rates than symptomatic infected individuals. Detected cases are supposed to reduce their contacts even further. Due to limitations in the identifiability of the parameters with the available data, we fix the ratio between β_I, respectively β_H and β_U and estimate the latter. Duration of latency ($1/\gamma_E$) and infectious periods ($1/\gamma_I$, $1/\gamma_H$ and $1/\gamma_U$, respectively) are assumed in accordance with available literature. Further details on parameter assumptions are given in Table 1. The dynamics of the core model described above and shown in Fig 1 is given by the following system of
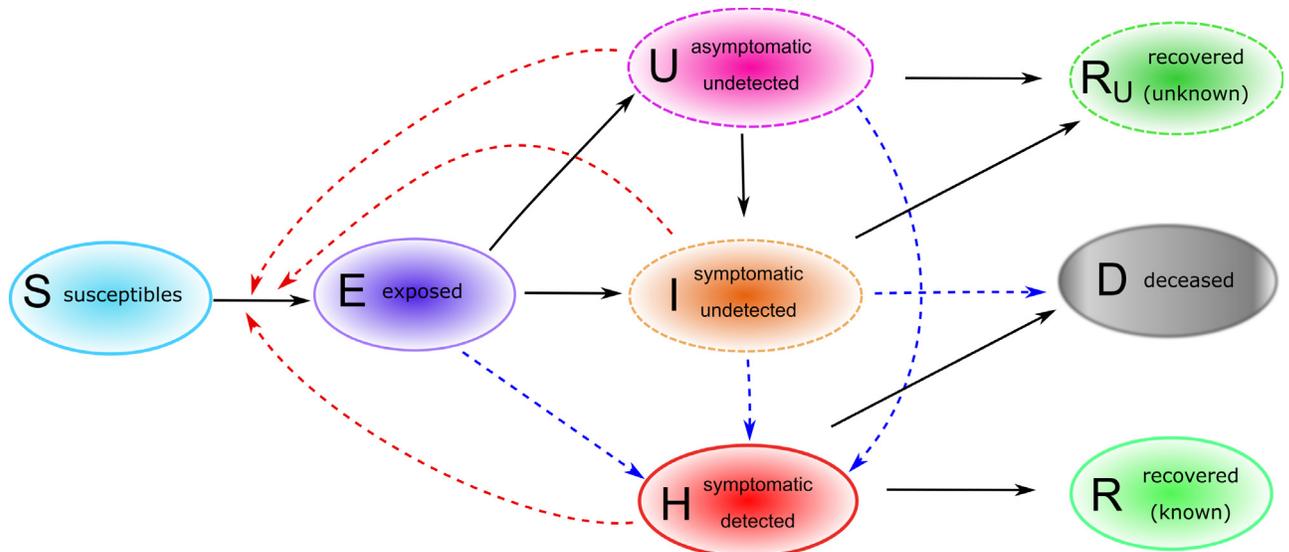
Fig 1. Core model structure for COVID-19 outbreak in Germany. Solid arrows indicate transition from one compartment to another, red dashed arrows indicate virus transmission due to contact with infectives, blue dashed arrows indicate detection of infections due to testing activities. Upon infection with SARS-CoV-2, susceptible (S) individuals enter a latent phase (E), in which they are not yet infectious nor symptomatic. After the latent phase, individuals become infectious, may develop symptoms and may be detected as COVID-19 cases. We distinguish between asymptomatic undetected (U), symptomatic undetected (I) and symptomatic detected (H) infections. Infected individuals who recovered from a detected (R) or an undetected (R_U) infection, as well as patients who died (D) upon infections, are removed from the chain of transmission.

https://doi.org/10.1371/journal.pone.0238559.g001

differential equations:

$$
\begin{aligned}
\dot{S} &= -\lambda S & \text{susceptibles} \\
\dot{E} &= \lambda S - \gamma_E E & \text{exposed} \\
\dot{U} &= (1-\rho_0)\gamma_E E - \gamma_U U & \text{asymptomatic infected} \\
\dot{I} &= (1-\eta_0)\rho_0\gamma_E E - \gamma_I I & \text{symptomatic infected} \\
\dot{H} &= \eta_0\rho_0\gamma_E E + \hat{\eta}_1\gamma_U U + \eta_1\gamma_I I - \gamma_H H & \text{diagnosed cases} \\
\dot{R} &= (1-\delta_H)\gamma_H H & \text{recovered from detected infection} \\
\dot{D} &= \delta_H\gamma_H H & \text{deceased} \\
\dot{R}_u &= (1-\eta_1)\gamma_I I + (1-\hat{\eta}_1)\gamma_U U & \text{recovered from undetected infection,}
\end{aligned}
\tag{1}
$$

Table 1. Main model parameters in the core model. Parameter descriptions, values and references.

	Description	Value (unit) [Refs.]
β_U	Transmission rate of asymp. undetected infectives	fitted, $1/(\text{d} \cdot \text{individual})$
β_I	Transmission rate of sympt. undetected infectives	$0.8\beta_U$ (assumed)
β_H	Transmission rate of detected infectives	$0.1\beta_U$ (assumed)
$1/\gamma_E$	Mean incubation period	5.5 d [22, 33, 34]
$1/\gamma_I$, $1/\gamma_H$	Mean duration of symptomatic infection	7 d [6, 23, 35]
$1/\gamma_U$	Mean duration of asymptomatic infection	6 d [6, 35]
δ_H	Case fatality ratio of detected infections	0.057 [8, 23]
ρ_0	Probability of developing symptoms	0.67 [7]
η_0	Probability of detection during latency	fitted
$\hat{\eta}_1$	Probability of detection while asymptomatic	0.07 (assumed)
η_1	Probability of detection while symptomatic	fitted

https://doi.org/10.1371/journal.pone.0238559.t001

where

$$\lambda = (\beta_I I + \beta_U U + \beta_H H)/(N - D)$$

and $N \approx 83$ million is the total population.

The basic reproduction number of system (1) can be calculated analytically, e.g, by means of the next-generation matrix approach [28], and is given by

$$\mathcal{R}_0 = \frac{(1 - \eta_0)\rho_0 \beta_I}{\gamma_I} + \frac{\beta_H}{\gamma_H}(\eta_0 \rho_0 + (1 - \eta_0)\rho_0 \eta_1 + (1 - \rho_0)\hat{\eta}_1) + \frac{\beta_U(1 - \rho_0)}{\gamma_U}. \tag{2}$$

The effective reproduction number $\mathcal{R}$ at time t can be obtained by the same formula, substituting the values of time-dependent parameters in (2) and multiplying this value by the susceptible fraction of the population, $S(t)/(N - D(t))$.

To determine model parameters that could not be inferred from the literature, we use the sum of reported cases from three different time periods (February 28 to March 11, March 11 to March 22, and March 22 to April 11) and fit the data to the sum of H, R, and D using the trust region reflective method [29] as implemented in SciPy [30]. The first period does not include any global contact reducing measures except for school closings in the county of Heinsberg. The second period includes nationwide school closings and an increased use of working from home where possible. The third period includes the closing of most stores and general constraints on contacts in the public sphere. For each period we fit η_1, $1 - \eta_0$, and β_U, while keeping β_I and β_H proportional to β_U with proportionality factors given in Table 1. Fitted model solutions at the end of each time interval are taken as initial values for integration of system (1) over the next time period. To test the sensitivity of the results to variations in the estimated parameters, we add Monte Carlo sampling of the parameters generating a large ensemble of model fits. Each of these fits contributes to the solution according to its Akaike weight. The Akaike weight is determined based on the Akaike information criterion [31]. For few data points the corrected Akaike information criterion (AICc) [32] takes the form

$$\text{AICc} = n\left[\ln\left(\frac{\text{SSE}}{n}\right)\right] + 2k + \frac{2k(k + 1)}{(n - k - 1)}, \tag{3}$$

where SSE is the sum of squared errors, n is the number of data points, and k is the number of degrees of freedom. The weight of each fit i in a set of J fits is then given as

$$w_i = \frac{e^{-\Delta_i/2}}{\sum_{j=1}^{J} e^{-\Delta_j/2}}, \tag{4}$$

where $\Delta_i = \text{AICc}_i - \text{AICc}_{\min}$. Based on the sample values and their weights, we can construct a histogram for each parameter and derived properties (c.f. Fig 2). A narrow peaked histogram indicates a well determined parameter with a small standard deviation and a corresponding narrow confidence interval, whereas a flat histogram shows that the parameter may be indeterminate and that the fit is not sensitive to its variations. For a sufficiently large number of points, we can determine, for example, a 95% confidence interval directly from the histogram by excluding the left- and rightmost 2.5% (marked in light blue in Fig 2).

