The impact of digital contact tracing on the SARS-CoV-2 pandemic - a comprehensive modelling study

Tina R Pollmann[1], Julia Pollmann[4], Christoph Wiesinger[1], Christian Haack[1], Lolian Shtembari[5], Andrea Turcati[1], Birgit Neumair[1], Stephan Meighen-Berger[1], Giovanni Zattera[1], Matthias Neumair[2], Uljana Apel[2], Augustine Okolie[2], Johannes Müller[2,3], Stefan Schönert[1], and Elisa Resconi[1]

[1]Physics Department, Technical University of Munich, 85748 Garching, Germany
[2]Department of Mathematics, Technical University of Munich, 85748 Garching, Germany
[3]Institute for Computational Biology, Helmholtz Center Munich, 85764 Neuherberg, Germany
[4]Department of Medical Oncology, University Hospital Heidelberg, National Center for Tumor Diseases (NCT) Heidelberg, 69120 Heidelberg, Germany
[5]Max Planck Institute for Physics, 80805 Munich, Germany

September 13, 2020

Abstract

Contact tracing is one of several strategies employed in many countries to curb the spread of SARS-CoV-2. Digital contact tracing (DCT) uses tools such as cell-phone applications to improve tracing speed and reach. We model the impact of DCT on the spread of the virus for a large epidemiological parameter space consistent with current literature on SARS-CoV-2. We also model DCT in combination with random testing (RT) and social distancing (SD).

Modelling is done with two independently developed individual-based (stochastic) models that use the Monte Carlo technique, benchmarked against each other and against two types of deterministic models.

For current best estimates of the number of asymptomatic SARS-CoV-2 carriers (approximately 40%), their contagiousness (similar to that of symptomatic carriers), the reproductive number before interventions (R_0 at least 3) we find that DCT must be combined with other interventions such as SD and/or RT to push the reproductive number below one. At least 60% of the population would have to use the DCT system for its effect to become significant. On its own, DCT cannot bring the reproductive number below 1 unless nearly the entire population uses the DCT system

and follows quarantining and testing protocols strictly. For lower uptake of the DCT system, DCT still reduces the number of people that become infected.

When DCT is deployed in a population with an ongoing outbreak where $\mathcal{O}(0.1\%)$ of the population have already been infected, the gains of the DCT intervention come at the cost of requiring up to 15% of the population to be quarantined (in response to being traced) on average each day for the duration of the epidemic, even when there is sufficient testing capability to test every traced person.

Contents

1 Introduction

Tracing and isolation of people who were in contact with an infectious person (contact tracing) can be used to control the spread of communicable diseases [1, 2]. In the traditional understanding of contact tracing (CT), public health employees interview known carriers (index cases) of the disease and then track down people who had the type of close contact with the index case necessary to transmit the disease. Contacts are then diagnosed and isolated. This implementation is only suited for infections that spread relatively slowly, and where cases can be easily diagnosed [3]. SARS-CoV-2, with its unspecific symptoms, high number of asymptomatic carriers, and incubation times as short as a day, does not fit this mold. Recent studies indicate that an outbreak of SARS-CoV-2 could be controlled using fast and efficient digital contact tracing (DCT) [4, 5]. DCT systems using cellphone applications based on Bluetooth proximity measurements are currently being developed and/or deployed in many countries [6, 7, 8, 9, 10, 11, 12, 13, 14, 15]. Predicting the effect of DCT on an outbreak is challenging, especially since the epidemiological parameters that describe the outbreak dynamics are currently still uncertain. Almost no practical experience for DCT is available. Partly automated CT systems were in use during the Ebola outbreak 2014-2016 [16] and it turned out that the technical difficulties that come with that approach must not be underestimated.

Simulation studies focusing on classical CT for COVID-19 consistently indicate that around 70% of the contacts need to be traced, and that the tracing delay has to be as short as 1 day [5, 17]. For these reasons, [17] doubt that COVID-19 can be controlled by traditional CT in practice. Ferretti et al. [4] however point out that DCT could significantly reduce tracing delays, so that the outbreak could be controlled for tracing probabilities much smaller than 70%. Meanwhile, several other simulation studies indicate that a high tracing probability and a combination of fast CT and testing is required to control SARS-CoV-2 [16]. However, it also became clear that not only the reduction of the reproduction number, or the final size of the epidemic needs attention, but also the number of persons that go to quarantine are of importance [18]: A naive application of DCT leads to a situation that resembles a lock-down as a large fraction of the population is quarantined. As we also find in the present study, an appropriate choice of the tracing and testing protocol is central.

We developed individual-based models with the Monte Carlo (MC) simulation technique, flanked and cross-checked by deterministic models, to evaluate the dynamics of a COVID-19 outbreak under different intervention protocols, focusing on DCT and DCT combined with random testing and social distancing. We determine not just the immediate effective reproductive number R_e, but also the daily R_e, the number of healthy people in quarantine, and the number of people infected, for up to a year of continuous interventions. The sensitivity of all outcomes to the reported ranges of values of the epidemiological parameters is studied in detail.

The goal of this paper is to study quantitatively under which conditions and to which degree DCT combined with fast testing and social distancing can replace rigorous shelter-in-place policies for keeping the effective reproduction rate $R_e \leq 1$.

2 Model inputs

Table 1 presents an overview of all model input parameters and the values considered for them. They will be discussed in detail in the following subsections.

Many model inputs that are properly described by a distribution rather than just a mean value are modelled using a shifted Gamma distribution, defined as

$$G(x;\mu,\gamma,\beta) = \frac{\left(\frac{x-\mu}{\beta}\right)^{\gamma-1}\exp(-\frac{x-\mu}{\beta})}{\beta\Gamma(\gamma)} \tag{1}$$

where $x \geq \mu$. As this distribution is flexible, we also use a discretized version ($n \in \mathbb{N}_0$)

$$\hat{G}(n;\gamma,\beta) = G(n;0,\gamma,\beta)\left(\sum_{i=0}^{\infty} G(i;0,\gamma,\beta)\right)^{-1}. \tag{2}$$

2.1 Social contact structure

Our deterministic models describe contacts between individuals in a population by the usual mass action law. The stochastic models allow for a more detailed investigation. Particularly, two different strategies are considered: For the first, in the following referred to as *homogeneous population*, we choose the contacts for each person randomly out of the entire population. The probability for a person to have n unique contacts close enough to transmit a respiratory virus on a given day is taken from the empirical distributions reported in [19]. These are well described by

$$P_{\text{social}}(n) = \hat{G}(n;\gamma=2,\beta=n_c/2)\,, \tag{3}$$

where n_c is the mean number of contacts per day. We also consider a more realistic contact pattern, in the following referred to as *social graph population*. The population is described by a social graph, were each individual is represented by a node and contacts are represented by edges. Each individual is given a fixed set of contact persons for the entire simulation. We employ a modified version of the *Lancichinetti–Fortunato–Radicchi* benchmark graph (LFR) [20] as shown in [21] (therein referred to as LFR-BA). In this model, the population is divided into communities with sizes distributed according to a power-law distribution. Node edges are constructed according to the linear preferential attachment model [22] under the constraint that an average fraction of $(1-\mu)$ of the edges of each node connect nodes within the same community. A graph constructed in this manner results in a power-law probability distribution with index $a = 2$ for the node degree n [22]:

$$P_{\text{social;SG}}(n;a,n_{min}) \sim a \cdot n_{min}^{a} n^{-(a+1)} \tag{4}$$

where n_{min} is the minimum node degree. Here, the mean number of contacts per day is given by $n_c = \frac{a \cdot n_{min}}{a-1}$. We assume that each edge is active once per day, so that the node degree corresponds to the number of contacts per day.

2.2 Epidemiological parameters

Fig. 1 schematically shows the probability distributions for symptom onset, transmission, and true positive test results. The parameters are explained in the following paragraphs.

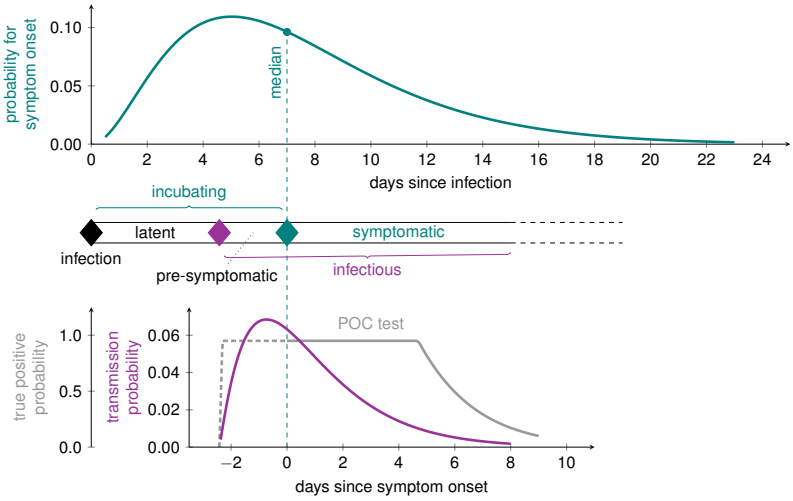


Figure 1: Course of the disease with probability distribution for the incubation period T_{Inc} (top) [23] and the fully correlated probability density function for the contagiousness T_{con} (bottom) [24]. The dotted vertical lines corresponds to the median of T_{Inc}. The probability for a true–positive point-of-care (POC) test is displayed on the bottom left (grey line). The diamonds correspond to exposure (contact, black), end of latency/begin of contagious period (magenta) and symptom onset (teal).

Incubation period $\mathbf{P(T_{Inc})}$. The distribution of incubation periods is taken as

$$P(T_{\text{Inc}}) = G(T_{\text{Inc}}, \mu, \gamma, \beta) \tag{5}$$

(Fig. 1 upper curve) with shape parameters chosen to match the curve reported in [23], which has a median and mean incubation period of 7 days and 7.44 days, respectively, with a range from 0-23 days. The authors included a relatively large group of patients (n=587) with a wide age (0-90 years) and symptom (asymptomatic to severe) range. Other studies, albeit with smaller number of patients and with a bias towards more severe symptoms, have reported lower medians [25, 26]. Therefore, in addition to the parameters matching [23] (printed in black in Table 1) we also model a curve with shorter median and mean incubation period of 3.6 days and 4 days (printed in gray in Table 1).

Latent period T_{lat}. The latent period for SARS-CoV-2 is shorter than the incubation period, meaning pre-symptomatic transmission can occur [1][27]. The latent period is difficult to determine empirically, as it requires exact information about the time of exposure and contagiousness. As, to our knowledge, no reliable, large-scale studies have been published on the latent period of SARS-CoV-2 so far, we use the measured contagiousness relative to the incubation time as an auxiliary mean to infer latency,

$$T_{\text{lat}} = \max\{T_{\text{Inc}} - 2.5, 0\}$$

where the value of 2.5 days comes from the transmission probability curve discussed in the next paragraph.

The transmission probability curve $P_{\mathbf{trans}}(\tau)$. The transmission probability is the probability that a contagious person infects someone they have contact with. This probability is often given as an average "infectivity per day", β_i, even though it changes significantly as a function of the time since infection τ. The infectivity was measured as a function of the time since onset of symptoms by He et al. [24]. They find that carriers become contagious approximately 2.5 days before the onset of symptoms, and that approximately 44% of transmissions occur during this pre-symptomatic phase.

We take as the contagious period T_{con} the time from the end of the latent period until 99% of the cumulative transmission probability is reached; a person is considered to be recovered afterwards. Therewith, the transmission probability as a function of time since infection is given by a scaled and truncated Gamma distribution (Fig. 1 lower curve). After infection, an individual has a (random) latent period T_{lat}, during which the transmission probability is zero. Afterward, for $\tau = T_{\text{lat}} + t$ (and $t \geq 0$), we have

$$P_{\text{trans}}(T_{\text{lat}} + t) = T_{\text{con}}\, \beta_i\, G(t; \mu, \gamma, \beta)\ \chi(t < T_{\text{con}}). \tag{6}$$

where χ is 1 if the condition in the argument is met, and 0 otherwise.

The shape parameters shown in black in Table 1 are our defaults, taken to match the curve in [24]. Pre-symptomatic infectivity is strongly debated. Therefore, we model a second shape where only 18% of the transmission occurs during the pre-symptomatic phase (shape parameter values printed in gray).

The course of the disease for asymptomatic carriers is the same as that for symptomatic carriers as shown in Figure 1, and the incubation time is then the time when symptoms would have started.

Fraction of asymptomatic cases $1-\alpha$. The proportion of asymptomatic carriers $(1-\alpha)$ described in the literature ranges from $\sim 4\%$ [28] to $\sim 40\%$[29] of all cases. Initial reports for $(1-\alpha)$ derived from testing of specific cohorts (cruise ship, returning travellers) ranged from 17-31% [27, 30, 31], however larger studies suggest even higher numbers. Analysis of the mass screening of the full population of the municipality of Vo', Italy [29] report that 41.1% of the confirmed SARS-CoV-2 infections were asymptomatic (as defined by the absence of fever and/or cough). Ferretti et al. [4] analyzed 40 selected transmission pairs and also derived a value of 40% for the proportion of asymptomatic infected individuals. We model several different values for α to cover the reported ranges.

Reduced asymptomatic transmission probability η_{as}. While initial studies assume that asymptotic cases are less contagious [32, 4], newer reports indicate that the viral load of asymptomatic cases is similar to symptomatic cases, which suggests similar contagiousness [33, 34, 35, 36, 37, 38]. We nevertheless introduce the parameter $\eta_{as} \in [0, 1]$, which scales the transmission probability for pre- and asymptomatic cases, and vary that parameter to explore its effects. We note that when we model with $\eta_{as} < 1$, we apply the scaling to both pre- and asymptomatic phases; since there most likely is no difference in viral load, if asymptomatic transmission is suppressed, this is likely due to circumstantial factors, like lack of coughing, which apply to the pre-symptomatic phase as well.

The basic reproductive number R_0. If the contagiousness is independent of the symptom status, i.e. η_{as} =1, the reproductive number is given by the number of contacts during the contagious period, $n_c T_{\text{con}}$, times the average

probability to transmit the infection in one contact, β_i. However, we need to distinguish between symptomatic, pre-, and a-symptomatic cases in order to allow for $\eta_{as} < 1$. We find

$$R_0 = n_c \left[\eta_{as} \int_{T_{\text{lat}}}^{T_{\text{Inc}}} P_{\text{trans}}(t)dt + ((1-\alpha)\,\eta_{as} + \alpha) \int_{T_{\text{Inc}}}^{T_{\text{lat}}+T_{\text{con}}} P_{\text{trans}}(t)dt\right] \quad (7)$$

Note that – though P_{trans} is a random function – the random part of the function is a pure translational offset (the latency period), s.t. the integral is deterministic.

R_0 is expected to be different in different populations because n_c is different by factors of up to 4 just within the populations of different European countries. Values of R_0 ranging from 1.4 to 6.5 have been reported [43], though it is not always clear whether or not the reported R_0 is for the case where symptomatic individuals are quarantined. Furthermore, R_0 is usually not corrected for the contact rate in the population where it is studied. We consider R_0 the reproductive number without any form of interventions. We take R_0 of 3 (approximately the median reported in [43]) as our default, but also run the models for R_0 of 2 and 4.

Following [44], we introduce the random reproduction number for an individual $R_{0,i}$. Then, of course, R_0 is the expectation value over the $R_{0,i}$. For simplicity (and since two of our models are not based on continuous time but on discrete time/days), we state the time-discrete formula for $P(R_{0,i} = n)$. Assume that an individual did infect n persons. These n persons can be arbitrarily distributed over the contagious period. In a slight abuse of notation, let T_c denote the number of days that a person is contagious for. Furthermore, let $C_n = \{\vec{n} \in \mathbb{N}_0^{T_c} : \sum_{i=1}^{T_c} \vec{n}_i = n\}$ be all possible ways to distribute the n infectees to T_c days. Then, for a homogeneous population,

$$P(R_{0,i} = n) \;=\; \frac{1}{|C_n|} \sum_{\vec{n} \in C_n} \left[\sum_{i=1}^{T_{\text{con}}} \text{Binomial}(\vec{n}_i, P_{\text{trans}}(T_{\text{lat}} + i), m) \circledast P_{\text{social}}(m)\right], \quad (8)$$

where the convolution $\circledast$ is over the parameter m. This distribution is shown for some parameter combinations in Fig. 2.

2.3 Intervention protocols

The interventions considered here are 1) DCT, 2) quarantining, 3) testing, and 4) social distancing. Reported symptomatic cases are quarantined starting right at the beginning of the epidemic. The remaining interventions are turned on once a fraction f_i of the population has become exposed.

(1) DCT. We assume that a fraction p_{app} of the population uses the DCT system and that we can trace all contacts between users of this system with time delay T_{delay}. In case that both infector and infectee have a DCT device, the probability for successful tracing is η_{DCT}, while tracing always fails if either infector or infectee do not use the system. η_{DCT} accounts for situations where cell phones run out of battery, are not with the owner at all times, Bluetooth is turned off, or where an alert is ignored. The DCTS system identifies contacts within the past ΔT_{trace} days.

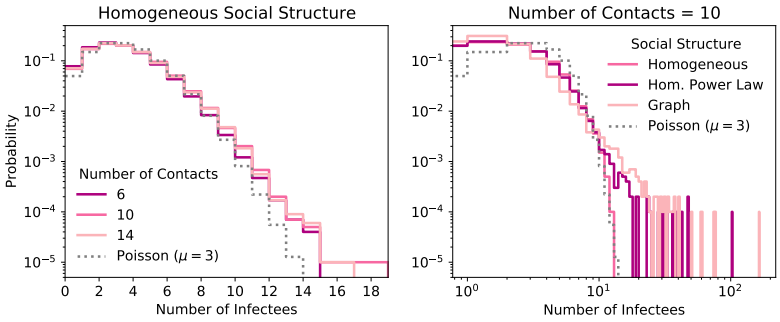


Figure 2: The distribution of the number of people a carrier infects (Eq. (8)) for 3 combinations of n_c (the mean number of contacts per day) and β_i (the average transmission probability per day) that result in R_0 =3 (see Tab. 1) (left) and for a fixed n_c but different social structures (right). Default values for the infection probability curve are used, $\eta_{as} = 1$, and no interventions are applied. Compared to a Poisson distribution with mean of 3, the distribution is over-dispersed.

In the literature, the overall tracing probability across the population is often taken as $p_{trace} = p_{app}^2\,\eta_{DCT}$. We note that this formula is not correct but becomes approximately right if $p_{\rm app}^2\,\eta_{DCT}$ is small (see supplemental materials).

To become an index case for tracing, a person must be reported. We assume that from the group of symptomatic carriers, a fraction f_m sees a doctor to get tested with a reliable laboratory test and is then reported. A fraction $(1-\alpha)$ of cases will go unreported because they do not exhibit symptoms, unless they get tested due to being traced. A fraction $\alpha(1-f_m)$ of symptomatic cases will go unreported due to lack of access to medical tests. In the case where η_{as} =1, α and f_m are degenerate.

First order tracing refers to a protocol where contacts of an index case are traced. DCT also allows immediate tracing of contacts-of-contacts. We refer to this as second order tracing. If the ΔT_{trace} is big enough, DCT will identify the infector. Second order tracing then can trace not just the people infected by an index case, but also the people who were infected by the same infector as the index case.

A DCT system will identify all contacts, regardless of their infection status. Since many models assume perfect accuracy in identifying only contacts that became infected, we run all parameters both with (closer to reality) and without (to be comparable to other models) tracing of the uninfected contacts.

All traced people immediately go into quarantine. This is necessary to suppress the pre- and asymptomatic transmission rates.

(2) Quarantining. Quarantining refers to any intervention that reduces the transmission probability significantly; this includes self-quarantine at home as well as being hospitalized. We assume that all reported symptomatic patients are immediately quarantined, regardless of any other interventions. This already

reduces the reproductive number to

$$R_{e,Q} = n_c \left[\eta_{as} \int_{T_{\text{lat}}}^{T_{\text{Inc}}} P_{\text{trans}}(t)dt + ((1-\alpha)\ \eta_{as} + \alpha\ (1-f_m)) \int_{T_{\text{Inc}}}^{T_{\text{lat}}+T_{\text{con}}} P_{\text{trans}}(t)dt \right] \quad (9)$$

Fig. 3 shows $R_{e,Q}$ for combinations of α, f_m, and η_{as}. In Fig. 3 (top), η_{as} =1, so only the product of α and f_m is relevant, and results are shown for R_0 =3 and for R_0 =2. In Fig. 3 (bottom), $\eta_{as} < 1$, so both R_0 and $R_{e,Q}$ depend on α and on f_m. For each combination of α and f_m, the transmission probability β_i was adjusted to obtain R_0 =3.

In response to being reported or being traced, people are quarantined by default for 14 days. Symptomatic cases may leave quarantine 8 days after the symptoms start. Uninfected contacts can leave quarantine early following a testing protocol.

(3) Testing. We consider two types of tests. A reliable laboratory test for symptomatic carriers seeing a medical professional, and a fast point of care (POC) test that can be performed at home or at mobile testing stations. In either case, carriers only test positive while they have a high enough viral load. We assume the tests have $p^{\text{true positive}} = 0.0$ while the carrier is in the latent period, that is up to approximately 2.5 days before symptom onset. The viral load rises quickly after the end of the latent period. We further assume that the laboratory test then has a true positive rate of 100% until the carrier has recovered. The POC test on the other hand has $p^{\text{max}}_{\text{true positive}} = 0.9$ until approximately 5 days after symptom onset. After this time, the true positive rate falls at the same rate as the transmission probability curve; the true positive rate as a function of days since symptom onset is shown in Fig. 1 (bottom gray curve).

All people who are traced must be tested for two reasons: a) A positive test result is the only way for asymptomatic individuals to become index cases for tracing, and index cases are needed for tracing to be effective, and b) so uninfected traced people can be released from quarantine. Keeping all traced people in quarantine for the full quarantining duration means that a large fraction of the uninfected population may end up quarantined on any given day of the outbreak. We use the following release protocol: All traced people go into quarantine and get tested with a POC test. Regardless of the test result, everyone stays in quarantine, because the person may still be in the latent period. Those who tested negative on the first day are re-tested $\delta T_{\text{re-test}}$ days later. If both tests were negative, the person may leave quarantine, but is tested again after another $\delta T_{\text{re-test}}$ days in case they were still in the latent period when the second test was done.

In addition to testing in response to being traced, we simulate the option of randomly testing a fraction f_{RT} of the population each day. This is done with a testing protocol assumed to have a negligible number of false positives.

(4) Social Distancing. Social distancing includes both a reduction of the total number of contacts per day n_c to $n_c \cdot f_{\text{SD}}$, and limiting the maximum number of contacts per day. In the absence of second-order effects, and if the upper limit does not change the distribution mean significantly, this reduces the reproductive number to

$$R_{e,SD} = f_{\text{SD}}\ R \quad (10)$$

where R is the reproductive number without social distancing.

Table 1: Key model input parameters and settings. The models are evaluated for all possible combinations of all parameter values shown in black. Parameter values in grey are used only with select other parameter combinations.

Parameter/Setting	Values	Notes/References
Disease and population		
Size of population	10k, 100k , 1M	
Population structure	uniform , social graph	
Transmission prob.(β_i)	1.89, 2.87, 3.74, 2.0, 4.67 [%]	
Contact rate (n_c)	10, 14, 6	
R0	2.0, 3.0, 4.0	Calculated from β_i and n_c.
Trans. prob. curve (μ, γ, β)	(-2.42, 2.08, 1.56) (-1, 1.9, 1.4)	[24, 36, 39, 33]
Incubation time curve (μ, γ, β)	(0, 3.06, 2.44) , (0, 3.06, 1.3)	[23, 40, 26]
Fraction symptomatic (α)	0.4, 0.6, 0.8, 0.95	[28, 27, 29, 30, 31, 41, 42]
Asymptomatic trans. scaling (η_{as})	0.1, 0.5, 0.8, 1.0	[33, 34, 35, 4, 37]
Interventions		
Interventions start (f_i)	0.00, 0.004, 0.04	The fraction of the population exposed when interventions start.
Quarantine duration	14 days	
Tracing		
Reported from symptoms (f_m)	0.5, 0.75, 1.0	Fraction of symptomatic carriers that see a doctor.
Trace back (ΔT_{trace})	7 , 14 [days]	Time window for CT.
App coverage (p_{app})	0.0, 0.6, 0.75, 0.9, 1.0	Fraction of the population that uses the DCTS.
Tracing efficiency (η_{DCT})	0.5, 0.75, 1.0	Chance that a contact between two users of the DCTS is successfully traced.
Tracing order	1, 2	
Trace uninfected contacts	True, False	
Tracing delay (T_{delay})	0 , 2, 4, 6 [days]	
Social distancing		
SD upper limit, factor	(60, 1.0), (12, 0.6), (16, 0.8)	Maximum number of contacts per day, factor by which mean number of contacts is scaled.
Testing		
Random testing rate (f_{RT})	0.00,0.01, 0.05, 0.1, 0.15, 0.2 [1/day]	Fraction of population tested per day.
Days to test result	0 [days]	
False positive rate	0.00, 0.01	
Re-test interval ($\delta T_{\text{re-test}}$)	5 [days]	Traced people that test negative on tracing day are tested again after this time interval.
True positive rate (p_m)	0.9 , 0.0 (no testing)	For the POC test on days with peak test efficiency

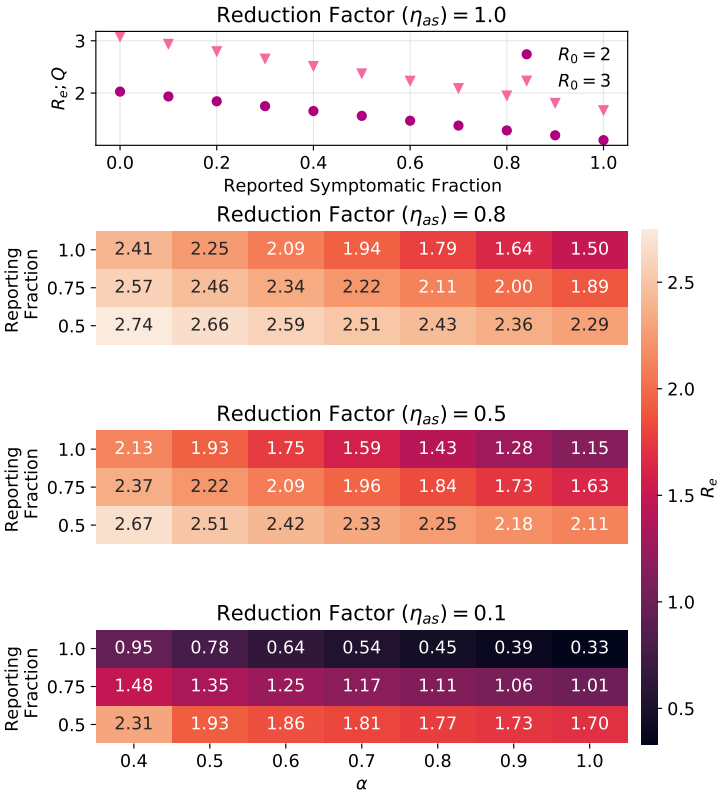


Figure 3: The effective reproductive number reached just from quarantining reported symptomatic carriers, $R_{e;Q}$, is shown for four different values of η_{as} (asymptomatic infectivity scaling) as calculated from Eq. (9). Top panel: In the case of η_{as} =1, $R_{e;Q}$ depends only on the product of α (symptomatic fraction) and f_m (fraction reported and tested) and is shown for two values of R_0. Lower three panels: For each combination of η_{as} and α, the infection probability was adjusted to obtain $R_0 = 3$.

3 Models

Epidemiological modelling is a well established scientific discipline and different approaches, including contact tracing, are described in the rich literature [45, 46, 47, 48]. Epidemiological models that account for CT date back to the 1980s [49]. The main challenge to modelling a CT system is the individual-based character of CT, and the handling of the resulting stochastic dependencies between individuals. Individual-based simulation models [50] readily describe this process. For the scope of this paper, we developed two deterministic and two individual-based models. This redundant approach serves to cross-validate the results.

3.1 Deterministic models

The early phase of an outbreak can be quantitatively described with compartmental models based on ordinary differential equations (ODE) [51] or with age-since-infection models [52].

The deterministic models used here bridge the different scales utilizing the mathematical analysis of the underlying, microscopic stochastic branching process with contact tracing. The effect of contact tracing on the removal rates are determined. These effective removal rates are then used in the deterministic models. Our first deterministic, compartmental model explicitly predicts the status (exposed/infectious) for a newly infected person, when he/she will eventually be traced. Eventually traced and never-traced individuals go to different compartments. In that, the (exponentially distributed) waiting times can be readily adapted. Particularly, the model is close to standard SEIR-models (see Fig. 4), and is feasible to analytical analysis (Appendix A). In contrast, the second model, based on age since infection, does not explicitly formulate an exposed and an infectious period. The basic assumption is that the state of an individual is a function of his/her age of infection, that is the time that has passed since he or she became infected. The structure is less pronounced, but it is possible to use transition rates that rate more realistic (Appendix B). At the present time, the analytical treatment of the interdependence of contact tracing and correlations between infected individuals at the plateau phase of an epidemic is not well understood. Therefore, both models focus on the onset of the epidemics, where the reduction of susceptible by quarantine does not play a central role. The main outcome of these two models are the doubling time T_2 and effective reproductive number R_{eff} for interventions starting on the first day of the epidemic, though both models are able to predict in a heuristic way the total course of the epidemics.

3.2 Individual-Based Models

We developed two independent individual-based models (IBMs), which use the Monte Carlo (MC) technique to simulate social interactions, the progression of the viral disease, and interventions, at the level of individual people. The MC simulations proceed through the outbreak in steps of one day. Each day of the outbreak, every infected person not in quarantine meets a number of other people randomly drawn either from $P_{\text{social}}(n)$ (for a homogeneous population structure) or from the person's social graph. The probability to infect each

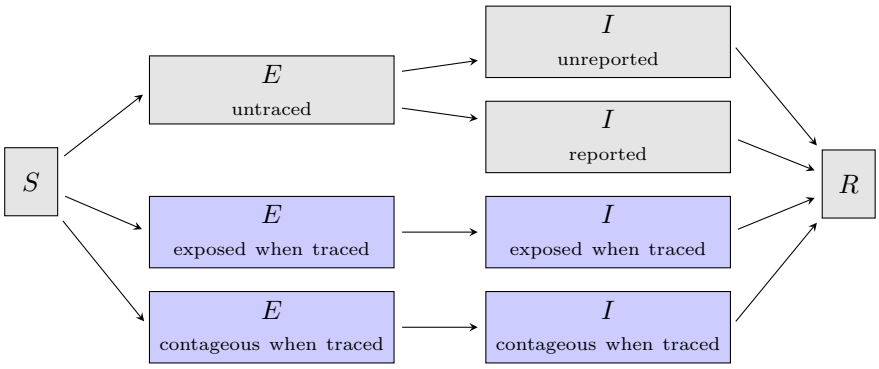


Figure 4: Simplified structure of the compartmental model. The model is based on a SEIR-type model from Ref. [53] and distinguishes between untraced (gray) and traced (blue) individuals. A detailed description of the model and its variables can be found in Appendix A.

contact is given by $P_{\text{trans}}(\tau)$. When a contact becomes infected, the incubation time is drawn from $P(T_{\text{Inc}})$. The intervention protocols are implemented as described in Sect. 2.

Figure 5 shows a chain of infections from one of the simulation runs. Each box represents a person, and arrows between boxes represent infections and tracing.

In this example, P936 was exposed to the virus on day 147 of the simulated epidemic and had a latent period of 5 days, but never developed symptoms (light blue background) or tested positive and was therefore never reported (R–). He or she infected 3 others – P576 on day 152, P747 on day 154, and P277 on day 155. All three infectees developed symptoms (purple background). P576 saw a doctor on day 155, tested positive, and become reported. This triggered tracing of his infector, P936, and of the person he or she infected, P392. Tracing to P392 failed because this person does not use the app. Since P392 also does not develop symptoms, he or she never gets reported or quarantined and infects 3 others. The backward trace from P576 to P936 puts P936 in quarantine on day 155 and thus prevents him or her from infecting more people after this time. P936 does not tests positive (dashed outline of the box indicates the person was traced but never reported), so never becomes an index case him- or herself. However, since second order tracing is active in this simulation, the 'siblings' of P576 are identified. P747 is put in quarantine before he or she can infect anyone else, and tests positive the same day. This makes him or her an index case, so that the common infector, P936, is traced again. P277 does not use the app, therefore the trace fails. However, P277 happens to not meet many people on the first 2 days of infectiousness, then developed symptoms, saw a doctor, and was quarantined.