Sensitivity analysis of $\mathcal{R}$ and peak cases in the core model

Sensitivity analysis of the reproduction number and of the number of cases at the outbreak peak was further investigated by means of scatter plots (results not shown here) and Sobol analysis (SALib library [36] in Python). The Sobol method is a variance-based method for
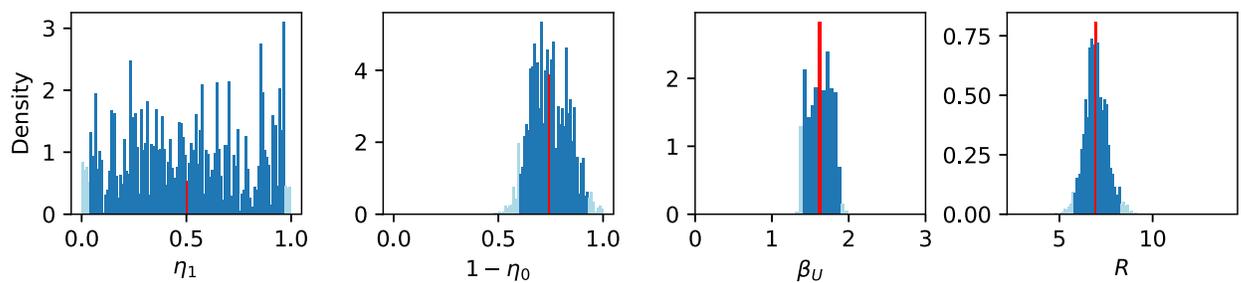
Fig 2. Weighted histograms of η_1, $1 - \eta_0$, β_U, and $\mathcal{R}$ (from left to right) based on stochastic variation of η_1, $1 - \eta_0$, and β_U. Each sample contributes with its Akaike weight, which is based on a fit to the data from February 28 to March 11, 2020, to the histogram (see Eq (4)). Dark blue bins lie within the 95% confidence interval, light blue bins lie outside. Apparently, η_1 cannot be well identified using these data. The red bin indicates the mean value (of $1 - \eta_0 = 0.74$ (0.59–0.92), of $\beta_U = 1.63$ (1.39–1.88), and of $\mathcal{R} = 6.95$ (5.76–8.25)).

https://doi.org/10.1371/journal.pone.0238559.g002

global sensitivity analysis that decomposes the variance or sensitivity of the model output into contributions of individual parameters or groups of parameters [37, 38]. The first-order sensitivity index ($S1$)—also called main effect—is the ratio of the partial variance of an individual parameter with respect to the total variance. The total-order sensitivity index (ST)—also called total effect—measures the overall influence (including higher-order interactions) of each parameter on the model output. The values of both $S1$ and ST range between 0 and 1, where higher values indicate a greater contribution of the parameter to the model output [39]. When the total effect due to a parameter is much larger than its main effect, one may need to look at higher-order sensitivity indices to look for interactions between parameters. However, this was not the case for the parameters that we have considered (results not shown here).

We computed the sensitivity indices of the parameters that could not be determined from literature or are expected to vary between populations and over time, viz the transmission parameters, the probability of developing symptoms, and the probability of being detected. Of these parameters ρ_0, η_0, and η_1 are varied between 0 and 1, β_U is varied between 0 and 3, and β_I and β_H are assumed to be fixed multiples of β_U as detailed in Table 1. All the parameters are varied simultaneously to also observe their interaction effects.

The first-order sensitivity indices ($S1$) indicate that β_U has the expected very strong effect on the calculated reproduction number ($S1(\beta_U) = 0.87\pm0.07$). Also, the probability of developing symptoms, ρ_0, has an effect on $\mathcal{R}$, but this is much smaller than the effect of β_U ($S1(\rho_0) = 0.04\pm0.02$). Early-stage detection also has a similar small impact ($S1(\eta_0) = 0.05\pm0.03$), while the probability of detection while symptomatic has an even lower impact ($S1(\eta_1) = 0.002 \pm0.005$), confirming the non-identifiability of this parameter. Both the transmission rate of asymptomatic individuals and the probability of developing symptoms seem to have strong effects on the maximum number of active reported infections ($S1(\beta_U) = 0.32\pm0.05$, $S1(\rho_0) = 0.42\pm0.06$). Detection probabilities also have a small impact on the maximum number of active infections reported ($S1(\eta_0) = 0.05\pm0.03$, $S1(\eta_1) = 0.05\pm0.03$).

Population structured by age and stages of infection

Refining the core model described above, we include age groups and stages within infective compartments. This allows considering features like the immune response of individuals during infection or social behavior, in particular interactions among individuals of the same or different age groups. Based on the statistical analysis of the RKI data in our previous study [24], we distinguish three groups: children (0-14y), adults (15-59y) and people 60y or older. For

each age group the model tracks susceptibles (S), exposed (E), undetected asymptomatic infectives (U), undetected symptomatic infectives (I), diagnosed infectives (H), recovered (R), undetected recovered (R_u), and deceased (D) individuals, as detailed above for the core model. To obtain more realistic distributions for the duration of the exposed and infective compartments, we split each of these in three stages (E_j, U_j, I_j, H_j, $j = 1, 2, 3$). This is a classical extension of the standard disease transmission model to account for non-exponential distributions of incubation and infectious periods (cf., e.g., [18] for another example of application in modeling COVID-19). This results in a total of one plus nine infective compartments per age class (stage E_3 is assumed to be infective as well, as individuals have been reported to be infectious before symptoms onset [6, 11]). The age classes evolve in parallel (maturation during the course of the outbreak is neglected along with demographics in general), but are coupled to one another by contact rates and disease transmission among individuals of different age groups.

After virus transmission, an exposed individual is assumed to travel through the stages E_1, E_2, and E_3 before entering either of the infection stages I_1, U_1, or H_1. In the course of the disease the individual will then pass through the infection stages of the respective compartment. Both symptomatic and asymptomatic infectives in stages U_i or I_i can enter stages H_{i+1} or H_i, respectively, by being tested. The probability of being tested and positively diagnosed is assumed to be larger for symptomatic than for asymptomatic individuals. Since individuals at stage E_3 are already infectious [11], late symptom onset can be viewed as a prolongation of the latency period. We have therefore omitted any transitions from U to I. The individual leaves the last stage of infection (U_3, I_3) by either recovering to R or R_u, depending on whether the infection has been diagnosed. Given the current efforts to detect and control the spread of the disease, it seems a justified assumption that the vast majority of deaths caused by SARS-CoV-2 are investigated, hence we assume only diagnosed individuals (H) to be fatally affected by the disease. Even in the case that the presence of the virus is only discovered postmortem, we model this as detection followed by death (transition I to D).

Equations and parameters of the age structured model. For each age class A we have the following compartments

$$
\begin{aligned}
\dot{S}^A &= -\lambda^A S^A &&\text{susceptibles} \\
\dot{E}_1^A &= \lambda^A S^A - \gamma_{E1}^A E_1 &&\text{exposed stage 1} \\
\dot{E}_j^A &= \gamma_{E(j-1)}^A E_{j-1}^A - \gamma_{Ej}^A E_j^A &&\text{exposed stage } j = 2, 3 \\
\dot{U}_1^A &= (1-\rho_0^A)\gamma_{E3}^A E_3^A - \gamma_{U1}^A U_1^A &&\text{asymptomatic stage 1} \\
\dot{U}_j^A &= (1-\hat{\eta}_{j-1}^A)\gamma_{U(j-1)}^A U_{j-1}^A - \gamma_{Uj}^A U_j &&\text{asymptomatic stage } j = 2, 3 \\
\dot{I}_1^A &= (1-\eta_0^A)\rho_0^A\gamma_{E3}^A E_3^A - \gamma_{I1}^A I_1^A &&\text{infected stage 1} \\
\dot{I}_j^A &= (1-\eta_{j-1}^A)\gamma_{I(j-1)}^A I_{j-1}^A - \gamma_{Ij}^A I_j^A &&\text{infected stage } j = 2, 3 \quad (5) \\
\dot{H}_1^A &= \eta_0^A\rho_0^A\gamma_{E3}^A E_3^A + (1-v_1^A)\eta_1^A I_1^A - \gamma_{H1}^A H_1^A &&\text{diagnosed stage 1} \\
\dot{H}_j^A &= v_{j-1}^A\eta_{j-1}^A\gamma_{I(j-1)}^A I_{j-1}^A + \hat{\eta}_{j-1}^A\gamma_{U(j-1)}^A U_{j-1}^A \\
&\quad + (1-v_j^A)\eta_j^A\gamma_{Ij}^A I_j^A + \gamma_{H(j-1)}^A H_{j-1}^A - \gamma_{Hj}^A H_j^A &&\text{diagnosed stage } j = 2, 3 \\
\dot{R}^A &= (1-\delta_H^A)\gamma_{H3}^A H_3^A &&\text{recovered (detected)} \\
\dot{D}^A &= \delta_H^A\gamma_{H3}^A H_3^A &&\text{deceased} \\
\dot{R}_u^A &= (1-\eta_3^A)\gamma_{I3}^A I_3^A + \gamma_U^A U_3^A &&\text{recovered (undetected)}
\end{aligned}
$$