In this example, the chain of infections was stopped at P747 through second order tracing, the chain was stopped at P277 due to luck, but the chain could not be interrupted at P576 because the person he or she infected did not use the tracing app.

For a given set of input parameters, that is for a specific scenario, each run of the MC simulation represents one possible course of the epidemic. To find

the most likely outcome for a scenario, the simulation is run 50 to 1000 times and the outcomes are averaged. As an example, Figs. 6 and 22 show the course of the epidemic for 20 MC runs each. The stochastic nature of the processes involved creates a spread in outcomes. Especially near the beginning of the epidemic where only few people are infected, statistical fluctuations cause large differences in the outbreak dynamics.

The R_e shown for each day is given as the average number of people infected by everyone who recovered on that day. After the interventions are turned on, R_e begins to decrease and in the absence of non-linear effects reaches a plateau after approximately 10 to 14 days. In runs where more than a few percent of the population has been exposed at that time, R_e declines naturally due to an increasing chance that contact persons are already infected or recovered, and therefore cannot be infected again. When reporting the R_e for a simulation run, $R_e(t)$ is averaged in the time span of 10 days to 28 days after interventions start, or from 10 days to the day more than 50% of the population has been exposed, whichever period is shorter. This time window is a compromise between being far enough away from the start of interventions for the effect of the interventions to fully manifest, and not getting to close to the region where R_e changes naturally. The R_e reported for a scenario is the average R_e over all the simulation runs for that scenario.

We consider the following outcomes:

- The fraction of the population exposed after one year of continuous interventions. The one year is counted from the day interventions start.
- The fraction of the population sick on the day when most people are sick.
- The average fraction of the population in quarantine each day over one year of continuous interventions.
- The fraction of the population in quarantine on the day when most people are in quarantine.
- The effective reproductive number after interventions.
- The fraction of simulations that did not generate an outbreak, where an outbreak is defined as at least 0.04% of the population becoming exposed in runs where interventions do not start on day 0, and is defined as at least 50 people becoming exposed in runs where interventions start on day 0.

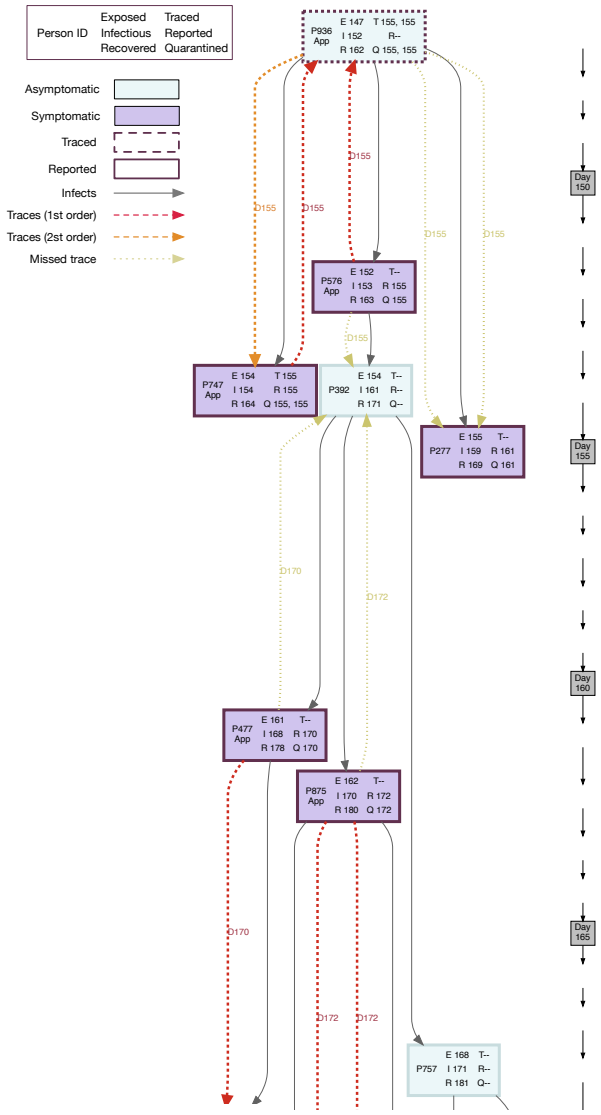


Figure 5: Spread and containment of SARS-CoV-2 using tracing and quarantine interventions in a Monte Carlo simulation including 1st and 2nd order tracing. The downward arrows at the right indicate the time axis, with days in the simulated epidemic given in the grey boxes. Shown are the first 3 generations of infectees originating from simulated person P936. The input settings for the simulation were: R_0 =2.0, ΔT_{trace} =7 days, p_{app} =0.6, η_{DCT} =0.9, α =0.6. Each box represents a person. The leftmost column in each box gives the person ID and whether or not they use the DCT app. The middle column indicates the days when the person was exposed, became infectious, and recovered. The rightmost column indicates if and when the person was traced, reported, and quarantined.

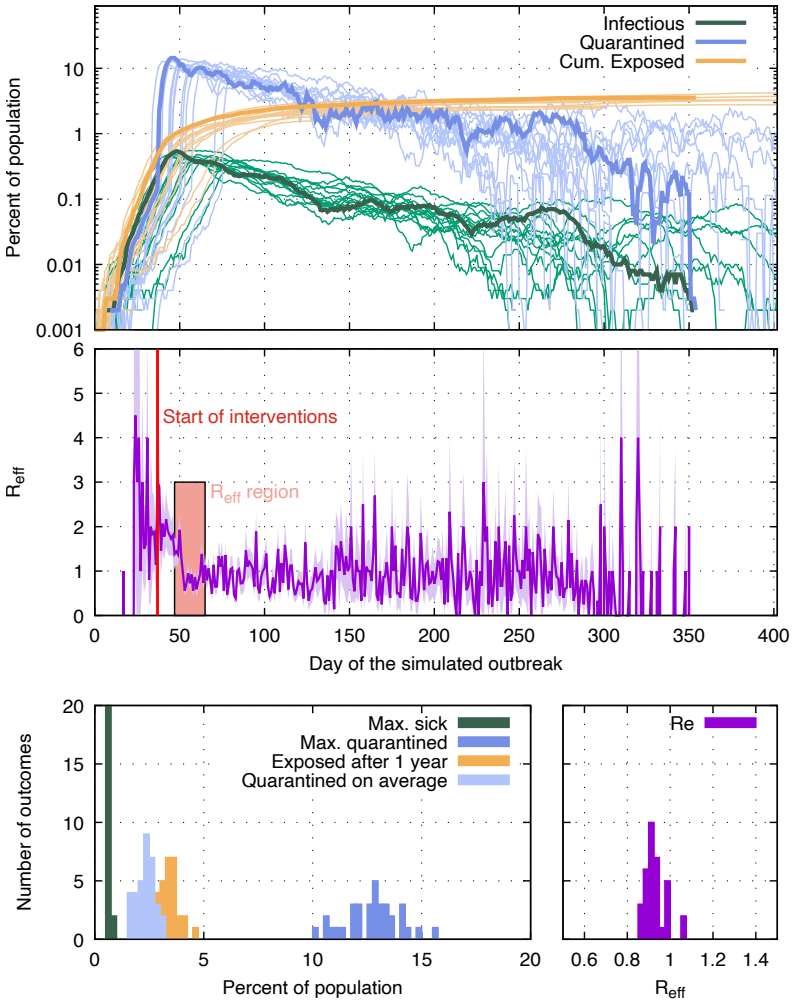


Figure 6: Stochastic variation in outbreak dynamics. The results are from 50 runs of the MC simulation; each run has the same input parameters (p_{app} =0.9, $\alpha \cdot f_m$ =0.95, η_{as} =1, trace uninfected = true). Top: The fraction of infectious (green), quarantined (blue) and cumulative exposed (yellow) people for each day of the simulated outbreak. Curves for only 20 out of the 50 runs are shown to improve legibility. Outcomes from one selected run are drawn as bold lines. Middle: R_e each day is shown for the MC run drawn in bold in the top plot. The red vertical line indicates the time when 0.4% of the population have been infected, which is when interventions (other than quarantining of reported symptomatic cases, which is enabled from the beginning) are turned on. To measure their effectiveness, R_e is averaged over 18 days (red area), starting 10 days after interventions commence. Bottom: Outcomes from the 50 MC runs, such as the maximum fraction of the population quarantined, are histogrammed to show the statistical variation more clearly.

4 Results

We highlight three outcomes for select scenarios and as function of the app coverage. Unless stated otherwise, the values printed in black in Tab. 1 are used for those parameters not explicitly varied in the figures or stated in the figure captions. The full set of outcomes for all scenarios is shown in the supplementary materials.

4.1 The effect of instantaneous contact tracing on an ongoing epidemic

Fig. 7 and Fig. 8 show three outcomes each for the four simulated symptom/reporting fractions and for R_0 =3 (Fig. 7) and R_0 =2 (Fig. 8). Results are in each case shown for the realistic case where tracing identifies contacts regardless of their infection status, and for the case where only infected contacts are traced. The latter is included so that results can be compared to other models, and because the difference in the number of quarantined people between the two cases indicates how many healthy people are quarantined when uninfected contacts are also traced. $R_0 = 2$ is likely too optimistic, however the results are also valid in the situation where R was lowered to R=2 by other interventions, such as mask wearing, before tracing and quarantining starts.

The R_e shown in the top panels of Fig. 7 and Fig. 8 should be compared to Fig. 3. For example for $R_0 = 3$ and $\alpha \cdot f_m = 0.6$, just quarantining reported symptomatic cases yields $\mathrm{R}_{e;q} = 2.2$, so DCT only lowers R_e by another 0.5 (if tracing is independent of infection status), or 0.3 (if tracing finds only infected contacts).

Assuming that α is about 60% in European populations and that not everyone who has symptoms sees a doctor or is tested, the region between $\alpha \cdot f_m = 0.4$ to 0.6 is likely relevant for Europe. If the reports of higher α in Asian countries are due to true differences in symptom fraction rather than to differences in study methods, the higher $\alpha \cdot f_m$ values simulated should be more relevant to Asian countries.

We will use $R_0 = 3$, $p_{\mathrm{app}} = 0.6$, $\alpha \cdot f_m = 0.6$, and $\eta_{DCT} = 1$ (points outlined in red in Fig. 7) as defaults.

In Fig. 9, the lightest points correspond to these defaults. The other colors indicate what happens when the tracing efficiency is reduced. For the lower app coverages, the results are barely sensitive to η_{DCT} because DCT is not very effective to begin with.

In the realistic case where traced uninfected contacts are quarantined until two test results are negative (see Sect.2.3), as many as 15% of the population are in quarantine on each day of the simulated outbreak, most of them healthy. Without a POC test to release healthy contacts, this number rises to 25%. At the peak of the outbreak, approximately 30% of the population is quarantined and half of those quarantined are actually sick.

People are not available to be infected while in quarantine, so the mean number of contacts per day, and with it the effective reproductive number, goes down and fewer people become exposed. The number of people quarantined rises with higher app coverage (because more people are traced in that case) and with a higher number of people exposed (because there are more index cases). A higher app coverage eventually leads to fewer exposed people though. Hence for

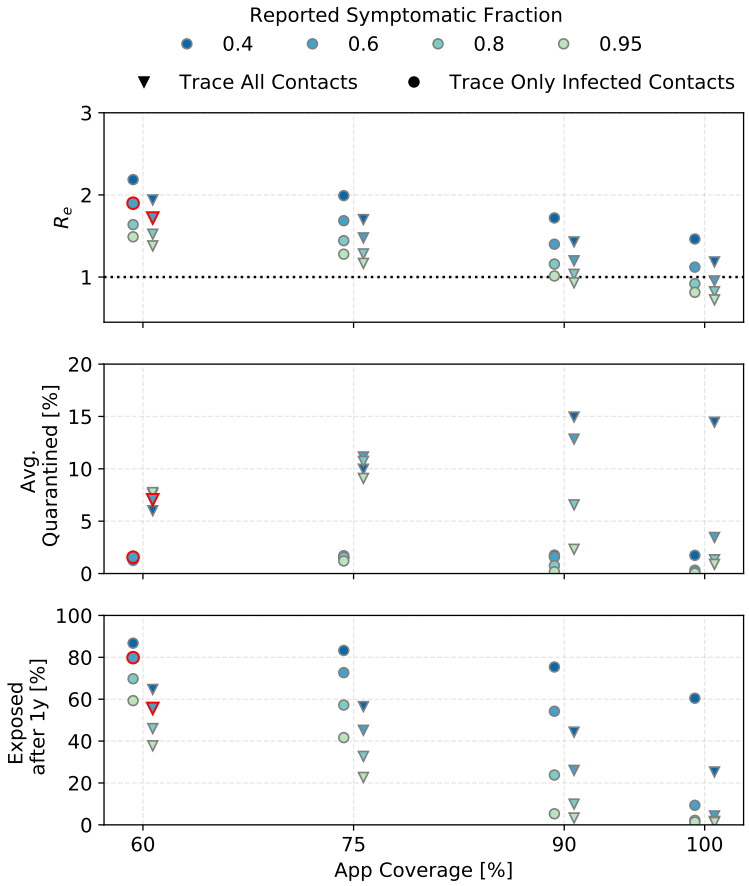


Figure 7: The effect of $\alpha \cdot f_m$ (tested symptomatic fraction) for different app coverages on the reproductive number after interventions (DCTS and quarantine) (top), the average of daily quarantined people as a percentage of the total population (middle) and the percentage of the population exposed after 1 year of continuous interventions (bottom). The points for 60% app coverage and 60% symptomatic fraction, outlined in red, will be studied further. This is for $R_0 = 3$ and $\eta_{DCT} = 1$.

a given reported symptomatic fraction, the number of people quarantined rises until an app coverage of approximately 75% (for lower reported symptomatic fractions) or 90% (for higher reported symptomatic fractions) and then falls sharply.

Contact tracing cannot reduce R below 1 in any of the simulations presented here except for $\alpha \cdot f_m \geq 0.8$ and $p_{\text{app}} \geq 0.9$ (if $R_0 = 3$) or $p_{\text{app}} \geq 0.7$ (if R when tracing and quarantining is started is 2), and perfect tracing probability.

We note that in some cases, the fraction of the population exposed after 1 year is higher than the herd immunity level. The herd immunity level is defined as the fraction of the population that must be immune for the increase

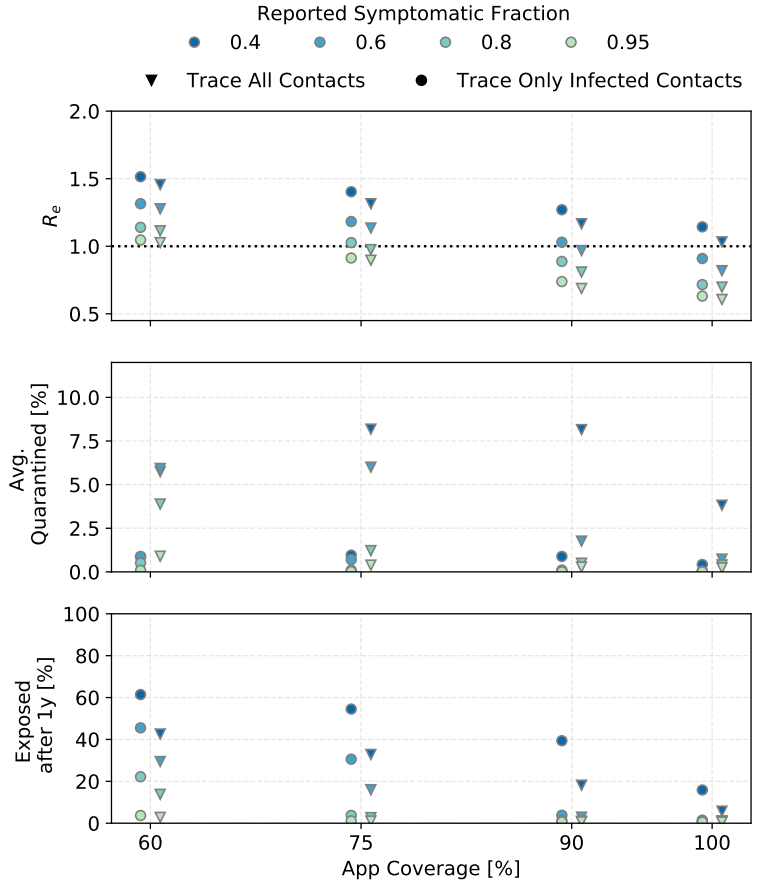


Figure 8: The same as Fig. 7 but for starting conditiosn where R=2. R=2 could be achieved by interventions other than DCT.