The transmission rate for S^A individuals is given by

$$\lambda^A = \frac{1}{N - \sum_B D^B} \sum_{B,X} \beta_X^{B,A} X^B,$$

where N is the total population and $\beta_X^{B,A}$ denotes the transmission rate due to contacts of infectious individuals in compartment $X \in \{E_3, U_j, I_j, H_j | j = 1, 2, 3\}$ and age class B, X^B with susceptibles in age class A. Contact rates can be summarized in a (3×30) matrix $\mathcal{B} = (\beta_X^{B,A})_{X,B,A}$. Entries of the contact matrix $\mathcal{B}$ are assumed to be influenced by (i) the overall aggressiveness of the virus, β_0, that describes how easily the virus is transmitted and is a common factor for all entries of $\mathcal{B}$, (ii) an age specific susceptibility factor σ^A that describes the average biological susceptibility to the virus and the average activity level (in terms of social contacts) of the given age group, A, and is a common factor for each row of $\mathcal{B}$, (iii) the effective infectivity ι_X^B of each given compartment X and age group B that depends on the activity level and the plain infectivity of X^B, and is a common factor in each column of $\mathcal{B}$, and finally (iv) a likeness factor φ^{BA} that describes the mixing of age group B with age group A. The factor β_0 has to be estimated by fitting the model to data provided by the RKI [34]. The same is true for the age specific susceptibilities that depend on the immune response of the individual (supposedly on average stronger in younger individuals). The effective infectivity ι_X^B is supposed to be lowest for diagnosed individuals (H_j) thanks to deliberate contact reduction. We also assume symptomatic individuals to be less active than asymptomatic ones which means that ι_X^B is smaller for $X = I_j$ than for $X = U_j$, even though the biological infectivity should be higher for symptomatic individuals than for asymptomatic ones. This activity reduction results from both the symptoms restricting the mobility of an affected individual as well as them being aware of their potential to spread a communicable disease. The likeness factor is hard to estimate. Following the work by Prem et al. [21], who focused on the population in Hubei province in China, we consider contacts in three different realms: at home, at work and school, and in other locations, most notably during leisure activities. For each activity, we assume specific contact distributions depending on the ages of the individuals. For example, we assume high contact rates among adults and between adults and children at home. The separation of the realms allows for each intervention to produce a specific effect. For example, school closures as part of the CS intervention (cf. Table 2) predominantly act on the child-child contacts in the school realm. The resulting matrix $\mathcal{B}$ under the influence of different interventions is graphically illustrated in Fig 3. Both the assumptions and the final overall contact matrices are largely in keeping with the findings of, e.g., [40] for general communicable diseases.

The progression rates γ_X are determined by the estimated mean durations of each stage of the infection. For simplicity, and analogously to [18], we choose all the rates for a given compartment to be the same (that is, $\gamma_{E1} = \gamma_{E2} = \gamma_{E3} =: \gamma_E$ and likewise for U, I, and H). In accordance with [35] we assume the mean incubation period to be 5.5 days and therefore put $\gamma_E = 3/5.5$, the factor 3 giving a mean combined duration of stay in E_1, E_2, and E_3 of 5.5 days. In the same manner we take the rates $\gamma_U = 3/7$ and $\gamma_H = \gamma_I = 3/8$ corresponding to mean durations of the infectious period of seven days for asymptomatic individuals and eight days for symptomatic ones, in accordance with previous studies [17, 21].

The probability ρ_0 of developing symptoms at the end of the incubation period is estimated to be $\rho_0^j = 0.25$ for juniors who are reportedly often asymptomatic [41], and $\rho_0^a = 0.7$ and $\rho_0^s = 0.8$ for adults and seniors, respectively.

The probability η_0 of an infection being discovered by the end of the incubation period is assumed to be small in the beginning of the simulation. The same is true for the rates $\hat{\eta}_m$ for discovering an infection in the absence of symptoms. For symptomatic individuals, these rates

Table 2. Summary of control measures for mitigation of COVID-19 infections adopted in Germany as of April 4, 2020. The indicated percentages are meant to be compared to the state without the respective measure (i.e., the whole has a multiplicative effect).

Label	Policy	Description/Working assumptions
CS	Closure of all schools, universities, sport clubs. Public events are canceled. Stricter social distancing encouraged	Overall reduction of activities, decreasing all entries by 10% (social distancing). Moreover, a strong effect on the shape matrix reducing child-child contacts by 80% (resulting from no contacts in the school realm), adult-adult contacts by 20% (students in high school and college, club activities, public events), and senior-senior by 10%. Adopted nationally from March 14 (day 28), fully in force March 16 (day 30).
HO	Remote working policy (home office), closure of all restaurants and bars, reduced (use of) public transport services.	Reducing contacts: child-child (- 20%), child-adult (- 20%), adult-adult (- 50%), child-senior (- 30%), adult-senior (- 30%), senior-senior (- 30%). Assumed to be applied at national scale from March 13 (day 27), contact reductions fully achieved after 13 days.
M	Awareness raising due to media, canceling big events	Information and media activities increase social distancing and personal hygiene (e.g., washing hands), starting on February 25 (day 10). Contacts are on average reduced by 20%, achieved by activity reduction (higher relative decrease for seniors than for juniors and adults).
T0	Testing	Increased testing activity starting February 28 (day 13), most significant increase between March 9 and March 15 (days 23 through 29)Approx. 10-fold increase in detection rates compared to previous levels.
IC0	Isolation of infected cases	Identified and suspected infected cases (self) quarantined for 2 weeks. Reduces *effective* contact rates (for individuals in specific groups by up to 40%); implemented starting on February 28 (day 13), in effect by March 7 (day 21).
LD	Partial economic and social activities	Closing non-critical businesses, gatherings of more than two or three persons sanctioned, decreases contact rates by another average 30%, starting March 22 (day 36), in full effect by March 27 (day 41).

https://doi.org/10.1371/journal.pone.0238559.t002

are larger as symptoms provide a strong suspicion of being infected. We assume a certain age dependence for testing. On the one hand, seniors are at high risk and might be expected to be tested more frequently but on the other hand, severe respiratory tract infections are common in people with limited immune competence and may not be taken as implicating a SARS-CoV-2 infection. Juniors are reportedly not seriously affected, hence we assume that they are less frequently tested. According to recent RKI reports [26], testing has been dramatically increased between March 9 and March 15, so we assume that all the testing rates mirror this increase. In particular, we assume the rates η_0 and $\hat{\eta}_{1/2}$ to be significantly larger than zero since individuals are encouraged to go for a test, even in the absence of symptoms if they were in close contact with a known infective.

The parameters ν_m describe the probability of progressing to the next stage while being diagnosed (i.e. transitioning from an I to an H compartment). Their values are assumed to be close to 1. Them being smaller would imply the assumption that being tested is somehow associated with more severe symptoms and would, therefore, slow down the progression toward recovery.

To follow weekly oscillations in the reported cases, which show a regular slump over the weekends, we include time dependent test rates. This allows the simulated case numbers to

(a) baseline

(b) cautious phase-out
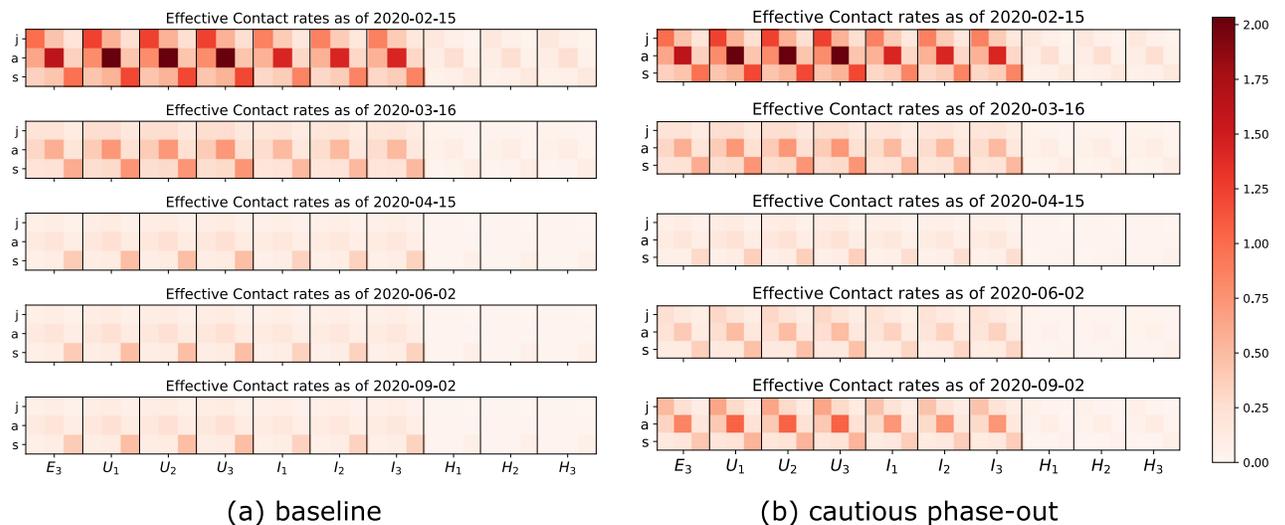
Fig 3. The evolution in time of contacts between juvenile (j), adult (a) and senior (s) susceptible individuals (columns) and infectious individuals in the late latent phase (E_3) or one of the infected phases ($I_1, \ldots H_3$). (a) In *baseline* scenario control measures will be maintained over time, hence contacts and transmission of the virus will remain low, whereas (b) a lift of all control measures (*cautious phase-out* scenario) after April 15 will slowly lead to a new increase in contacts and virus transmission.