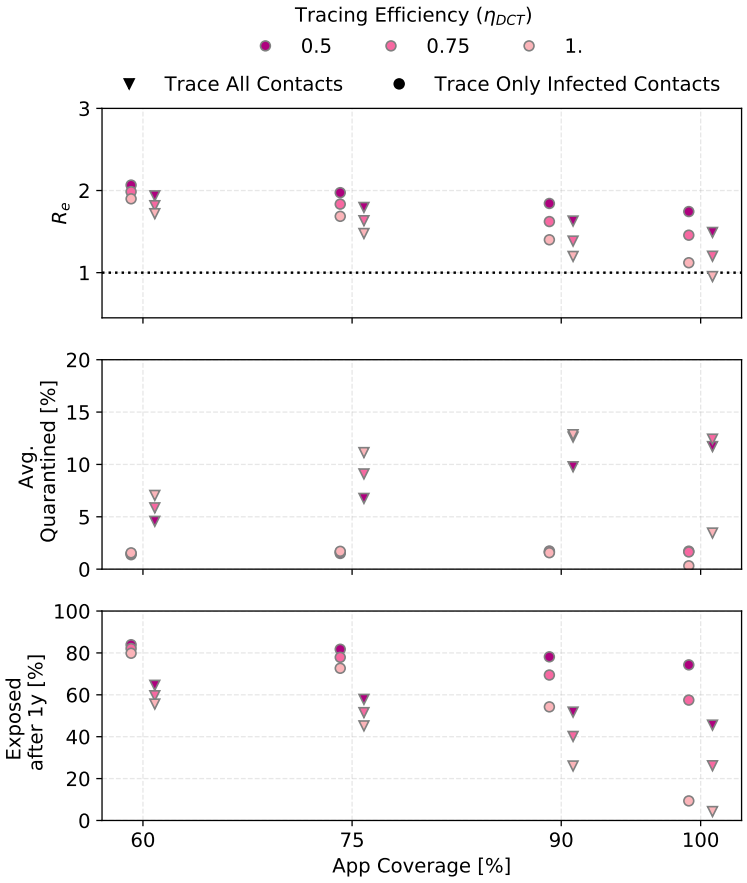


Figure 9: The effect of the tracing probability η_{DCT} for different app coverages (R_0 =3, $\alpha \cdot f_m$ =0.6).

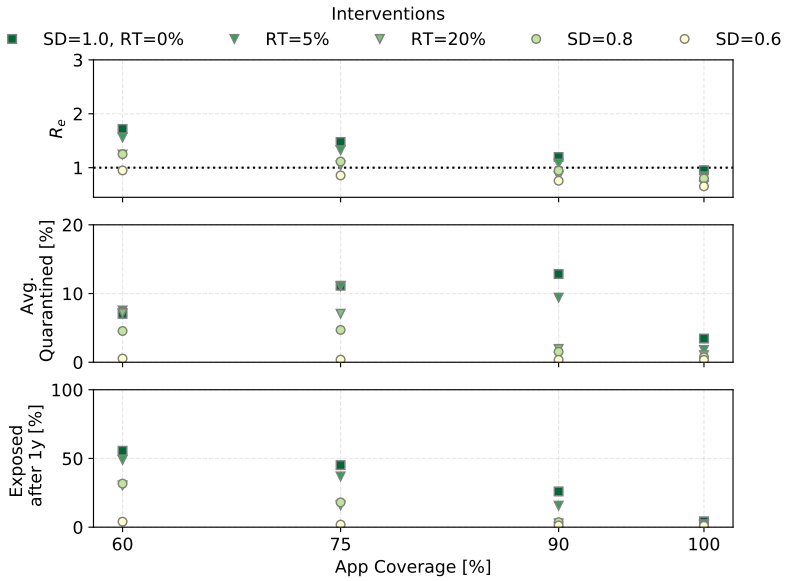


Figure 10: The effect of social distancing (SD) and random testing (RT) for different app coverages in combination with CT (R_0 =3, η_{DCT} =1, $\alpha \cdot f_m$ =0.6).

in new infections to not be able to grow exponentially, that is for R_e to become 1. In an ongoing epidemic, many people are infectious when this point is reached, and the number of exposed people continues to rise until enough people are immune for $R_e = 0$, therefore the curve overshoots herd immunity level.

4.2 Contact tracing in combination with random testing and social distancing

To control the epidemic, R must be reduced by additional measures. We simulated the effect of random testing (RT) and social distancing (SD). Fig. 10 shows the outcomes for our standard scenario with the addition of RT of 5% and 20% of the population per day, and social distancing bringing n_c to 0.8 and 0.6 of its original value. The reduction in contact rate is always connected to an upper limit in the number of contacts as shown in Table. 1.

Random testing even at 20% of the population per day in combination with contact tracing can only achieve $R_e \leq 1$ for $p_{\text{app}} \geq 0.75$. It does however bring R_e close enough to 1 to significantly reduce the fraction of the population that becomes exposed, even for lower app coverages.

Social distancing reliably reduces the reproductive number. Social distancing to just 80% of the contact rate does as well as randomly testing 20% of the population each day. Reducing the contact rate to 60% pushes R_e below 1 for 60% app coverage.

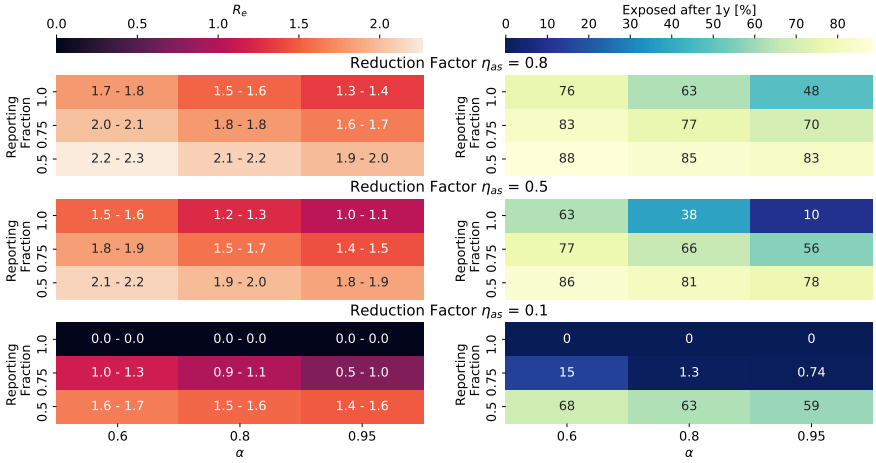


Figure 11: Outcomes when the contagiousness of a- and pre-symptomatic carrieres, η_{as}, is smaller than 1. Settings are R_0 =3 p_{app} =0.6, η_{DCT} =1.0, trace uninfected contacts = false. The values printed for R_e correspond to: (first number) the mean minus the standard deviation, and (second number) the mean plus the standard deviation, of the distribution of R_e from 100 simulations (compare Fig. 6 bottom right panel), while the color of the field shows the mean. Where values are exactly 0.0, none of the 100 simulations had an outbreak (compare Sect. 4.5).

4.3 The effect of reduced contagiousness of asymptomatic carriers

As we saw in Fig. 3, in a situation where reported symptomatic cases are quarantined, down-scaling the contagiousness of asymptomatic carriers reduces R significantly. Fig. 11 shows the outcomes when DCT is then applied.

In the case where $\eta_{as} = 0.1$ and all symptomatic carriers are reported when symptoms start, the simulations generated no outbreaks (fewer than 400 people became infected in total). When 75% of symptomatic cases are reported, R_e has large fluctuations between simulation runs, and outcomes are very sensitive to the fraction of symptomatic cases.

4.4 The effect of timing, delays, and second order tracing

Fig. 12 shows the outcomes as a function of tracing delay, that is the time in days between when an index case is identified and when his or her contacts are traced and quarantined. Outcomes are again grouped by whether or not uninfected contacts are identified by tracing. Results are also shown for both 1st order tracing and 2nd order tracing and for the two incubation time curves (the default one with mean incubation time (IT) of 7 days and the alternative one with the shorter IT=4 days). The yellow marker uses the default incubation time curve with the alternate transmission probability curve (IC) where there is less pre-symptomatic transmission. In addition to the simulation results, the calculated R_e from the age-of-infection model is shown for the settings using the default incubation time and transmission probability curves, and first order

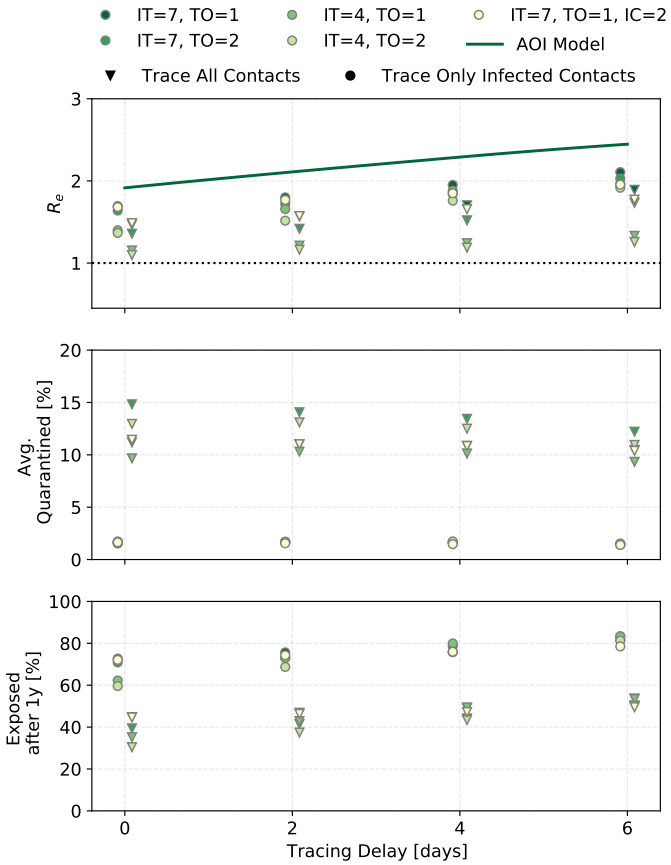


Figure 12: The effect of incubation time (IT) and tracing order (TO) for different tracing delays. IT=7 refers to the curve with mean incubation period of 7 days, and IT=4 to the one with mean incubation period of 3.6 days. The yellow point shows results for the alternate transmission probability curve (IC) as shown in Tab. 1 (grey values). Predictions from the "age-of-infection" (AOI) model, where only infected contacts are traced, are shown as the dark green line for parameters IT=7, TO=1 - the result shown is not exact (see appendix). This is for $p_{\text{app}} = 0.75$, $\eta_{DCT} = 1$, and $\alpha \cdot f_m = 0.6$.

tracing. Approximations had to be made in the calculation, hence the absolute value is not expected to match the simulation results perfectly.

Results are shown for an app coverage of 75%. Some of the dynamics are quite sensitive to the app coverage (see appendix), and at 75% trends are clearer than at 60% app coverage.

The difference between first and second order tracing is small in all three outcomes (this changes in some situations for higher app coverages). For the default incubation time curve with mean of 7 days, tracing delays of up to 6 days have only a small effect, increasing the number of people exposed after 1 year

from approximately 72% to 82%. When the mean incubation time is only 4 days, tracing delays more strongly reduce R_e The infection probability curve with less pre-symptomatic transmission probability only very slightly improves the outcomes, though the effect becomes bigger with increasing tracing delays.

Second order tracing can find the infector and through him or her, the 'siblings' of the index case. The chance that the infection took place within ΔT_{trace} is higher with a longer ΔT_{trace}. However even for ΔT_{trace} =14 days, the outcomes are not significantly different. The look-back time must be balanced against the number of healthy people quarantined. People typically become index cases before they have recovered, and thus would have had a chance to infect others in the approximately 7 days prior. Looking back longer than that means one has a bigger chance of finding the infector, but it also means tracing many uninfected contacts.

When considering the realistic case where uninfected contacts are traced, second order tracing with a 7 day look-back time sends about 1.3 times as many people into quarantine on average over 1 year as first order tracing.

4.5 Outbreak probability

The results discussed so far consider situations where an outbreak is ongoing and interventions are started at some point into the outbreak. But not all simulation runs result in an outbreak. The stochastic nature of the outbreak means that there are large statistical variations at the beginning of the chain. For example, if patient zero happens to not infect anyone, no outbreak happens.

The chance for an outbreak to occur increases with R. The more people a case typically infects, the less likely it is that cases at the beginning of the infection chain do not infect anyone. Therefore, keeping interventions in place even in populations without an ongoing outbreak can be useful to decrease the probability that an outbreak will occur when a case is introduced into the population, for example through travel.

Figure 13 shows the outbreak probability as a function of the reproductive number when an infected person enters a fully susceptible population.

4.6 Sensitivity of results to the social contact structure

The results presented so far assumed a homogeneous population of size 1×10^5 and a distribution of the number of contacts with infection potential per day from Eq. (3). We also ran some sets of parameters for different population sizes and for different contact structures. The results are shown in Fig. 14.

The introduction of a social graph introduces non-linear effects that change R_e on timescales much longer than what is captured by our standard analysis. In some cases, this means that fewer people are exposed after one year, even though R_e is higher (see Fig. 21 in Appendix C).

5 Discussion

Contact tracing relies on index cases from which to trace. When there is a large fraction of mildly symptomatic and asymptomatic carriers who never go to the doctor or get tested, many carriers do not become index cases, so DCT does

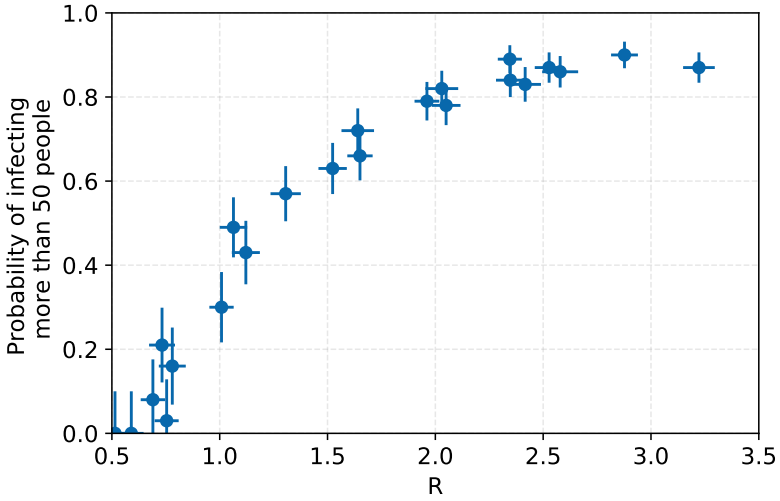


Figure 13: The probability for an outbreak to start as a function of the reproductive number at the time when patient 0 enters a fully susceptible population. An outbreak here is defined as more than 50 people becoming infected. The error bars shown are statistical.

not have a large impact. Outcomes improve strongly the higher the fraction of reported symptomatic carriers. This is partially because DCT is more efficient, and partially because R is additionally reduced just from quarantining the index cases. Therefore it is crucial that every person with even the mildest symptoms has easy access to a COVID-19-test.

The extend to which pre- and asymptomatic carriers drive the outbreak depends on their contagiousness. If for some reason they are less contagious than symptomatic carriers, missing them as index cases does not worsen outcomes much. In the case where η_{as} is 0.1, as proposed for example in [4], quarantining of index cases, without CT, reduces R from $R_0 = 3$ to $R_e < 1$ even when just 40% of cases are symptomatic.

Randomly testing a fraction of the population regularly to find unreported carriers helps to make up for the large fraction of asymptomatic carriers. We find that a very large fraction of the population must be tested daily to significantly improve outcomes. For our default parameters, even when testing 20% of the population daily, at least 90% of the population would have to use the DCTS for R_e to become smaller than one.

Reducing the contact rate (social distancing) by as little as 20% is as effective as testing 20% of the population every day while requiring fewer people to be quarantined.

Tracing delays of a few days do not significantly worsen the outcomes. Two studies, Ferretti et al. [4] and [5], indicate that a DCTS could control a SARS-CoV-2 outbreak (that is achieve $R_e < 1$) because it allows for contact tracing without delays. We find that the asymptomatic infectiousness scaling of η_{as} =0.1 used by [4] is the main driver of their R_e and given these starting conditions, DCTS only has to lower R by a small amount to achieve outbreak control and

is therefore then effective. Kretzschmar et al. [5] are more careful about the reduction in R achievable with DCTS, but do confirm the improved outcomes with short tracing delays. However, [5] use a very short latency period. With the longer median latency periods consistent with recent large-scale studies, this effect is small. Therefore, the advantage of a DCT in the case of COVID-19 lies mostly in the possibility to scale tracing to a large number of cases without needing a large increase in the number of manual contact tracers.

Most models consider that contacts that were actually infected are traced with some probability. In reality, it is impossible to tell immediately whether or not a traced contact has been infected. Even if a test performed immediately on tracing is negative, it could just mean that the person is still in the latent period. Therefore, all traced contacts should be quarantined and tested multiple times. In principle, one could devise other schemes, such as testing each traced person every morning (e.g. with a POC tests that can be done at home and gives results within minutes) for a few days without requiring quarantine unless the test comes up positive. Right now, such frequent testing is not realistic in most countries.

We find that including the effect that quarantining of uninfected contacts has on the outbreak dynamics can lead to significantly different, typically more positive, outcomes compared to models where this effect is ignored. The improvement in outcomes is due to the large number of people quarantined even though they are healthy. Our simulations probably underestimate this number, because we use contact rates for the types of contacts that have a high chance of transmitting a respiratory virus. A DCT system will typically pick up many contacts who were in spacial proximity to the index case, but not in a manner that was likely to transmit the virus, so the number of contacts traced per index case could be bigger in reality. For any serious large-scale use of a DCT system during an ongoing epidemic, dealing with these uninfected contacts in quarantine is going to be a major challenge, especially as the compliance of the population with quarantining procedures may decrease once someone has been traced and quarantined multiple times.

The statistical nature of virus transmission and contact rates leads to large variations in outbreak dynamics at the start of the outbreak. Sometimes, an infectious person entering a susceptible population does not start an outbreak. This becomes less likely the higher R is. This also means that under identical conditions, one population could have hundreds of cases within a week of the arrival of patient 0, while in another population the case number does not start rising for several weeks, just by luck.

Beside control of the outbreak, that is achieving R_e below 1, an important outcome is how many people will have been exposed by the time a vaccine might be available. Due to the heavy social and economical burden imposed by virus control, some countries are aiming at R_e around 1, rather than total control. Our simulations assume that the same interventions are applied throughout the epidemic, so the outcomes over 1 year are guidelines rather than realistic predictions for any real country. They do give a qualitative idea of what achieving a given R_e means in terms of the number of people exposed (and with that, the number of fatalities).

For reasons of computing power, most simulations were run for a homogeneous population. In reality, society is organized into social units. Introducing such social units into the simulation means that an infectious person tends to

meet the same people every day, exposing them again and again. This leads to slight changes in outcome, while the qualitative results remain the same.

5.1 Limitations

We assumed that all people, once they have recovered from the infection, are immune to a secondary infection. Whether and for how long a recovered person is immune remains to be answered. Studies show that neutralizing antibodies are produced during infection and to a higher degree in symptomatic carriers, but decline significantly 2-3 months after recovery [54, 55]. The minimal antibody titer to confer protection is, however, still unclear. Furthermore, memory T cells to SARS-CoV-2 have been found in patients including asymptomatic and mildly symptomatic ones, which likely contribute to protective immunity as well [56].

In our models, everyone adheres to quarantine protocols. That is, every time someone is alerted by the DCT system to having been in proximity of a contagious individual, this person must follow the quarantining and testing procedure. This is crucial to suppress pre- and a-symptomatic transmission, but may be difficult to achieve in reality. We also assume that the fraction of symptomatic individuals who see a doctor/get tested do so the day they become symptomatic.

The transmission probability in our models changes with the time since infection, but not between individuals. Current research, however, suggests that COVID-19 is overdispered, meaning some individuals spread the virus to many others, in so-called "superspreading events", while most do not transmit the virus at all or only to very few people [57]. Part of this overdispersion is due to the random nature of the contact number - some people just meet more others, and is therefore included in our models (see Sect. 2.2 and Fig. 2).

We assume that no manual tracing is performed at all. Typically, the types of close contact persons to whom spreading the disease is most likely, that is friends and family, can be manually traced without much effort, hence the fraction of infected contacts traced could be larger in reality.

6 Summary

Many countries enforced a policy of 'shelter–in–place' and/or extreme social distancing, effectively putting most of the population into quarantine. This significantly slowed down the infection rate [58][59], but came with large economical and social costs to society. World-wide, a lot of effort has been put into the development of CT systems, in the hope that large-scale CT could replace other public health measures at much smaller cost to society.

We modelled the effect of instantaneous DCT in combination with a testing and quarantining protocol, as well as random testing and social distancing, on an ongoing COVID-19 epidemic. Results were validated by running the scenarios with two independently developed individual-based models, which were further cross-checked by two types of deterministic models. We modelled many different parameter values for the still not well-known properties of SARS-CoV-2, COVID-19 and for the interventions, leading to well over 10 000 simulated scenarios. The goal was to find the regions in this parameter space where CT

without additional interventions could lower the effective reproductive number enough to halt exponential growth.

Wherever modelling approximations had to be made, we chose defaults that lead to better outcomes, hence these results are likely on the optimistic side. Our results are stable under different simulated social structures and epidemiological parameters, with significantly different outcomes seen only when varying the fraction of asymptomatic individuals or down-scaling the contagiousness of pre- and asymptomatic cases.

We find that for large regions of the parameter space, including the currently most likely parameter values, an outbreak of COVID-19 cannot be fully controlled by DCT even if a large fraction of the population uses the system. Furthermore, if interventions are started once an outbreak is already ongoing, DCT causes a large fraction of the healthy population to be traced and quarantined.

DCT can be combined with other measures, such as face-mouth coverings, social distancing, and/or random testing, to achieve outbreak control.

The availability of fast testing, and coordination of test results with the DCT system, are crucial to allow symptomatic cases to become index cases for tracing, and to release traced healthy contacts from quarantine. Since SARS-CoV-2 symptoms are unspecific, everyone with even a slight cough of fever must be able to get a test (a) quickly, because the infection probability peaks just before symptom onset and then falls quickly and people who are not sure they are infected likely will not effectively quarantine themselves, and (b) easily, so that a large fraction of symptomatic cases do seek out testing. The gains of a DCT system in outbreak control quickly vanish if many symptomatic cases do not seek out testing, or if positively tested individuals do not become index cases.

7 Acknowledgements

We acknowledge the support by the DFG Cluster of Excellence "Origin and Structure of the Universe". The simulations were carried out on the computing facilities of the Computational Center for Particle and Astrophysics (C2PAP) and on the Max Planck Institute for Physics computing cluster. This work was undertaken in the context of the ContacTUM collaboration.

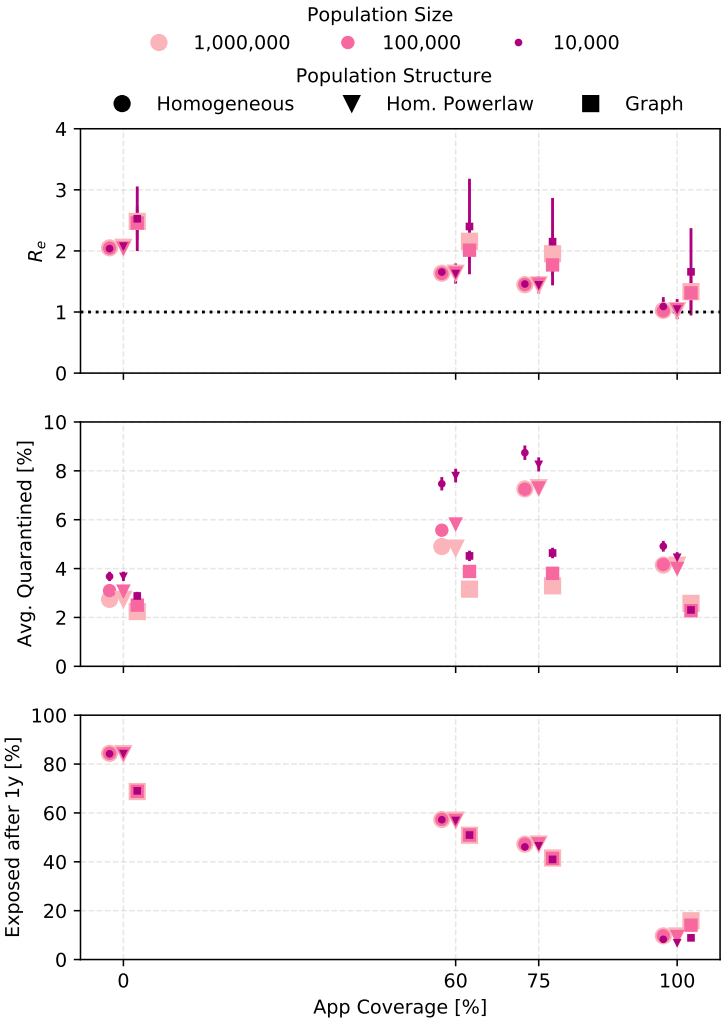


Figure 14: The sensitivity of the outcomes to the size and the structure of the simulated population is shown. The three population structures (described in Sect. 2.1) correspond to the homogeneous population with the gamma distribution (Eq. (3)) describing the number of contacts per day, a homogeneous distribution using the power law (Eq. (4)) for the number of contact per day, and the social graph population. Settings are $R_0 = 3$, $\alpha \cdot f_m = 0.6$, $\eta_{DCT} = 1$, trace uninfected=true. For the 10 000 people population, outcomes have large statistical fluctuations and error bars show the error on the mean. For the other points, the error bars are smaller than the marker size.

References

[1] Camilla Rothe, Mirjam Schunk, Peter Sothmann, et al. Transmission of 2019-NCOV infection from an asymptomatic contact in Germany. *New England Journal of Medicine*, 382(10):970–971, 2020 DOI:10.1056/NEJMc2001468.

[2] Vernon J Lee, Calvin J Chiew, and Wei Xin Khong. Interrupting transmission of COVID-19: lessons from containment efforts in Singapore. *Journal of Travel Medicine*, 27(3), 2020 DOI:10.1093/jtm/taaa039.

[3] Christophe Fraser, Steven Riley, Roy M Anderson, and Neil M Ferguson. Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences*, 101(16):6146–6151, 2004 DOI:10.1073/pnas.0307506101.

[4] Luca Ferretti, Chris Wymant, Michelle Kendall, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, 2020 DOI:10.1126/science.abb6936.

[5] Mirjam Kretzschmar, Ganna Rozhnova, Martin Bootsma, et al. Time is of the essence: impact of delays on effectiveness of contact tracing for COVID-19. *medRxiv*, 2020 DOI:10.1101/2020.05.09.20096289.

[6] Australia. COVIDSafe app. https://github.com/AU-COVIDSafe. Accessed June 09, 2020.

[7] Austria. StoppCorona app. https://github.com/austrianredcross. Accessed: June 09, 2020.

[8] France. StopCovid app. https://gitlab.inria.fr/stopcovid19. Accessed: June 09, 2020.

[9] India. Aarogya Setu app. https://www.mygov.in/aarogya-setu-app. Accessed: June 09, 2020.

[10] Iceland. Rakning C19. https://github.com/aranja/rakning-c19-app. Accessed: June 09, 2020.

[11] Italy. Immuni App. https://github.com/immuni-app. Accessed: June 09, 2020.

[12] Norway. Smittestopp app. https://github.com/djkaty/no.simula.smittestopp. Accessed: June 09, 2020.

[13] UK. NHS Covid-19 app. https://github.com/nhsx/. Accessed: June 09, 2020.

[14] Singapore. TraceTogether. https://github.com/OpenTrace-community and https://bluetrace.io. Accessed: June 09, 2020.

[15] Switzerland. SwissCovid App. https://github.com/DP-3T/dp3t-app-android-ch. Accessed: 2020-06-09.

[16] Isobel Braithwaite, Thomas Callender, Miriam Bullock, and Robert W Aldridge. Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19. *The Lancet Digital Health*, 2020 DOI:10.1016/s2589-7500(20)30184-9.

[17] Joel Hellewell, Sam Abbott, Amy Gimma, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health*, 8(4):488–496, 2020 DOI:10.1016/S2214-109X(20)30074-7.

[18] Josh A. Firth, , Joel Hellewell, et al. Using a real-world network to model localized COVID-19 control strategies. *Nature Medicine*, 2020 DOI:10.1038/s41591-020-1036-8.

[19] Joë L Mossong, Niel Hens, Mark Jit, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine*, 5, 2008 DOI:10.1371/journal.pmed.0050074.

[20] Andrea Lancichinetti, Santo Fortunato, and Filippo Radicchi. Benchmark graphs for testing community detection algorithms. *Phys. Rev. E*, 78: 046110, 2008 DOI:10.1103/PhysRevE.78.046110.

[21] Günce Keziban Orman, Vincent Labatut, and Hocine Cherifi. Towards realistic artificial benchmark for community detection algorithms evaluation. *International Journal of Web Based Communities*, 9, 2013, arxiv:1308.0577v1 DOI:10.1504/IJWBC.2013.054908.

[22] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, 2002 DOI:10.1103/RevModPhys.74.47.

[23] Shujuan Ma, Jiayue Zhang, Minyan Zeng, et al. Epidemiological parameters of coronavirus disease 2019: a pooled analysis of publicly reported individual data of 1155 cases from seven countries. *medRxiv*, 2020 DOI:10.1101/2020.03.21.20040329.

[24] Xi He, Eric HY Lau, Peng Wu, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature medicine*, 26, 2020 DOI:10.1038/s41591-020-0869-5.

[25] Steven Sanche, Yen Ting Lin, Chonggang Xu, et al. High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2. *Emerging infectious diseases*, 2020 DOI:10.3201/eid2607.200282.

[26] Stephen A Lauer, Kyra H Grantz, Qifang Bi, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases : Estimation and Application. *Annals of Internal Medicine*, 2020 DOI:10.7326/M20-0504.

[27] Kenji Mizumoto, Katsushi Kagaya, Alexander Zarebski, and Gerardo Chowell. Estimating the Asymptomatic Proportion of 2019 Novel Coronavirus onboard the Princess Cruises Ship , 2020. *Euro surveillance : European communicable disease bulletin*, 25(10), 2020 DOI:10.1101/2020.02.20.20025866.

[28] Shin Young Park, Young-Man Kim, Seonju Yi, et al. Coronavirus Disease Outbreak in Call Center, South Korea. *Emerging Infectious Diseases*, 2020 DOI:10.3201/eid2608.201274.

[29] Enrico Lavezzo, Elisa Franchin, Constanze Ciavarella, et al. Suppression of COVID-19 outbreak in the municipality of Vo ', Italy. *medRxiv*, 2020 DOI:10.1101/2020.04.17.20053157.

[30] Victor M. Corman, Holger F. Rabenau, Ortwin Adams, et al. SARS-CoV-2 asymptomatic and symptomatic patients and risk for transfusion transmission. *medRxiv*, 2020 DOI:10.1101/2020.03.29.20039529.

[31] Hiroshi Nishiura, Tetsuro Kobayashi, Ayako Suzuki, et al. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *International Journal of Infectious Diseases*, 2020 DOI:10.1016/j.ijid.2020.03.020.

[32] Ruiyun Li, Sen Pei, Bin Chen, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science*, 3221, 2020 DOI:10.1126/science.abb3221.

[33] Lirong Zou, Feng Ruan, Mingxing Huang, et al. SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. *New England Journal of Medicine*, 382(12):1177–1179, 2020 DOI:10.1056/NEJMc2001737.

[34] Kelvin Kai-Wang To, Owen Tak-Yin Tsang, Wai-Shing Leung, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *The Lancet Infectious Diseases*, 3099(20), 2020 DOI:10.1016/s1473-3099(20)30196-1.

[35] Nathan W. Furukawa, John T. Brooks, and Jeremy Sobel. Evidence supporting transmission of severe acute respiratory syndrome coronavirus 2 while presymptomatic or asymptomatic. *Emerging infectious diseases*, 26, 2020 DOI:10.3201/eid2607.201595.

[36] Roman Wölfel, Victor M. Corman, Wolfgang Guggemos, et al. Virological assessment of hospitalized patients with COVID-2019. *Nature*, 2020 DOI:10.1038/s41586-020-2196-x.