closely follow the data as illustrated in Fig 4. For long term simulations of scenarios we assumed these fluctuations to fade away and removed the time-dependency from the testing rates. This implicitly represents an assumed increase in testing activities, and de facto increases the reporting ratio by more than 10%.
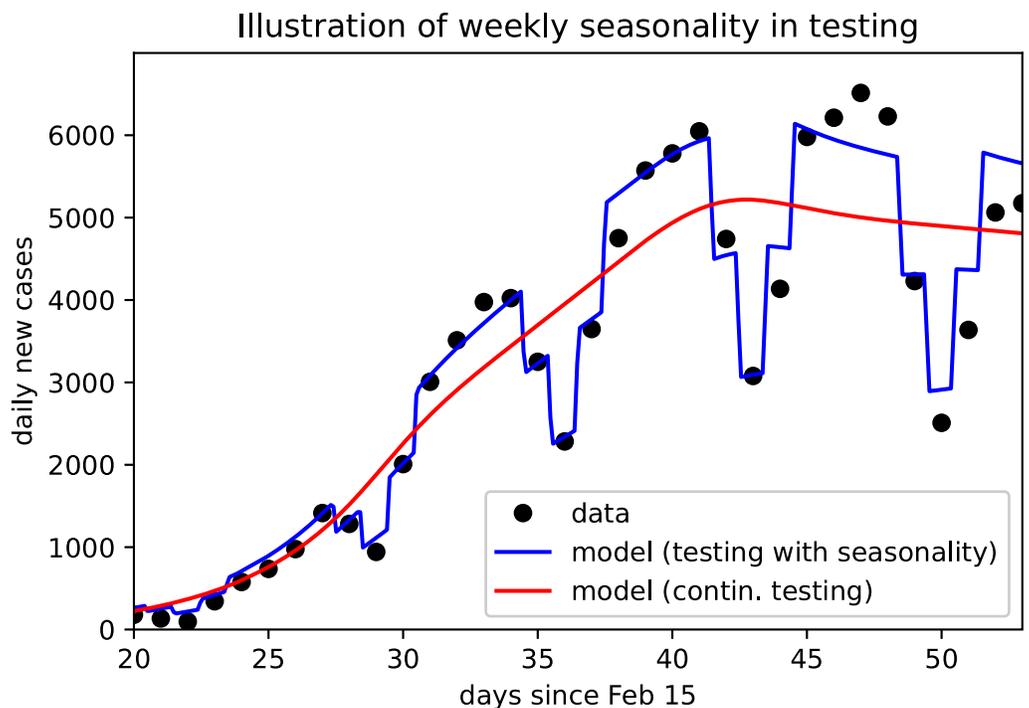

Fig 4. The *weekend effect* accounting for fluctuations in case detection and reporting. Black dots denote daily reported new cases, continuous curves show the model solutions with (blue) and without (red) time-depending testing rates.

Modeling control measures

In the structured model we can include control measures that explicitly affect different age groups. In particular we consider (i) general increased awareness in the population due to the effect of media (M), as well as (ii) active control due to main intervention measures adopted throughout Germany since March 2020. These control measures include: (CS) Closure of all schools, universities, sports clubs and canceling public events; (HO) reduced contacts in workplaces and outside the household (restaurants, bars, public transport); (T0) initial efforts to improve detection by more testing; (IC0) isolation of infected cases; and lock-down measures (LD) closing most activities and prohibiting more than two people gatherings in effect by the end of March. Details are summarized in Table 2.

Reported cases for the three age groups were used to fit the model including control measures as indicated in Table 2 until April 5 (Fig 5). The obtained setting was used to predict data from April 6 to April 14, as well as long term predictions of the *baseline*(BSL) scenario, in which the control measures applied as of early April (CS, HS, TO, ICO and LD) and the awareness due to the effect of media are maintained until the end of the year (Fig 6(b)). The *baseline* scenario has then been modified to simulate different possible future scenarios (cf. Table 3). On the one hand, we have enriched the current control measures by further increased testing activity (T1+), stricter isolation of known cases (IC) and increased social distancing from individuals at risk (IO). On the other hand, we have considered possible rollback scenarios (cf. Results).
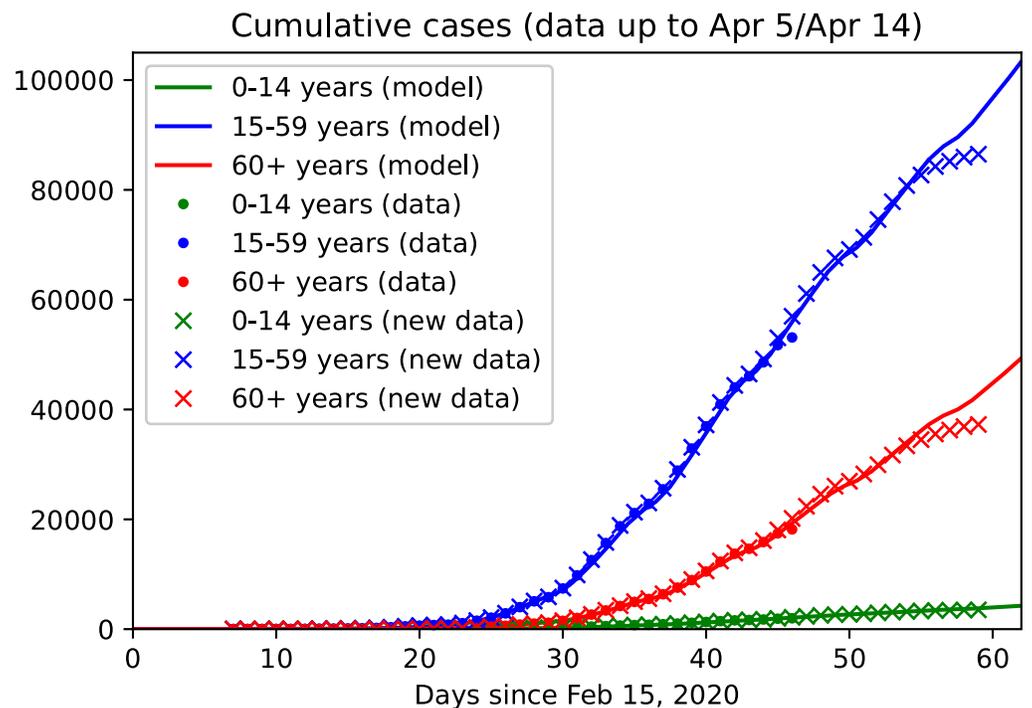
Fig 5. Age and stage structured model: Data fit and predictions. Continuous curves show model solutions, dots are reported data up to March 25 (as used in our previous work [24]), cross denotes reported data as of April 15. The model is calibrated on collected data up to April 5 (day 50), data from April 6 to 14 are used for validation. Colors denote the three different age groups: juveniles (0-14y, green), adults (15-59y, blue) and seniors (60y and older, red). It should be noted that the most recent data tend to be lower than expected since not all cases detected on these days have been reported to the RKI, yet.

https://doi.org/10.1371/journal.pone.0238559.g005
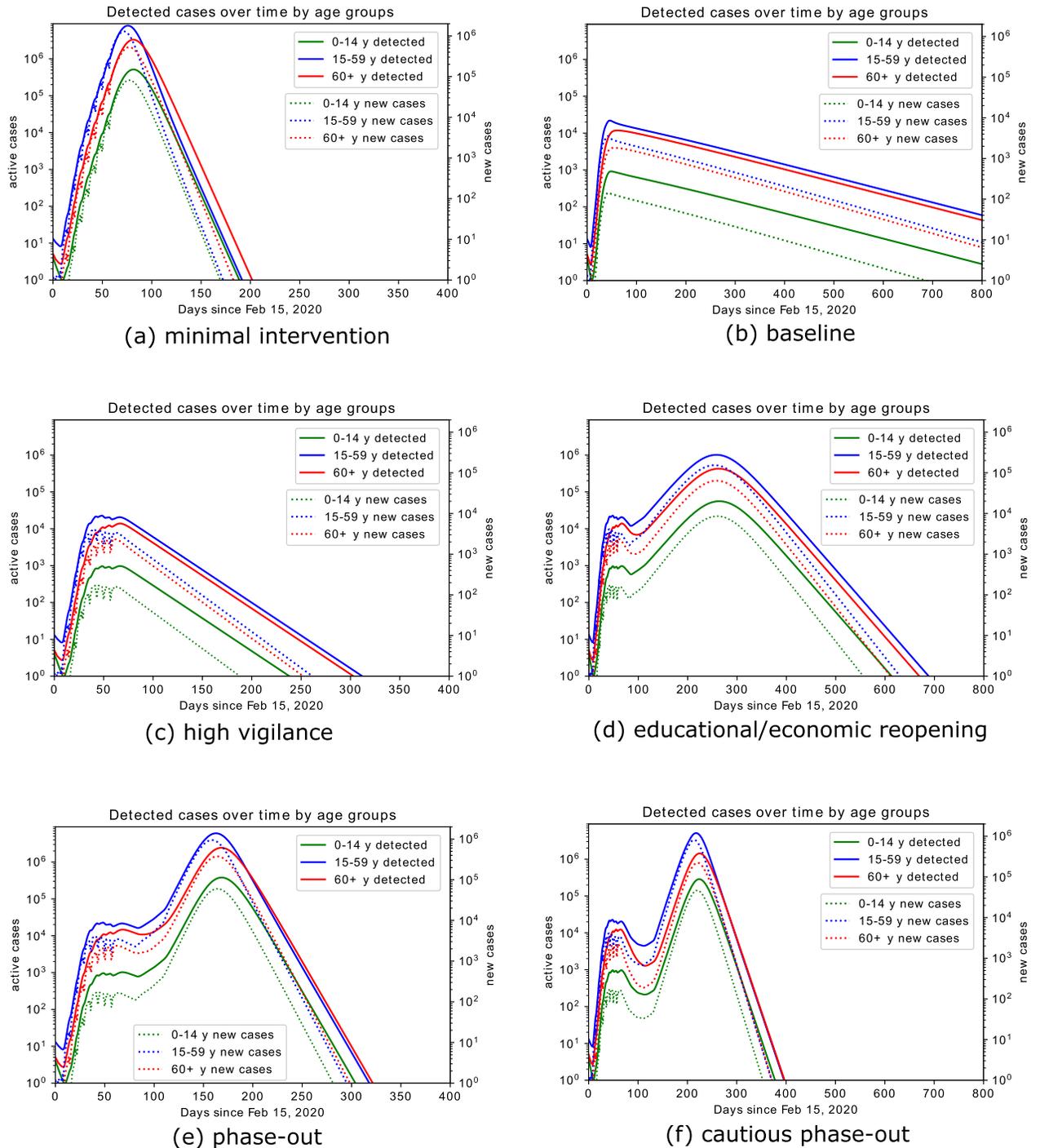
Fig 6. Model prediction for simulated scenarios. (a) *minimal intervention*: increased awareness, quarantine of known or suspected cases, testing of patients with symptoms and contact history; (b) *baseline* scenario: *minimal intervention* scenario increased with school closure, high reduction in economical activities, contact limitation, high testing activity; (c) *high vigilance*: *baseline* scenario enriched by isolation of detected cases, combined with increased testing activity; (d) *educational/economic reopening*: reintroducing in three phases contacts at schools, workplaces, public transportation service; (e) *phase-out*: rollback of all introduced control measures, up to minimal interventions, accompanied by increased testing also of asymptomatic individuals and strict isolation of identified cases; (f) *cautious phase-out*: similar to (e), but with slower rollback to regular activities, accompanied by strongly increased testing also of asymptomatic individuals, strict isolation of identified cases, and reduced contacts with elderly and risk groups. Solutions are shown in logarithmic scale for both new and active cases, in order to make peak heights in different orders of magnitude visible. Oscillations up to day 60 are due to the weekend-effect (Fig 4), which is relaxed for long term projections.

Table 3. Possible measures included in simulated scenarios, starting from April 7. Some of them are by now in effect.

| Label	Policy	Description/Working assumptions
T1+	Significant further increase of testing	Further increase in testing by at least four-fold, not only of individuals with relevant symptoms and/or contact to identified cases but also of asymptomatic persons with any suspicion of being infected; starting on April 15 (day 60).
RB	(Partial) Rollback of measures HO, CS, and/ or LD	Gradually taking back the effects of contact rate reductions according to a possible schedule
	CS:	(Partially) Reopening schools and child care facilities from May 4 (day 79), reduces effect, particularly on children, by approx 70%, reopening universities etc. on May 18 (day 93), reopening clubs and social activities on June 1 (day 107)
	HO:	Restricting HO policy to cases where easily applicable and cautiously reopening restaurants decreases the effect by approx. 60% by May 18 (day 93), final return to full activity by June 1 (day 107)
	LD:	Lifting most severe curfew-like restrictions on May 4 (day 79) reduces effect by 50%, gradual further reductions on May 18 and June 1 to pre-measure level
IC	Stricter isolation of known infectives compared to IC0	Reduces activities rates of identified SARS-CoV-2 infectives by another -50% (compared to IC0), starting on April 10 (day 55), in full effect by April 20 (day 65).
IO	General social distancing (incl. self-separation) of individuals at risk, in particular elderly	Includes particular hygienic care when dealing with risk group individuals, reduction of senior susceptibility by another 40%, starting on April 7 (day 52), in full effect on April 10 (day 55)

Control measures have the fundamental effect to reduce contacts between individuals, hence transmission of the virus from person to person. Recall that in our modeling assumptions, we consider contacts in three different realms: at home, at work and school, and in other locations. Treating different contact categories separately allows for an easier estimate of the effects of contact reducing interventions. School closures are an example which beautifully illustrates the use of separate contact matrices for different realms. A visualization of how control measures reduce the transmission rates in the *baseline* scenario is given in Fig 3(a). In contrast, in rollback scenarios a partial lift of control measures leads to increased transmission rates (Fig 3(b)).

Data

The publicly available dataset provided by the Robert Koch Institute (RKI) [34] was used for this study. Statistical analysis of the dataset was performed (results not shown here) analogously to our previous work [24].

Results

In the core (non-structured) model (1) we estimated how contacts, hence the reproduction number, decreased in the three considered time slots since the beginning of the epidemic fitting the model to reported data [34]. The left panel of Fig 7 shows the fits on top of the data. Each period is described by a different value for the reproduction number: $\mathcal{R} = 6.95$ for the first period, $\mathcal{R} = 3.38$ for the second period, and $\mathcal{R} = 0.97$ for the third period. The value of $\mathcal{R}$ in the early stage may appear rather large when compared to the widely reported values for different countries. On the other hand, reproduction numbers around 10 have been reported
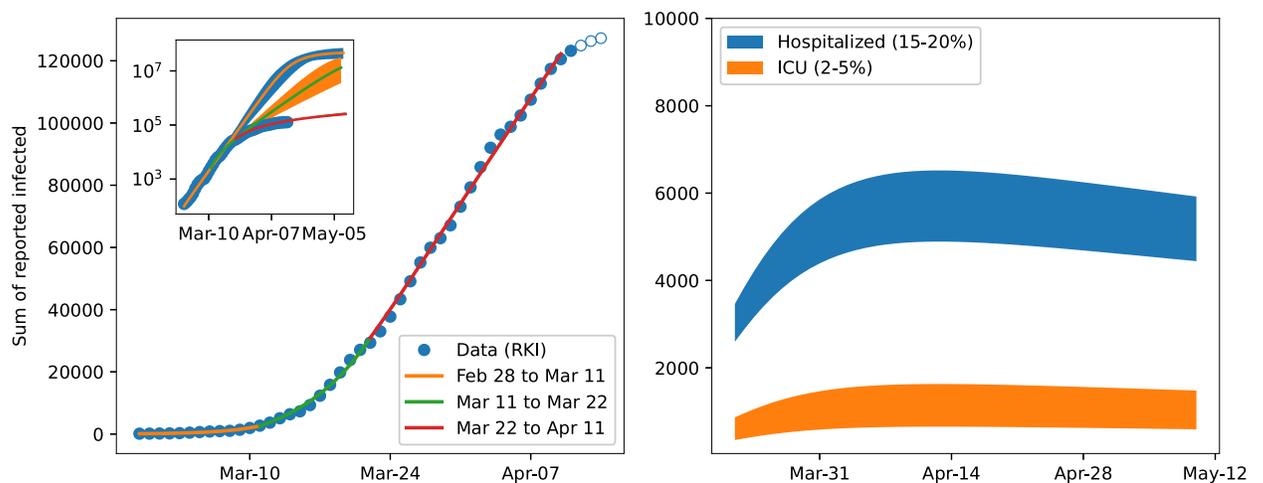
Fig 7. Model fit results and extrapolation. The left panel shows the fits in the 3 periods, clearly indicating the ability of the model to capture the dynamics. The open circles are data points not used for the fit. Due to the inherent latency of the reporting in Germany, these points are not reliable, yet. The inset shows in addition the prediction for the sum of reported cases. Without intervention the number of infections could have surpassed ten million cases by mid April. The bands around the lines indicate the 95% confidence level of the fit based on stochastic variation of the parameters (cf. Fig 2). The right plot shows the number of individuals that require hospitalization either in the low care ward (15-20% of identified cases) indicated by the blue band or ICU (2-5% of identified cases) indicated by the orange colored band.

by other authors as well, e.g., [16] for China or [42] for China and the Italian province of Lombardy. An obvious reason for large reproduction numbers could be super-spreading events, such as large public gatherings. Moreover, decreasing detection ratios upon rising case numbers, due to testing capacities being overwhelmed, may lead to lower apparent values of $\mathcal{R}$ when only data for confirmed cases are considered. The significant reduction of the reproduction number over time clearly indicates the success of the contact reducing measures.

The inset in Fig 7 shows the predictions based on the fits until mid May showing the catastrophic increase of infections that would have resulted if no or only insufficient measures had been taken. The right panel in Fig 7 shows a range for cases requiring hospitalization and intensive care using the latest prediction of the reproduction number. These are computed from the number of detected cases predicted by the model, assuming 15-20% hospitalizations in low care units and 2-5% in intensive care units. Under the current prediction these numbers could be handled by German hospitals. Though this is an indicative estimate only, it clearly shows that the reproduction number consequently decreased from $\mathcal{R} \approx 7$ at the end of February to almost $\mathcal{R} \approx 1$ at the beginning of April.

In the *baseline* scenario for the refined age and stage structured model, where control measures apply continuously over time, we observe that the dominant eigenvalue of the linearization of the system about the disease-free equilibrium (DFE, i.e., a completely susceptible population) decreases in time and eventually crosses zero (Fig 8). This corresponds to the reproduction number dropping below $\mathcal{R} = 1$ (cf. [28]). The seemingly erratic fluctuations which can be observed in Fig 8 are caused by the weekly oscillations in testing that we call the *weekend-effect* and that we illustrate in Fig 4.