[37] Wei Li, Ying-Ying Su, Shen-Shen Zhi, et al. Viral shedding dynamics in asymptomatic and mildly symptomatic patients infected with sars-cov-2. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 2020 DOI:10.1016/j.cmi.2020.07.008.

[38] Seungjae Lee, Tark Kim, Eunjung Lee, et al. Clinical Course and Molecular Viral Shedding Among Asymptomatic and Symptomatic Patients With SARS-CoV-2 Infection in a Community Treatment Center in the Republic of Korea. *JAMA Internal Medicine*, 2020 DOI:10.1001/jamainternmed.2020.3862.

[39] Yang Liu, Li Meng Yan, Lagen Wan, et al. Viral dynamics in mild and severe cases of COVID-19. *The Lancet Infectious Diseases*, 2019(20):2019–2020, 2020 DOI:10.1016/S1473-3099(20)30232-2.

[40] Jantien A. Backer, Don Klinkenberg, and Jacco Wallinga. Incubation period of 2019 novel coronavirus (2019- ncov) infections among travellers from wuhan, china, 20 28 january 2020. *Eurosurveillance*, 25, 2020 DOI:10.2807/1560-7917.ES.2020.25.5.2000062.

[41] Hendrik Streeck, Bianca Schulte, Beate M Kümmerer, et al. Infection fatality rate of SARS-CoV-2 infection in a German community with a superspreading event. *medRxiv*, 2020 DOI:10.1101/2020.05.04.20090076.

[42] Oyuka Byambasuren, Magnolia Cardona, Katy J L Bell, et al. Estimating the extent of true asymptomatic covid-19 and its potential for community transmission: systematic review and meta-analysis. *medRxiv*, 2020 DOI:10.1101/2020.05.10.20097543v2.

[43] Ying Liu, Albert A Gayle, Annelies Wilder-Smith, and Joacim Rocklöv. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of travel medicine*, 27(2), 2020 DOI:10.1093/jtm/taaa021.

[44] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz. Superspreading and the effect of individual variation on disease emergence. *Nature*, 438(7066):355–359, 2005 DOI:10.1038/nature04153.

[45] O. Diekmann and JAP Heesterbeek. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. Wiley, 2000. ISBN 9780471492412.

[46] F Brauer. *Mathematical Epidemiology: Compartmental Models in Epidemiology, Lecture Notes in Mathematics*. Springer, 2008. ISBN 978-3-540-78911-6.

[47] Odo Diekmann, Hans Heesterbeek, and Tom Britton. *Mathematical tools for understanding infectious disease dynamics, Princeton Series in Theoretical and Computational Biology*. Princeton University Press, 2012. ISBN 9781400845620.

[48] Johannes Müller and Christina Kuttler. *Methods and Models in Mathematical Biology Deterministic and Stochastic Approaches*. Springer-Verlag Berlin Heidelberg, 2015. ISBN 978-3-642-27250-9.

[49] H W Hethcote and J A Yorke. *Gonorrhea transmission dynamics and control*. Springer, 1984. ISBN 978-3-662-07544-9.

[50] Istvan Z. Kiss, Darren M. Green, and Rowland R. Kao. Infectious disease control using contact tracing in random and scale-free networks. *Journal of the Royal Society Interface*, 2006 DOI:10.1098/rsif.2005.0079.

[51] Glenn Webb, Cameron Browne, Xi Huo, et al. A model of the 2014 ebola epidemic in West Africa with contact tracing. *PLoS Currents*, 2015 DOI:10.1371/currents.outbreaks.846b2a31ef37018b7d1126a9c8adf22a.

[52] Johannes Müller, Mirjam Kretzschmar, and Klaus Dietz. Contact tracing in stochastic and deterministic epidemic models. *Mathematical Biosciences*, 164, 2000 DOI:10.1016/S0025-5564(99)00061-9.

[53] Cameron Browne, Hayriye Gulbudak, and Glenn Webb. Modeling contact tracing in outbreaks with application to ebola. *J. Theor. Biol.*, 384:33–49, 2015 DOI:10.1016/j.jtbi.2015.08.004.

[54] Jeffrey Seow, Carl Graham, Blair Merrick, et al. Longitudinal evaluation and decline of antibody responses in sars-cov-2 infection. *medRxiv*, 2020 DOI:10.1101/2020.07.09.20148429.

[55] Quan Xin Long, Xiao Jun Tang, Qiu Lin Shi, et al. Clinical and immunological assessment of asymptomatic sars-cov-2 infections. *Nature Medicine*, 26, 2020 DOI:10.1038/s41591-020-0965-6.

[56] Takuya Sekine, André Perez-Potti, Olga Rivera-Ballesteros, et al. Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. *bioRxiv*, 2020 DOI:10.1101/2020.06.29.174888.

[57] Akira Endo, Sam Abbott, Adam J. Kucharski, and Sebastian Funk. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Research*, 5:67, 2020 DOI:10.12688/wellcomeopenres.15842.1.

[58] Adam J. Kucharski, Timothy W. Russell, Charlie Diamond, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases*, 20, 2020 DOI:10.1016/S1473-3099(20)30144-4.

[59] Bruce (WHO) Aylward and Wannian (PRC) Liang. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19). *The WHO-China Joint Mission on Coronavirus Disease 2019*, pages 16–24, 2020.

[60] Robert L. Harrison. The construction of next-generation matrices for compartmental epidemic models. *The Royal Society Interface*, 7:873–885, 2010 DOI:10.1098/rsif.2009.0386.

A The ODE model

The compartment model is based on the SEIR-type model from Reference [53] and is extended to visualize infectious individuals for their whole infectious period T_{con} even though they might be quarantined or hospitalized (see Figure 15). To do so, we introduce the convalescing compartment 'C' in between the infectious I and recovered R status.

I denotes the infectious individuals that are able to infect susceptible individuals S for the mean infectious time T. In contrast, C are convalescing individuals that are still infectious but cannot infect further as they are either isolated or the probability to infect is effectively too small. The latter case reflects the fact that the probability to infect is reduced for large times post infection as shown in Figure 1.[1] After the residual infectious time $T_{\text{con}} - T$, C enter the recovered status R. In Fig. 18, the infectious population is defined as the sum of all I and C subgroups.

[1] A similar second infectious state has been implemented by Ref.[29] to describe individuals that can be tested positive beyond their infectious period. While the viral load in this second state is still large enough for detection, it is too small for further infection.

The model takes into account contact tracing and differentiates between untraced and traced individuals. Contact tracing is triggered by reported infectious individuals that appear with the fraction α of the untraced individuals. The model incorporates forward contact tracing, e.g. one traces the contacts that have been infected by the reported index case and predicts the probability p_e that the traced contact is still exposing when traced. This prediction relies on exponential probability distributions for the latent and infectious period. We assume that traced contacts that are still exposing E_t^e are immediately isolated. In comparison, contacts that have been traced during their infectious period I_t^i could have already infected further people. The corresponding mean infectious time T_t of I_t^i is calculated within the model and depends on the latent and the infectious period of the reported infectious individuals. Traced infectious contacts I_t^i can trigger contact tracing as well and the probability that the traced contact of I_t^i is still exposing is given by p_e^i that depends then on T_t. Table A summarizes the parameters and definitions of the compartmental model. For a detailed description and the proof of the computation of p_e, p_e^i, and T_t, we refer the reader to Reference [53].

In order to compute the doubling time T_2 and the basic reproduction number R_e, we focus on the initial infection-free state, in which the entire sample population is given by S_0 [53, 60]. At this stage, the system of differential equations can be linearised, and the aforementioned parameters can be extracted from the Jacobian of the system J. In particular, the doubling time is computed as $\ln(2)/\lambda$, where λ is the largest eigenvalue of J, known as the exponential-growth parameter. On the other hand, the basic reproduction number is obtained using the *next-generation analysis*, in which the Jacobian is split into two matrices: F, containing the terms relative to the generation of new infections, and V containing the transfer from one infectious compartment to another. The basic reproduction number R_e is then obtained as the spectral radius of the matrix FV^{-1}.

We derive the mean latent and infectious times T_{lat}, T_r, and T_u using the probability functions described in Section 2 to 5.0 d, 1.4 d, and 2.5 d, respectively. Further, the total infectious time T_{con} is assumed to be 10.4 d.

The model is given by the following set of ordinary differential equations

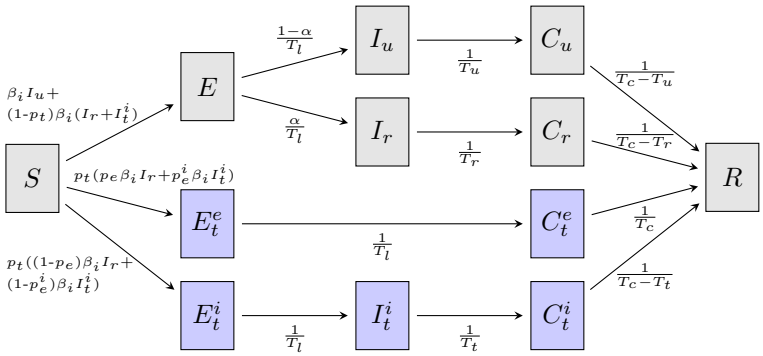


Figure 15: Structure of the compartmental model. The model is based on a SEIR-type model from Ref. [53] and is extended for the convalescing (C) compartment. A person in the convalescing status is technically still infectious, but the chance to infect further is drastically reduced, as this person is either isolated or the viral load is effectively too low. The sum of the specific infectious periods is given by T_{con}. Untraced compartments are indicated in gray and traced compartments in blue. See Table A for the definitions of the variables.

(ODEs):

$$\frac{d}{dt}S = -\beta_i \frac{S}{N} I_u - \beta_i \frac{S}{N} I_r - \beta_i \frac{S}{N} I_t^i, \tag{11}$$

$$\frac{d}{dt}E = \beta_i \frac{S}{N} I_u + (1-p_t)\beta_i \frac{S}{N} I_r + (1-p_t)\beta_i \frac{S}{N} I_t^i - \frac{1}{T_{\text{lat}}} E, \tag{12}$$

$$\frac{d}{dt}E_t^e = p_t p_e \beta_i \frac{S}{N} I_r + p_t p_e^i \beta_i \frac{S}{N} I_t^i - \frac{1}{T_{\text{lat}}} E_t^e, \tag{13}$$

$$\frac{d}{dt}E_t^i = p_t(1-p_e)\beta_i \frac{S}{N} I_r + p_t(1-p_e^i)\beta_i \frac{S}{N} I_t^i - \frac{1}{T_{\text{lat}}} E_t^i, \tag{14}$$

$$\frac{d}{dt}I_u = \frac{1-\alpha}{T_{\text{lat}}} E - \frac{1}{T_u} I_u, \tag{15}$$

$$\frac{d}{dt}I_r = \frac{\alpha}{T_{\text{lat}}} E - \frac{1}{T_r} I_r, \tag{16}$$

$$\frac{d}{dt}I_t^i = \frac{1}{T_{\text{lat}}} E_t^i - \frac{1}{T_t} I_t^i, \tag{17}$$

$$\frac{d}{dt}C_u = \frac{1}{T_u} I_u - \frac{1}{T_{\text{con}} - T_u} C_u, \tag{18}$$

$$\frac{d}{dt}C_r = \frac{1}{T_r} I_r - \frac{1}{T_{\text{con}} - T_r} C_r, \tag{19}$$

$$\frac{d}{dt}C_t^e = \frac{1}{T_{\text{lat}}} E_t^e - \frac{1}{T_{\text{con}}} C_t^e, \tag{20}$$

$$\frac{d}{dt}C_t^i = \frac{1}{T_t} I_t^i - \frac{1}{T_{\text{con}} - T_t} C_t^i, \tag{21}$$

$$\frac{d}{dt}R = \frac{1}{T_{\text{con}} - T_u} C_u + \frac{1}{T_{\text{con}} - T_r} C_r + \frac{1}{T_{\text{con}}} C_t^e + \frac{1}{T_{\text{con}} - T_t} C_t^i \tag{22}$$

Table 2: Parameters and Definitions of the compartmental model.

Parameter	Definition
$S(t)$	number of susceptible individuals
$E(t)$	number of untraced exposing individuals
$E_t^e(t)$	number of exposed individuals that will be traced during exposed phase
$E_t^i(t)$	number of exposed individuals that will be traced during infectious phase
$I_u(t)$	number of unreported infectious individuals
$I_r(t)$	number of reported infectious individuals
$I_t^i(t)$	number of infectious individuals that will be traced during infectious phase
$C_u(t)$	number of unreported convalescing individuals
$C_r(t)$	number of isolated convalescing reported individuals
$C_t^e(t)$	number of isolated convalescing individuals that were traced during exposed phase
$C_t^i(t)$	number of isolated convalescing individuals that were traced during infectious phase
$R(t)$	number of recovered and immune individuals. Note that individuals in R could also be dead.
$N(t) = N$	constant sum of all individuals
β_i	transmission rate
α	reporting fraction of all individuals. Note that we do not distinguish between asymptomatic and symptomatic cases.
T_{lat}	latency period
T_{con}	total infectious period
T_u	infectious period of unreported individuals
T_r	infectious period until reporting
T_t	infectious period of traced individuals I_t^i. $1/T_t = 1/\bar{T} + 1/\hat{T}_t$ with $\hat{T}_t = \frac{1}{T_r}\int_0^{T_r}\left[T_r + \delta - \left(\int_r^{T_r+\delta} \frac{x \exp\left(-(x-r)/T_{\text{lat}}\right)}{\int_0^{T_r-r+\delta} \exp\left(-s/T_{\text{lat}}\right) ds} dx\right)\right] dr$ and $\bar{T} = \alpha T_r + (1-\alpha)T_u$ [53].
p_t	tracing probability. Note that we do not distinguish between p_{app} and η_{DCT}
p_e	probability that individual is traced during exposed phase. Probability depends on T_{lat}, T_r, δ: $p_e = \frac{T_{\text{lat}}}{T_{\text{lat}}+T_r} e^{-\delta/T_{\text{lat}}}$ [53]
p_e^i	probability that individual who got infected by I_t^i is traced during exposed phase. Probability depends on T_{lat}, T_t, δ: $p_e^i = \frac{T_{\text{lat}}}{T_{\text{lat}}+T_t} e^{-\delta/T_{\text{lat}}}$ [53]
δ	time between reporting of index case and tracing of contact

B The age since infection model

We propose here a simple deterministic model for contact tracing, where the class of infeceteds is structured by age since infection.

Let $S(t)$ denote the density of susceptible at time t, $I(t,a)$ the density of infecteds at time t with age of infection a, and $R(t)$ the removed individuals at time t (recovered, quarantined, dead – in any case, not infectious any more). Note that the infected individuals may not be infectious, if they are still in the latent period – exposed and infectious are distinguished by the age of infection. The total amount of the infecteds at time t is given by $I(t) = \int_0^\infty I(t,a)\,da$. $N(t) = S(t) + \int_0^\infty I(t,a)\,da + R(t)$ denotes the total population size. Since we do not consider population dynamics (i.e. the population size does not change), $N(t) = N$ is a constant.

We first describe the model without contact tracing and discuss how to incorporate contact tracing afterwards. Infected individuals with age of infection a have infectivity $\beta(a)$, an infected person recovers spontaneously without diagnosis at rate $\mu(a)$; alternatively, an infected person develops symptoms and gets diagnosed at rate $\sigma(a)$. In that case, he/she is quarantined immediately and will not infect further susceptibles. We chose standard incidence, s.t. the model equations become

$$\frac{d}{dt}S(t) = -S(t)\int_0^\infty \beta(a)I(t,a)\,da/N, \tag{23}$$

$$(\partial_t + \partial_a)I(t,a) = -(\mu(a)+\sigma(a))I(t,a), \tag{24}$$

$$\frac{d}{dt}R(t) = \int_0^\infty (\mu(a)+\sigma(a))I(t,a)\,da. \tag{25}$$

In order to prepare for the effect of contact tracing, we slightly rewrite the model equations. Thereto, we note that

$$\kappa(a) = e^{\int_0^a \mu(\tau)+\sigma(\tau)\,d\tau}$$

is the probability to be in the class I at time of infection a. Below, we will modify $\kappa(a)$. For now, we note that

$$\text{removal rate} = (\mu(a)+\sigma(a)) = -\frac{d}{da}\ln\Big(\kappa(a)\Big)$$

(which is also called the hazard rate). therewith, our model becomes

$$(\partial_t + \partial_a)I(t,a) = \left(\frac{d}{da}\ln\Big(\kappa(a)\Big)\right)I(t,a), \tag{26}$$

$$\frac{d}{dt}R(t) = \int_0^\infty -\left(\frac{d}{da}\ln\Big(\kappa(a)\Big)\right)I(t,a)\,da. \tag{27}$$

In order to compute the effect of contact tracing, we only need to adapt $\kappa(a)$ accordingly. The overall model structure is not touched. The analysis of the probability to be infectious at age of infection a as given below is mathematically precise for the onset of the infection.