Let us now discuss results on the simulated scenarios in detail. In the *minimal intervention* scenario, we assume that no specific measures to reduce contacts between individuals (school closures, interruption of most social and economical activities) nor increased testing activity were undertaken. Under these assumptions, the initial rapid increase of cases would have gone on unabatedly, and the number of infected individuals requiring hospitalization or even
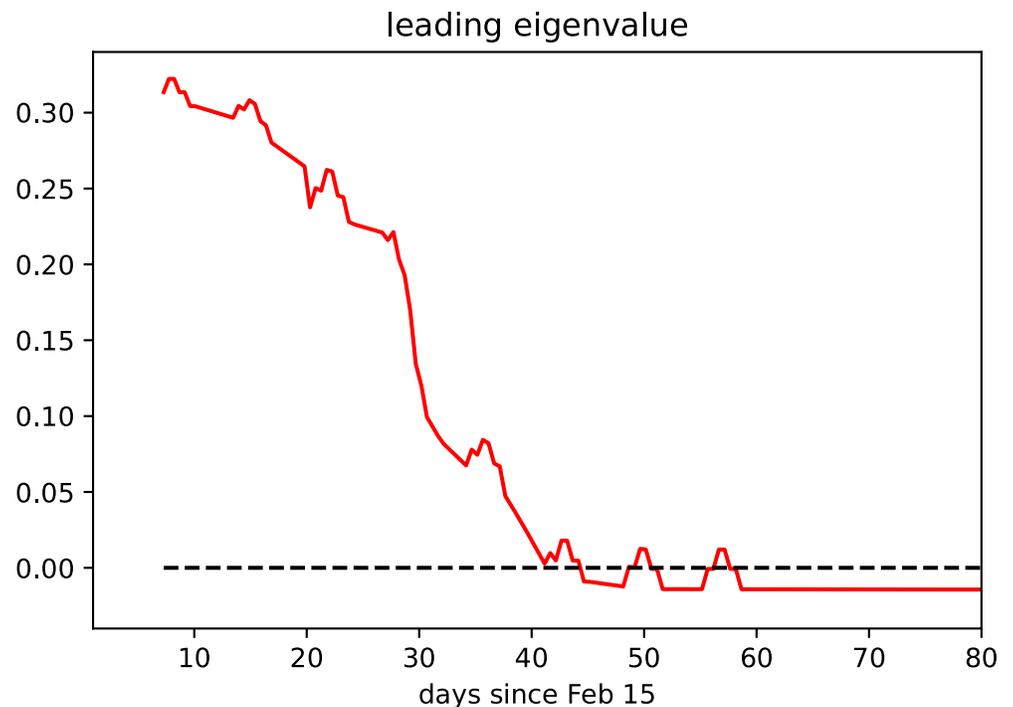
Fig 8. Evolution in time of the dominant eigenvalue of linearization of the age and stage structured model at the disease free equilibrium (DFE). The dominant eigenvalue crossing the zero axis corresponds to the reproduction number crossing the threshold value 1.

intensive care would have reached unmanageable levels within weeks. This scenario is purely counterfactual and is only detailed here to evaluate the effects of the measures adopted so far. Model simulations (Fig 6(a)) show that in this scenario a peak in infections would have been reached at the beginning of May 2020 (day 79 since February 15), with about 12 million active infections on the peak day. Over the course of the infection about 75 million people would have been infected and 1.6 million would have died. Fortunately, taking into account interventions adopted so far, the actual course of the epidemic is less dramatic, and the model simulations predict a significantly better outcome.

Compared to our initial investigations [24], we have adjusted the *baseline scenario* by taking into account the further restrictive measures adopted in the last week of March 2020 (reduction of economic activity, restrictions on meetings in public space, and further increased remote working activity). Said enhanced intervention scenario suggests that the number of active cases peaked with about 33,000 infected individuals at the beginning of April 2020 (Fig 6(b)). This is in accordance with recent modeling studies [23], which suggest that the effective reproduction number $\mathcal{R} \approx 1$ as of April 2, meaning that the disease free equilibrium is on the verge of instability (i.e., the leading relevant eigenvalue of the linearization about the equilibrium with only susceptible individuals is located close to the imaginary axis, cf. Fig 8). Since, at least for small numbers of active infections, the leading eigenvalue lying on the imaginary axis corresponds to the reproduction number $\mathcal{R}$ being close to 1, this makes the predictions for this scenario particularly sensitive to assumptions about the model parameters. The reason is that variations of $\mathcal{R}$ make a striking qualitative difference between further exponentially rising case numbers or a slowdown of the epidemic.

(a)

(b)

(c)

(d)
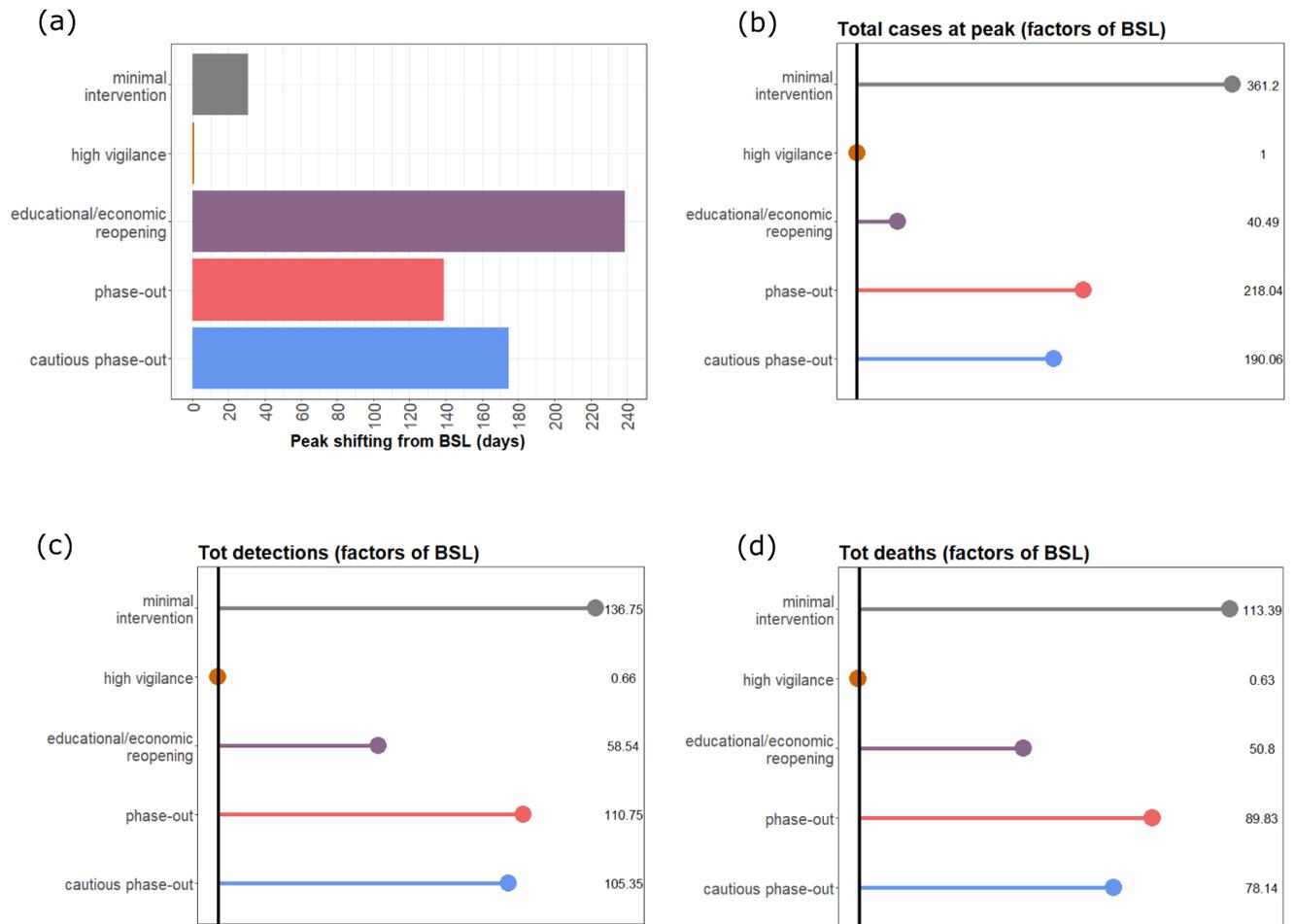
Fig 9. Statistical comparison of model output for the *baseline* (BSL) scenario and the considered possible alternatives. (a) Peak shifting (in days) compared to BSL; (b) Differences in reported cases (factor) at the day of the peak; (c) Differences in total detected cases (factor); (d) Differences in total deaths (factor). For all rollback scenarios, results refer to the second peak of the outbreak.

Taking the latest case numbers (up to April 8) into consideration and assuming that in the weeks to come the testing activity does not take significant slumps on weekends anymore (cf. Fig 4), we project the number of infected individuals, both detected and undetected, to decrease over the coming months. The total number of fatalities would be reduced by more than 90% as compared to the *minimal intervention* scenario (Fig 9(d)), and the capacity of the health care system would not be severely challenged (see also Fig 9(b)).

In the reported data [34], individuals are counted as "recovered" if they are no longer symptomatic and 14 days (supposedly the longest infectious period under normal circumstances) have passed since the positive test. In contrast, the *average* infectious period, as chosen for the simulation, is about $1/\gamma_I = 7$ days, meaning that individuals on average recover seven days after becoming infectious. Therefore, the number of recovered individuals according to the model is higher than the officially recorded figure, and that in turn leads to lower numbers of active reported cases in the model as compared to the official data. Using data up to April 5 to estimate parameters, the model predicts 116,000 cumulative reported cases for April 9, and about 149,000 cases for April 16. Maintaining the *baseline* measures throughout the year (we are assuming no seasonality of the disease) would lead to the eradication of the epidemic.