Contact tracing

We now slightly switch the perspective, and consider single individuals. What follows was basically published in [52]; we summarize these considerations for the convenience of the reader. Below, we investigate a forest: An infected individual infects other individuals, in that, we get a random tree with a directed node from infector to infectee. Individuals who recover leave this tree, that is, we are left with a forest.

Contact tracing acts on this forest. At rate $\sigma(a)$ an individual is (directly) detected/diagnosed and forms an index case. All neighbouring individuals (within the forest of infecteds) have probability p to be detected by contact tracing. Either we stop here (one step tracing) or the individuals detected by CT form recursively new index cases (recursive tracing). This process modifies $\kappa(a)$: contact tracing increases the probability to be removed at age of infection a.

In order to exactly understand how $\kappa(a)$ does look like, we use a mathematical fiction: First, we assume that we can only trace contacts from infectee to infector (backward tracing). Then, we assume that we can only trace contact from infector to infectee (forward tracing). Last, we combine both approaches to understand the full tracing we aim at.

Backward tracing

Proposition B.0.1. *The probability to be infectious for an individual at age of infection a for recursive contact tracing follows the following system of integro-differential equations,*

$$\frac{d}{da}\kappa^-(a) = -\kappa^-(a)\left(\mu(a)+\sigma(a)+\eta_{DCT}\int_0^a \beta(c)\,\kappa^-(a-c)\left(\frac{-\kappa^{-\prime}(a-c)}{\kappa^-(a-c)}-\mu(a-c)\right)\right) \quad (28)$$

with $\kappa^-(0)=1$.

Proof: Clearly, without contact tracing, we have $\kappa^{-\prime}(a) = -(\mu(a)+\sigma(a))\,\kappa^-(a)$. If backward tracing is active, we have an additional component of the removal rate that is caused by a tracing event triggered by an infectee. As only infectees cause tracing (tracing events are only triggered by "children"), the probability to be infectious at a given age of infection a is the same for infector and infectee. Hence the recovery rate of an infectee can be written as the hazard rate

$$\frac{-\kappa_*^{-\prime}(b)}{\kappa_*^-(b)}.$$

This hazard rate includes the rate of direct observation (that triggers a backward tracing event), the rate at which the infectee is discovered by recursive tracing (which triggers tracing event), and by spontaneous removal (which does not trigger a tracing event). Spontaneous removal does not lead to contact tracing, hence we subtract this rate and find the rate at which direct diagnosis or detection by contact tracing happens,

$$\frac{-\kappa_*^{-\prime}(b)}{\kappa_*^-(b)}-\mu(b).$$

The focal individual (for which we compute $\kappa^-(a)$) produces during his/her infectious time span (so far) $[0, a]$ infectees. When he/she has had age $c \in [0, a]$, his/her infection rate was $\beta(c)$. By now, the age of the focal individual/the infector is a. The probability that the infectee is now still infectious reads $\kappa(a-c)$). The rate of direct or indirect observation of the infectee with age of infection $a-c$ is $\frac{-\kappa_*^{-\prime}(a-c)}{\kappa_*^-(a-c)} - \mu(a-c)$. A detected individual triggers a successful tracing event with probability η_{DCT}. Hence, the contribution to the removal rate of our focal individual due to tracing is given by

$$\eta_{DCT} \int_0^a \beta(c)\, \kappa^-(a-c) \left(\frac{-\kappa^{-\prime}(a-c)}{\kappa^-(a-c)} - \mu(a-c) \right) dc.$$

□

Proposition B.0.2. *The probability to be infectious for a symptomatic/asymptomatic individual at age of infection a for one-step contact tracing follows the following system of integro-differential equations,*

$$\frac{d}{da}\kappa^-(a) = -\kappa^-(a) \left(\mu(a) + \sigma(a) + \eta_{DCT} \int_0^a \beta(c) \left\{ \kappa^-(a-c)\, \sigma(a-c) \right\} \right) \quad (29)$$

$$(30)$$

with $\kappa^-(0) = 1$.

Proof: The proof parallels that of proposition B.0.1; we only need to know that individuals have the rate of direct detection $\sigma(a-c)$; this rate replaces the expression (detection rate for recursive tracing) $\frac{-\kappa_*^{-\prime}(a-c)}{\kappa_*^-(a-c)} - \mu(a-c)$.

□

Forward tracing

In contrast to backward tracing, in forward tracing the position of a focal individual in the tree/forest of infecteds matter. If I'm the primary infected person (generation 0), I have no infector (inside of the population). I cannot be traced by forward tracing. The first generation (infectee of generation 0) can only be traced via the primary infecetd person. The second generation can be traced by the zeroth and first generation, and so on.

That is, for forward tracing, the "generation" of an individual does influence the probability to be infectious at age of infection a. "Generation" does refer to generation of infection; the primary case has generation 0, those infected by the primary case have generation 1 etc. We denote by $\kappa_i^+(a)$ the probability to be infectious at age of infection a for an individual of generation $i \in \mathbb{N}_0$ under forward tracing.

Furthermore, we introduce the probability to be infectious at age of infection a in case that we do not have contact tracing ($p = 0$) for symptomatic/asymptomatic infecteds,

$$\hat{\kappa}(a) = e^{-\int_0^a \mu(s)+\sigma(s)\, ds}. \quad (31)$$

In order to determine $\kappa_i^+(a)$, we first introduce and investigate

$$\kappa_i^+(a|b),$$

which is the probability that an individual in generation i is infectious at age of infection a, if we condition on the fact that the infector has age of infection $a+b$ (s.t. the infector – at the time of the corresponding infectious contact – has has age of infection b). We find the following result

Proposition B.0.3. *In case of recursive tracing we have for $i > 0$*

$$\kappa^+(a|b) \quad = \quad \hat{\kappa}(a)\left\{1 - \eta_{DCT}\int_0^a \left(\frac{-(\kappa^+_{i-1}(b+c))'}{\kappa^+_{i-1}(b+c)} - \mu(b+c)\right)\frac{\kappa^+_{i-1}(b+c)}{\kappa^+_{i-1}(b)}\,dc\right\} \tag{32}$$

Proof: Our focal individual is infectious if it did not recover already without contact tracing, times the probability that no tracing event did remove the individual from the class of infecteds,

$$\kappa^+_i(a|b) \quad = \quad \hat{\kappa}(a)\Big\{1 - \text{ a successful tracing event did happen}\Big\}.$$

In order to obtain the probability for a successful tracing event, we first note that we know that the infector has been infectious at (his/her) age of infection b, s.t. the probability for him/her to be infectious at age of infection $b+c$ reads

$$\frac{\kappa^+_{i-1}(b+c)}{\kappa^+_{i-1}(b)}.$$

As before, for recursive tracing, the detection rate is the hazard rate minus the rate to recover spontaneously/unobserved,

$$\frac{-\kappa^+_{*_2,i-1}(b+c)'}{\kappa^+_{*_2,i-1}(b+c)} - \mu(b+c).$$

Hence, the desired probability reads

$$\text{a successful tracing event did happen} = \eta_{DCT}\int_0^a \left(\frac{-\kappa^+_{i-1}(b+c)'}{\kappa^+_{i-1}(b+c)} - \mu(b+c)\right)\frac{\kappa^+_{i-1}(b+c)}{\kappa^+_{i-1}(b)}\,dc.$$

□

Proposition B.0.4. *In case of one-step-tracing we have*

$$\kappa^+_i(a|b) \quad = \quad \hat{\kappa}(a)\left\{1 - \eta_{DCT}\int_0^a \sigma(b+c)\,\frac{\kappa^+_{i-1}(b+c)}{\kappa^+_{i-1}(b)}\,dc\right\}. \tag{33}$$

Proof: The argument parallels that of proposition B.0.3. We only need to take into account that in one-step tracing the infectee has to be detected directly, what happens at rate $\sigma(.)$. This rate replaces the hazard rate minus the spontaneous recovery rate.

□

In order to determine the desired probability $\kappa^+_i(a)$ we remove the condition in $\kappa^+_i(a|b)$. Thereto we determine the probability density for an infector to have age b of infection b. The net infection rate is $\beta_i(b)\,\kappa_{i-1}(b)$. Therefore, the distribution of the age of the infector at the time of infection is given by

$$\varphi_{i-1}(b) = \frac{\beta_i(b)\,\kappa_{i-1}(b)}{\int_0^\infty \beta_i(c)\,\kappa_{i-1}(c)\,dc}.$$

Corollary B.0.5. *In one-step tracing as well as in recursive tracing, we have for $i > 0$*

$$\kappa_i^+(a) \quad = \quad \int_0^\infty \kappa_i^+(a|b)\varphi_{i-1}(b)\,db \tag{34}$$

where we have to use for $\kappa_i^+(a|b)$ the solution for one-step or recursive tracing, depending on the scenario chosen.

We find an iterative formula. Analysis for p small as well as numerical analysis (for general $\eta_{DCT} \in [0,1]$) shows that the convergence is rather fast: after 3-5 generations, the $\kappa_i^+(a)$ are basically converged.

Full tracing

Full tracing is just a combination of forward- and backward tracing. Let $\kappa_i(a)$ denote the "survival probability" for a target individual of generation i under full tracing.

In order to find $\kappa_0(a)$, we only need to understand that the primary infected individual can only be traced by downstream infecteds, that is, is only exposed to backward tracing. Also the next generations have – without forward tracing – just the "survival" probability $\kappa^-(a)$. We need to multiply this probability with the probability not to be target of a forward tracing event in order to obtain $\kappa_i(a)$.
That is, we use the argument for forward tracing, where we replace $\hat{\kappa}(a)$ (survival probability without tracing) by $\kappa^-(a)$ (survival probability under backward tracing only), and get immediately the following result (notation is an obvious extension of the notation above).

Proposition B.0.6. *In case of recursive tracing we have for $i > 0$*

$$\kappa(a|b) \quad = \quad \kappa^-(a)\left\{1 - \eta_{DCT}\int_0^a \left(\frac{-(\kappa_{i-1}(b+c))'}{\kappa_{i-1}(b+c)} - \mu(b+c)\right)\frac{\kappa_{i-1}(b+c)}{\kappa_{i-1}(b)}\,dc\right\} \tag{35}$$

Proposition B.0.7. *In case of one-step-tracing we have*

$$\kappa_i(a|b) \quad = \quad \kappa^-(a)\left\{1 - \eta_{DCT}\int_0^a \sigma(b+c)\,\frac{\kappa_{i-1}(b+c)}{\kappa_{i-1}(b)}\,dc\right\}. \tag{36}$$

With

$$\varphi_{i-1}(b) = \frac{\beta_i(b)\,\kappa_{i-1}(b)}{\int_0^\infty \beta_i(c)\,\kappa_{i-1}(c)\,dc}$$

and for recursive contact tracing (one-step: parallel formula)

$$\begin{aligned}
\int_0^\infty \kappa_i(a|b)\varphi_{i-1}(b)\,db &= \frac{\int_0^\infty \kappa_i(a|b)\beta_i(b)\,\kappa_{i-1}(b)\,db}{\int_0^\infty \beta_i(c)\,\kappa_{i-1}(c)\,dc} \\
&= \frac{\int_0^\infty \left[\kappa^-(a)\left\{1 - \eta_{DCT}\int_0^a \left(\frac{-(\kappa_{i-1}(b+c))'}{\kappa_{i-1}(b+c)} - \mu(b+c)\right)\frac{\kappa_{i-1}(b+c)}{\kappa_{i-1}(b)}\,dc\right\}\right]\beta_i(b)\,\kappa_{i-1}(b)\,db}{\int_0^\infty \beta_i(c)\,\kappa_{i-1}(c)\,dc} \\
&= \kappa^-(a)\,\frac{\int_0^\infty \left\{1 - \eta_{DCT}\int_0^a \left(\frac{-(\kappa_{i-1}(b+c))'}{\kappa_{i-1}(b+c)} - \mu(b+c)\right)\frac{\kappa_{i-1}(b+c)}{\kappa_{i-1}(b)}\,dc\right\}\beta_i(b)\,\kappa_{i-1}(b)\,db}{\int_0^\infty \beta_i(c)\,\kappa_{i-1}(c)\,dc}
\end{aligned}$$

we finally get the result for full tracing (we summarize all necessary equations):

Theorem B.0.8. [recursive tracing] *Let*

$$\frac{d}{da}\kappa^-(a) \quad = \quad -\kappa^-(a)\left(\mu(a)+\sigma(a)+\eta_{DCT}\int_0^a \beta_i(c)\,\kappa^-(a-c)\left(\frac{-\kappa^{-\prime}(a-c)}{\kappa^-(a-c)}-\mu(a-c)\right)\right) \tag{37}$$

with $\kappa^-(0)=1$. *Then,*

$$\kappa_i(a) \quad = \quad \kappa^-(a)\,\frac{\int_0^\infty\left\{1-\eta_{DCT}\int_0^a\left(\frac{-(\kappa_{i-1}(b+c))'}{\kappa_{i-1}(b+c)}-\mu(b+c)\right)\frac{\kappa_{i-1}(b+c)}{\kappa_{i-1}(b)}\,dc\right\}\beta_i(b)\,\kappa_{i-1}(b)\,db}{\int_0^\infty\beta_i(c)\,\kappa_{i-1}(c)\,dc} \tag{38}$$

Theorem B.0.9. [one-step tracing] *Let*

$$\frac{d}{da}\kappa^-(a) \quad = \quad -\kappa^-(a)\left(\mu(a)+\sigma(a)+\eta_{DCT}\int_0^a \beta_i(c)\,\kappa^-(a-c)\;\sigma(a-c)\right) \tag{39}$$

with $\kappa^-(0)=1$. *Then,*

$$\kappa_i(a) \quad = \quad \kappa^-(a)\,\frac{\int_0^\infty\left\{1-\eta_{DCT}\int_0^a\sigma(b+c)\;\frac{\kappa_{i-1}(b+c)}{\kappa_{i-1}(b)}\,dc\right\}\beta_i(b)\,\kappa_{i-1}(b)\,db}{\int_0^\infty\beta_i(c)\,\kappa_{i-1}(c)\,dc} \tag{40}$$

Note that it is straightforward (but tedious) to simplify these equations, s.t. they become handy for the numerical analysis.

Reproduction number and exponential growth

A direct consequence of the dynamic, age-structured model is the possibility to determine the reproduction number and the exponential growth rate right away from the survival probability $\kappa_\infty(a)$. The effective reproduction number is simply given by

$$R_e \quad = \quad \int_0^\infty \beta_i(a)\kappa_\infty(a)\,da. \tag{41}$$

The exponent λ of the exponential growth in the onset $S(t)\approx S_0\approx N$ can be determined by the unique real root of the equation

$$1=\int_0^\infty e^{-\lambda a}\beta_i(a)\kappa_\infty(a)\,da. \tag{42}$$

From here, we obtain $T_2=\ln(2)/\lambda$.