Simulations of this scenario (Fig 6(b)) suggest about 550,000 infections (about 185,000 thereof asymptomatic), and 14,000 deaths over the course of the epidemic. Enriching the *baseline* scenario with further increased testing activities and even stricter isolation of detected and suspected cases (*high vigilance* scenario, Fig 6(c)) would shorten the time necessary to call the disease eradicated (beginning of September 2020), but widen the peak of active cases. Stricter isolation of confirmed and suspected cases and improved testing activity would further reduce the spread of the virus, forecasting about 361,000 infections (out of which about 108,000 asymptomatic), and 9,000 deaths over the course of the epidemic.

The above scenarios assume that the current restrictions on public life remain in effect over a long period. As this does not seem to be feasible in practice, we consider further scenarios which include an at least partial lifting of the restrictions imposed thus far. In contrast to our previously simulated scenarios [24], we assume a gradual reopening of economical and educational activity (Fig 6(d)). We assume this to start in about three weeks from now (May 4) with reopening schools and childcare facilities as well as many shops, gradually proceeding to reopening universities, restaurants and other economic activities, and finally resuming on-site work and most club activities from June 1 on. Combining this partial rollback with further increased testing activity and isolation of identified cases (*educational/economic reopening* scenario) would lead to a (second) peak in active infections (1.3 million, detected and undetected combined) towards the end of November 2020. If no restrictive measures and interventions were to be (re)introduced, the simulation of the model results in about 32 million total infections and 730,000 deaths over the course of the epidemic, which seems to occur only by the end of the summer 2021 (notice the different time scales for different scenarios in Fig 6) under the assumption that no reliable treatment becomes available by then.

The last two scenarios that we present here suggest that a complete, though gradual, rollback of all introduced control measures would lead to a second peak towards the end of August 2020 (Fig 6(e), *phase-out* scenario), or end of September 2020 in case of slower reintroduction of regular activities (Fig 6(f), *cautious phase-out*). In both cases, the second peak would be anticipated (Fig 9(a)) and the number of infections at this peak would be way larger than in the *educational/economic reopening* scenario (7.2 million and 6.3 million active infections in the *phase-out* and *cautious phase-out* scenarios, respectively; cf. Fig 9(b)).

Discussion

In this work we proposed a mathematical model for predicting the evolution in time of detected COVID-19 infections in Germany taking into account the age distribution of cases. Distinguishing between people in different age groups allows the model to better characterize contacts between individuals (e.g., child-child contacts being typically different than senior-adult contacts, cf. [21]) and to fine-tune the effect of intervention measures on contacts reduction (cf. Fig 3).

Given the limited knowledge about the novel virus' properties and the unprecedented control measures, there are significant uncertainties regarding the precise effects of single measures on effective contact rates. For example, while school closures can be clearly modeled as reducing child-child contacts in the school domain (cf. Methods), the effectiveness of this specific intervention in curbing an epidemic can vary dramatically [43, 44], depending on, e.g., the pathogen and its interaction with the immune system of children. Moreover, it is hard to predict how the impact of control measures might wear off as the population grows tired and attention fades, or, rather the measures become even more effective as habituation makes following the guidelines easier. Our results indicate that the current measures lead to a significant reduction in the reproduction number, $\mathcal{R}$, which is approximately 1 in the second week of

April ([Fig 8]). This matches with estimations of the RKI [34] and findings by other groups which are also currently studying the situation in Germany [23, 45]. This naturally results in a significant uncertainty of the projection since small deviations of the parameters can make the difference between further growing active case numbers or slowly declining numbers.

We parametrized the core model using, if possible, known or previously estimated parameter values, in particular those concerning the evolution of the disease (latency time, duration of infection, death rates), or plausible assumptions (e.g., for the relation between infectivity of detected and undetected infectives) as explained in the Methods section. Uncertainties of parameters that could not be inferred from the literature or well identified from the data were investigated via stochastic sampling and Sobol analysis. Such uncertainties could be reduced by integrating further data in the model, e.g., estimated contacts or estimates for the number of unreported cases.

For performing the simulations of the age-structured model some of the parameters were taken from literature, some were inferred from the fit of the core model, and others were fixed by means of plausible assumptions (cf. Methods) that provided good agreement with the data reported in the early phase of the outbreak. The parametrization of the structured model could be improved in the future, in particular assuming that better quality data would be available. Analogously to what was done for the core model uncertainties could be systematically analyzed for this model as well.

In our *baseline* scenario we assumed the current reporting ratio in Germany to be rather high (over 50% as of mid April). Assuming a lower reporting ratio yields more undetected cases and therefore leads to a higher estimate of the reproduction number. Analogously, dropping the weekend effect increases the number of active cases at the peak. On the other hand, lower reporting ratios in early April provide a larger margin for improvement by enhanced testing as in our *high vigilance* scenario. While reporting ratios are notoriously hard to estimate in early phases of an epidemic, and have a potentially enormous impact on the predictions made by any model [46], several models completely neglect the presence of undetected cases [33, 45]. The aggressiveness of the virus and hence the mortality among all affected individuals (whether diagnosed or not) is another unknown, but different assumptions about this parameter can be expected to have similar impacts on all the scenarios discussed here. It may be assumed that earlier detection (as in all our scenarios with enhanced testing) and hence earlier and better care for high risk patients may result in a mortality even below the one observed in Germany so far. The limited capacity of the health care system, in particular of intensive care units, was not yet directly considered as a parameter of our refined model. However, the predicted number of infected individuals at its peak can be used as a proxy for the expected demand for health care resources at the height of the epidemic. Assuming that a fixed proportion of infected individuals will require intensive care, the maximal number of infectives for a given scenario directly indicates the maximal load on the health care system for this scenario.

For an initial study, considering the limited available data, we decided to apply the model to the entire German population, treating it as spatially homogeneous. Compartmental models assuming geographically homogeneous regions were also proposed by other authors for the spread of COVID-19 in Germany [17, 45] or other countries [47, 48]. Of course, this is an important simplifying assumption, as the population density and disease incidence vary from region to region, and from rural to urban areas. Features of the COVID-19 outbreak that emerged from reported data [34] clearly indicate the heterogeneous distribution of cases with respect to the age groups, which led us to highlight this property instead of geographic inhomogeneities. While the transmission dynamics would essentially remain as presented in this manuscript, the deterministic (based on ordinary differential equations) approach that we adopted here could be too coarse when considering much smaller populations distributed over

more geographic patches. When the number of infectious individuals is small, or demographic/environmental variability could significantly impact the epidemic outcome, stochastic models might be better suited [49].

We have conducted simulations covering one year and more starting from the beginning of the epidemic. Some scenarios predict high peaks in active cases and alarmingly high numbers of deaths far into the future. However, these scenarios should only be understood as predictions for the future if no appropriate measures were taken to contain the epidemic. In particular, postponing the date of the peak sufficiently far into the future would provide time for the development of a vaccine or an effective treatment.

Author Contributions

Conceptualization: Maria Vittoria Barbarossa.

Data curation: Hridya Vinod Varma, Noemi Castelletti.

Formal analysis: Maria Vittoria Barbarossa, Jan Fuhrmann, Jan H. Meinke, Stefan Krieg, Hridya Vinod Varma.

Funding acquisition: Thomas Lippert.

Investigation: Maria Vittoria Barbarossa, Jan Fuhrmann, Jan H. Meinke, Stefan Krieg, Hridya Vinod Varma.

Methodology: Maria Vittoria Barbarossa.

Software: Jan Fuhrmann, Jan H. Meinke, Stefan Krieg, Hridya Vinod Varma.

Supervision: Maria Vittoria Barbarossa.

Validation: Jan Fuhrmann, Hridya Vinod Varma.

Visualization: Maria Vittoria Barbarossa, Jan Fuhrmann, Jan H. Meinke, Noemi Castelletti.

Writing – original draft: Maria Vittoria Barbarossa, Jan Fuhrmann, Jan H. Meinke, Stefan Krieg.

Writing – review & editing: Maria Vittoria Barbarossa, Jan Fuhrmann, Jan H. Meinke, Stefan Krieg, Thomas Lippert.

References

1. ECDC *COVID-19* European Center for Disease Prevention and Control 2020. Available online: https://www.ecdc.europa.eu/en/covid-19-pandemic (last accessed on April 2, 2020)

2. WHO *Coronavirus disease 2019 (COVID-19) Situation Report—73* World Health Organisation 2020. Available online: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf?sfvrsn=5ae25bc7_2

3. Johns Hopkins University *Coronavirus Resouce Center.* Johns Hopkins Center for Systems Science and Engineering 2020. Available online: https://coronavirus.jhu.edu/#covid-19-basics (last accessed on April 2, 2020)

4. RKI *Coronavirus SARS-CoV-2—dedicated homepage of the* Robert Koch-Institute, Berlin, Germany 2020. Available online: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/nCoV_node.html (last accessed on April 2, 2020)

5. WHO Team *Modes of transmission of virus causing COVID-19:implications for IPC precaution recommendations* World Health Organisation—Department of Communication 2020. Available online: WHO Reference Number: WHO/2019-nCoV/Sci_Brief/Transmission_modes/2020.2 (last accessed on April 2, 2020)

6. RKI *SARS-CoV-2 Steckbrief zur Coronavirus-Krankheit-2019 (COVID-19)* Robert Koch-Institute, Berlin, Germany 2020. Available online: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html (last accessed on April 4, 2020)

7. EpiCentro *Epidemia COVID-19—Aggiornamento nazionale del 2 Aprile 2020*. Istituto Superiore di Sanitá, Rom, Italy 2020. Available online: https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_2-aprile-2020.pdf

8. EpiCentro *Caratteristiche dei pazienti deceduti positivi a COVID-19 in Italia—Dati del 2 Aprile 2020*. Istituto Superiore di Sanitá, Rom, Italy 2020. Available online: https://www.epicentro.iss.it/coronavirus/bollettino/Report-COVID-2019_2_aprile.pdf

9. WHO Team *Statement—Older people are at highest risk from COVID-19, but all must act to prevent community spread* World Health Organisation—Department of Communication 2020. Available online: http://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/statements/statement-older-people-are-at-highest-risk-from-covid-19,-but-all-must-act-to-prevent-community-spread (last accessed on April 19, 2020)

10. Rothe C and Schunk M and Sothmann P and Bretzel G and Froeschl G and Wallrauch C et al. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany N. Engl. J. Med. 382:970–971. (2020) https://doi.org/10.1056/NEJMc2001468 PMID: 32003551

11. He X and Lau EHY and Wu P and Deng X and Wang J and Hao X et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat Med 26 672–675 (2020). https://doi.org/10.1038/s41591-020-0869-5 PMID: 32296168

12. Zhang B and Zhou X and Zhu C and Feng F and Qiu Y and Feng J et al. *Immune phenotyping based on neutrophil-to-lymphocyte ratio and IgG predicts disease severity and outcome for patients with COVID-19*. MedRxiv.

13. Li Q and Guan X and Wu P and Wang X and Zhou L and Tong Y et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. N. Engl. J. Med. (2020) https://doi.org/10.1056/NEJMoa2001316

14. Riou J and Althaus CL Pattern of early human-to-human transmission of Wuhan 2019-nCoV. Eurosurveillance 25(4) 2020 https://doi.org/10.2807/1560-7917.ES.2020.25.4.2000058 PMID: 32019669

15. Shen M and Peng Z and Xiao Y and Zhang L *Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China. bioRχiv*2020.

16. Tang B and Wang X and Li Q and Bragazzi NL and Tang S and Xiao Y et al. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions J. Clin. Med. 9, 462 (2020)

17. Maier BF and Brockmann D Effective containment explains sub-exponential growth in confirmed cases of recent COVID-19outbreak in Mainland China Science 368 (6492) 742–746 (2020) https://doi.org/10.1126/science.abb4557 PMID: 32269067

18. Boldog P and Tekeli T and Vizi Z and Dénes A and Bartha FA and Röst G. Risk Assessment of Novel Coronavirus COVID-19 Outbreaks Outside China J. Clin. Med. 9(2), 571 (2020)

19. Jung S and Akhmetzhanov AR and Hayashi K and Linton NM and Yang Y and Yuan B et al. Real-time estimation of the risk of death from novel coronavirus (COVID-19) infection: inference using exported cases. J. Clin. Med. 9(2) 523 (2020)

20. Imperial College COVID-19 Response Team. *Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand*. March 16, 2020.

21. Prem K and Liu Y and Russell TW and Kucharski AJ and Eggo RM and Davie N 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 5(5) e261–e270 (2020). https://doi.org/10.1016/S2468-2667(20)30073-6 PMID: 32220655

22. an der Heiden M and Buchholz U *Modellierung von Beispielszenarien der SARS-CoV-2-Ausbreitung und Schwere in Deutschland* (2020)

23. Khailaie S and Mitra T and Bandyopadhyay A and Schips M and Mascheroni P and Vanella P et al. *Estimate of the development of the epidemic reproduction number Rt from Coronavirus SARS-CoV-2 case data and implications for political measures based on prognostics medRχiv*

24. Barbarossa MV and Fuhrmann J and Heidecke J and Vinod Varma H and Castelletti N and Meinke JH et al. *A first study on the impact of current and future control measures on the spread of COVID-19 in Germany*—March 28, 2020 medRxiv (2020)

25. Wiestler OD and Marquardt W and Heinz D and Meyer-Hermann M *Stellungnahme der Helmholtz-Initiative "Systemische Epidemiologische Analyse der COVID-19-Epidemie"*, April 13, 2020 https://www.helmholtz.de/fileadmin/user_upload/01_forschung/Helmholtz-COVID-19-Papier_02.pdf

26. RKI *Erfassung der SARS-CoV-2-Testzahlen in Deutschland*. Epidemiologisches Bulletin 15—Robert Koch-Institute, Berlin, Germany 2020. Available online: https://www.rki.de/DE/Content/Infekt/EpidBull/Archiv/2020/Ausgaben/15_20.pdf?__blob=publicationFile

27. Brauer F and Castillo-Chavez C and Feng Z *Mathematical Models in Epidemiology*, Springer (2019)

28. Diekmann O and Heesterbeek JAP and Roberts MG (2010) The construction of next-generation matrices for compartmental epidemic models. J R Soc Interface 7:873–885. https://doi.org/10.1098/rsif.2009.0386

29. Branch MA and Coleman TF and Li Y A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems SIAM Journal on Scientific Computing, Vol. 21(1) 1–23 (1999) https://doi.org/10.1137/S1064827595289108

30. Virtanen P and Gommers R and Oliphant TE and Haberland M and Reddy T and Cournapeau D et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods. 17, 261–272 (2020). https://doi.org/10.1038/s41592-019-0686-2 PMID: 32015543

31. Akaike HA new look at the statistical model identification. IEEE transactions on automatic control, 19 (6), 716–723 (1974) https://doi.org/10.1109/TAC.1974.1100705

32. Hurvich CM and Tsai CL Regression and time series model selection in small samples. Biometrika. 76, 297–307 (1989). https://doi.org/10.1093/biomet/76.2.297

33. an der Heiden M and Hamouda O Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland—Nowcasting (2020)

34. RKI *Table with the current COVID-19 infections per day (time series) from the Robert Koch-Institute.* Robert Koch-Institute, Berlin, Germany 2020. Available online: https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/dd4580c810204019a7b8eb3e0b329dd6_0/ (last accessed on April 4, 2020)

35. Lauer SA and Grantz KH and Bi Q and Jones FK and Zheng Q and Meredith HR et al. *The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application.* Annals of Internal Medicine, March 2020

36. Herman J and Usher W *SALib: An open-source Python library for Sensitivity Analysis* (2017)

37. Sobol' IMGlobal sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates2001 Mathematics and Computers in Simulation 55(1-3):271–280.

38. Saltelli A and Ratto M and Andres T and Campolongo F and Cariboni J and Gatelli D et al. *Global sensitivity analysis: the primer*, John Wiley & Sons (2008)

39. Wu J and Dhingra R and Gambhir M and Remais JV Sensitivity analysis of infectious disease models: methods, advances and their application. (2013) J. Royal Soc. Interface. https://doi.org/10.1098/rsif.2012.1018

40. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med 5(3) e74 (2008). https://doi.org/10.1371/journal.pmed.0050074 PMID: 18366252

41. Zimmermann P, Curtis N, Coronavirus Infections in Children Including COVID-19: An Overview of the Epidemiology, Clinical Features, Diagnosis, Treatment and Prevention Options in Children. The Pediatric Infectious Disease Journal (2020) 39(5):355–68 https://doi.org/10.1097/INF.0000000000002660 PMID: 32310621

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

43. Cowling BJ, Ali ST, Ng TWY, Tsang TK, Li JCM, Fong MW, Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. Lancet Public Health 5:e279–88 (2020) https://doi.org/10.1016/S2468-2667(20)30090-6 PMID: 32311320

44. Jackson C, Mangtani P, Hawker J, Olowokure B, Vynnycky E, The Effects of School Closures on Influenza Outbreaks and Pandemics: Systematic Review of Simulation Studies. PLoS ONE 9(5): e97297 (2013). https://doi.org/10.1371/journal.pone.0097297

45. Dehning J and Zierenberg J and Spitzner FP and Wibral M and Neto JP and Wilczek M et al. Inferring COVID-19 spreading rates and potential change points for case number forecasts Science 369 (6500), eabb9789 https://doi.org/10.1126/science.abb9789 PMID: 32414780

46. Fuhrmann J and Barbarossa MV *The significance of the detection ratio for predictions on the outcome of an epidemic—a message from mathematical modelers* Archives of Public Health 78, Article number 63 (2020)

47. Giordano G and Blanchini F and Bruno R and Colaneri P and Di Filippo A and Di Matteo A et al. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy Nature Medicine 26, 855–860(2020) https://doi.org/10.1038/s41591-020-0883-7 PMID: 32322102

48. Anastassopoulou C and Russo L and Tsakris A and Siettos C *Data-based analysis*, *modelling and forecasting of the COVID-19 outbreak* PLOS ONE (2020)

49. Allen LJA primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. Infectious Disease Modelling, 2(2), 128–142 (2017) https://doi.org/10.1016/j.idm.2017.03.001 PMID: 29928733