Reproduction number in case of DCT

Recall that we consider a homogeneously mixing population, where a fraction p_{app} of individuals have a DCT device. Only contacts between these individuals can be traced, with probability η_{DCT}. Let $R_{eff}(p)$ denote the reproduction number of a homogeneous model, where each contact is traced with probability p. Then, a straightforward generalization of the considerations above yields that non-app-users have the reproduction number $R_{eff}(0)$, while those with app have the reproduction number

$$R_{eff}=R_{eff}(\eta_{DCT}\;p_{\text{app}}).$$

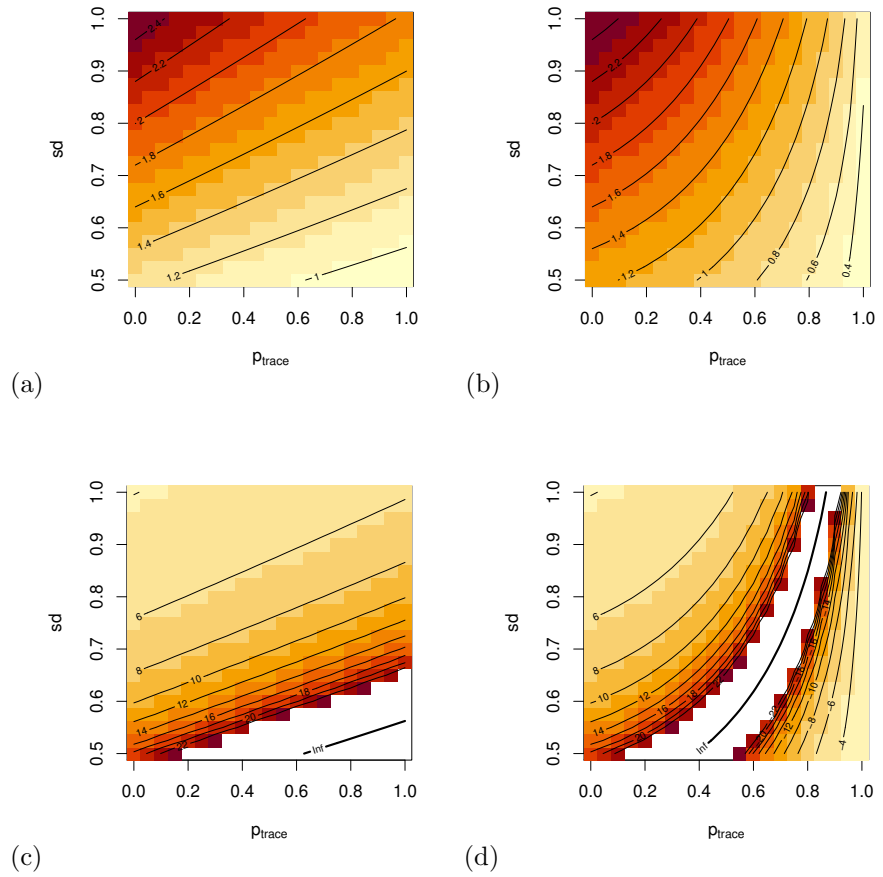


Figure 16: Influence of social distancing (reduction of the contact rate) and contact tracing on R_e and T_2. (a) and (c): One-step tracing, (b) and (d) recursive tracing. (a) and (b) R_e, (b) and (d) T_2. Note that in (b) and (c), there is a singularity for T_2 at the line $R_e = 1$, where T_2 becomes infinite.

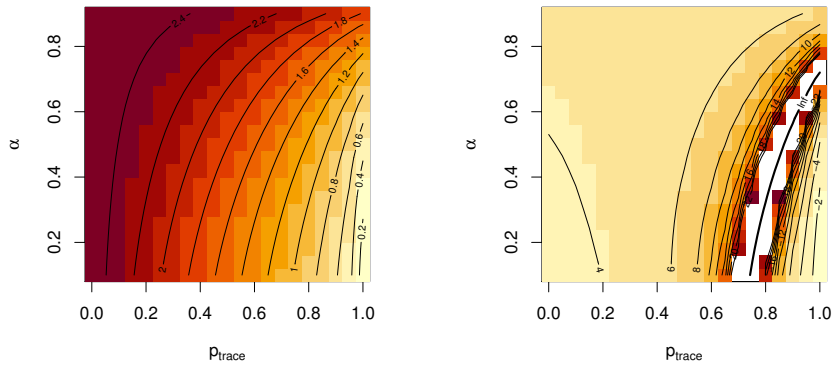
Since we assume homogeneous mixing, we obtain the overall reproduction number as

$$R_{eff} = p_{\rm app}\, R_{eff}(\eta_{DCT}\ p_{\rm app}) + (1 - p_{\rm app}) R_{eff}(0).$$

This formula is exact, but due to the nonlinearity in $R_{eff}(.)$ it cannot be simplified. However, if $p_{\rm app}, \eta_{DCT} \ll 1$, we can linearize at $p_{\rm app} = 0$, and find

$$\begin{aligned} R_{eff} &\approx p_{\rm app}\, R_{eff}(\eta_{DCT}\ p_{\rm app}) + (1 - p_{\rm app}) R_0 = p_{\rm app}\, (R_0 + p_{\rm app} p\, R_{eff}^{ct}{}'(0)) + (1 - p_{\rm app}) R_0 \\ &= R_0 - p_{\rm app}^2\, \eta_{DCT}\, R_{eff}^{ct}{}'(0)) \approx R_{eff}(\eta_{DCT}\, p_{\rm app}^2) \end{aligned}$$

In lowest order, tracing within a subgroup of relative size $p_{\rm app}$ and a tracing probability η_{DCT} is equivalent with tracing the total population with a tracing probability $p_{\rm app}^2 \eta_{DCT}$.



(a) (b)

Figure 17: Influence of the assumed fraction of asymptomatic and unreported cases on (a) R_e and (b) T_2, where we also take η_{DCT} into account (only recursive tracing considered).

Results

We identify the parameters of the models as described below. Using numerical analysis, it is possible to estimate the influence of control measures on the spread of the infection. Particularly, we are interested in the effect of contact tracing and social distancing on the effective reproduction number and the doubling time. Moreover, the sensitivity of the results on parameters for which there is little data is explored. Here, our focus is the fraction of asymptomatic cases.

If we inspect Fig. 16, we find first of all that recursive tracing is more efficient than one-step tracing (one must not confuse one-step tracing with level-1 tracing: here, we only follow infectious contacts for one step). In practice, the tracing delay will lead to a situation between one-step tracing and recursive tracing: even if recursive tracing is the aim, as we loose time in tracing each contacts, after a few steps persons might already be recovered when traced. In any case, small tracing probabilities have only a minor impact. If we increase the tracing probability to 0.7-0.8, we find that an eradication of the infection is possible. On the other side, social distancing is most effective for small tracing probabilities. The combination of contact tracing and social distancing is most likely best in the middle range: a reduction of contacts by 30%, and a tracing probability around 0.6-0.7 will be able to reduce the doubling time considerably, and even bring R_e down to values around 1.

The parameter values are not all precisely known. In particular, at the present time, the fraction of undiagnosed cases is rather unclear. In the literature, there are numbers between 1.2% and 95%. The parameter scan in Fig. 17 indicates that the results are rather stable for a wide range of α (from 10-50%). A major change can be observed above 70%. This figure might imply that the results are rather stable against the parameter choice in α.

Choice of parameter functions

The medical investigations yield particularly data on:
– Incubation period (time to the onset of symptoms/diagnosis)
– latent period (time until infectivity becomes positive) and viral load (a proxy for infectivity)
– fraction of asymptomatic cases.

Incubation period

Time to symptoms: We use a Γ distribution, density

$$f(a) = \frac{1}{\beta_i \Gamma(\gamma_I)} \left(\frac{a}{\beta_i}\right)^{\gamma_I - 1} e^{-a/\beta_i} \tag{43}$$

with parameters $\beta_i = 2.44$, $\gamma_I = 3.06$ days. Let $\hat{\sigma}(a)$ be the conditioned detection rate: It is only valid under the condition that an individual is detected, indeed. That rate is just the hazard rate of $f(a)$,

$$\hat{\sigma}(a) = \frac{f(a)}{1 - \int_0^a f(a')\, da'}.$$

We later need to compute the rate $\sigma(a)$ for the model, where we take into account the fraction of asymptomatic persons.

Infectivity

Given that the onset of symptoms of a person at age A_0 (a Gamma-distributed random variable, as described above), the infectious period starts $A_0 - \Delta_{onset}$, where Δ_{onset} is a fixed time span (we choose $\Delta_{onset} = 3$ days). If $a_0 - \Delta < 0$, then the infectivity period starts right away at the time of the infection. That is, $\max\{A_0 - \Delta_{onset}, 0\}$ is the latent period. We furthermore assume that the infectious period is a fixed (deterministic) time span T_{con}.
This assumption about the latent period is an input for the age-dependent infection rate as well as for the recovery rate.

(b.1) Recovery rate

If $A_0 \sim \text{Gamma}(\beta_i, \mu_I)$ is a random variable that states the onset of symptoms, we aim at the distribution of the recovery age $A_r = \max\{A_0 - \Delta_{onset}, 0\} + T_{\text{con}}$. Clearly, $P(A_r < a) = 0$ for $a \leq T_{\text{con}}$. A short computation yields for $a > T_{\text{con}}$

$$\begin{aligned} P(A_r < a) &= P(A_r < a | A_0 < \Delta_{onset}) P(A_0 < \Delta_{onset}) + P(A_r < a | A_0 > \Delta_{onset}) P(A_0 > \Delta_{onset}) \\ &= P(A_0 < \Delta_{onset} + a - T_{\text{con}}). \end{aligned}$$

Hence, in total we have

$$P(A_r < a) = \begin{cases} 0 & \text{if} \quad a < T_{\text{con}} \\ 1 - \int_0^{\Delta_{onset} + a - T_{\text{con}}} f(a')\, da' & \text{if} \quad a > T_{\text{con}} \end{cases}.$$

As the cumulative distribution involves a jump, the hazard rate (recovery rate) incorporates a delta peak. However, for the practical implication, we replace the

jump by a steep linear increase during a small age interval. The corresponding hazard rate yields the removal rate $\hat{\mu}(a)$, conditioned on the fact that a person will not be diagnosed by recover spontaneously. The parameter $T_{\rm con}$ is computed below.

(b.2) Infectivity

We again denote by A_0 the random variable that states the onset of the symptoms, $A_0 \sim \text{gamma}(\beta_i, \gamma_I)$.

The ingredient for the infectivity is an approximation of the virus load, given by a shifted Gamma distribution,

$$\beta_i(a|A_0) = \frac{1}{\beta_x \Gamma(\gamma_x)} \left(\frac{(a - (\mu_x + A_0))_+}{\beta_x} \right)^{\gamma_x - 1} e^{-(a-(\mu_x+A_0))/\beta_x}$$

with the understanding that this formula is only valid for $A_0 > -\mu_x$. Here, $(x)_+ = 0$ for $x < 0$ and $(x)_+ = x$ if $x > 0$. The parameters are given by:

$$\mu_x = -2.43 \text{ days} , \quad \beta_x = 1.56, \quad \gamma_x = 2.08 \text{ days} .$$

If $A_0 + \mu_x < 0$, we assume that symptoms start right away. Hence,

$$\beta_i(a|A_0) = \frac{1}{\beta_x \Gamma(\gamma_x)} \left(\frac{(a - \max\{A_0 + \mu_x, 0\})_+}{\beta_x} \right)^{\gamma_x - 1} e^{-(a-\max\{A_0+\mu_x,0\})/\beta_x}.$$

Then (as in our case $\mu_x < 0$). β_i is proportional to

$$\begin{aligned} \beta_i(a) &\sim \int_0^\infty \beta_i(a|b) f(b)\, db = \int_0^{-\mu_x} \beta_i(a|b) f(b)\, db + \int_{-\mu_x}^\infty \beta_i(a|b) f(b)\, db \\ &= P(A_0 < -\mu_x) \beta_i(a|b=0) + \int_{-\mu_x}^\infty \beta_i(a|b) f(b)\, db, \end{aligned}$$

where $f(b)$ is given in (43). The unknown proportionality constant models the number of contacts per day. This constant is calibrated s.t. the basic reproduction number R_0 is 2.5.

We define $T_{\rm con}$ by

$$\int_0^{T_{\rm con}} \beta_i(a|A_0 = -\mu_x)\, da = 0.99.$$

Asymptomatic cases

We define the effective detection rate $\sigma(a)$ and the effective removal rate $\mu(a)$, based on the the conditioned rates $\hat{\sigma}(a)$ and $\hat{\mu}(a)$, and the fraction of asymptomatic cases α. The probability to be infectious is a convex combination of the probabilities to be infectious for a symptomatic resp. asymptomatic individual (where we understand that an "asymptomatic individual" is not only asymptomatic at a given time, but will stay undiagnosed until recovery). That is,

$$\kappa(a) = \alpha\, e^{-\int_0^a \hat{\mu}(\tau)\, d\tau} + (1 - \alpha)\, e^{-\int_0^a \hat{\sigma}(\tau)\, d\tau}.$$

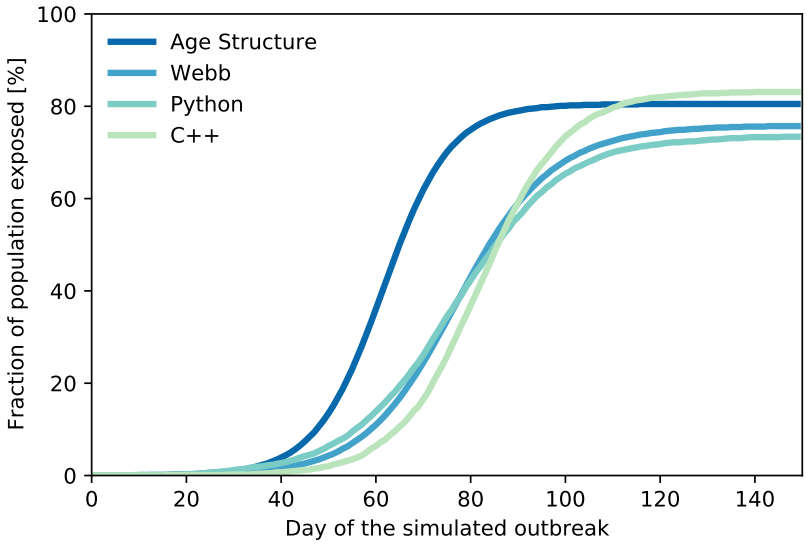


Figure 18: The cumulative number of exposed people from the four models and for the settings from Tab. 3.

In order to find the effective rates, we compute the hazard rate of $\kappa(a)$,

$$\frac{\alpha\,\mu(a)\,e^{-\int_0^a \hat{\mu}(\tau)\,d\tau} + (1-\alpha)\,\sigma(a)\,e^{-\int_0^a \hat{\sigma}(\tau)\,d\tau}}{\alpha\,e^{-\int_0^a \hat{\mu}(\tau)\,d\tau} + (1-\alpha)\,e^{-\int_0^a \hat{\sigma}(\tau)\,d\tau}}$$

From here, we immediately find the desired definitions for $\mu(a)$ and $\sigma(a)$,

$$\begin{aligned}\sigma(a) &= \frac{(1-\alpha)\hat{\sigma}(a)\,e^{-\int_0^a \hat{\sigma}(\tau)\,d\tau}}{\alpha e^{-\int_0^a \hat{\mu}(\tau)\,d\tau} + (1-\alpha)e^{-\int_0^b \hat{\sigma}(\tau)\,d\tau}},\\ \mu(a) &= \frac{\alpha\hat{\mu}(a)\,e^{-\int_0^a \hat{\mu}(\tau)\,d\tau}}{\alpha e^{-\int_0^a \hat{\mu}(\tau)\,d\tau} + (1-\alpha)e^{-\int_0^b \hat{\sigma}(\tau)\,d\tau}}.\end{aligned}$$

C Model benchmarking

Monte Carlo models rely on computer code, and therefore must be verified to work as designed. We do this in two ways: a) by comparing results to deterministic models, and b) by comparing the results from two implementations written independently by two teams using two different programming languages and toolsets. One example for (a) is shown in Fig. 18. The deterministic models cannot model all effects we consider, therefore POC testing and social distancing was turned off in the Monte Carlo models, and other parameters were chosen as summarized in Tab. 3. The parameters in the ODE model do not have straightforward correspondence to the parameters in the table, so they were tuned such that the outcome matches the other models as well as possible.

The analytic models are, strictly speaking, not valid once a sizable fraction of the population has been exposed, since they do not model non-linear effects

Table 3: Table of settings used to benchmark the four models against each other.

Parameter/Setting	Benchmark value
Disease and population	
Size of population	10k
Population structure	uniform
Transmission probability (β_i)	$\begin{pmatrix} 0.0595 \\ 5.0 \\ 3.2 \end{pmatrix}$
Contact rate	
R0	
Trans. prob. curve (μ, γ, β)	(-2.42, 2.08, 1.56)
Incubation time curve (μ, γ, β)	(0, 3.06, 2.44)
Fraction symptomatic (α)	0.5
Asymptomatic trans. scaling (η_{as})	1.0
Interventions	
Interventions start (f_i)	0.00
Quarantine duration	14 days
Tracing	
Reported from symptoms (f_m)	1.0
Trace back (ΔT_{trace})	14 [days]
App coverage (p_{app})	1
Tracing efficiency (η_{DCT})	0.7
Tracing order	1
Trace uninfected contacts	False
Tracing delay (T_{delay})	0 [days]

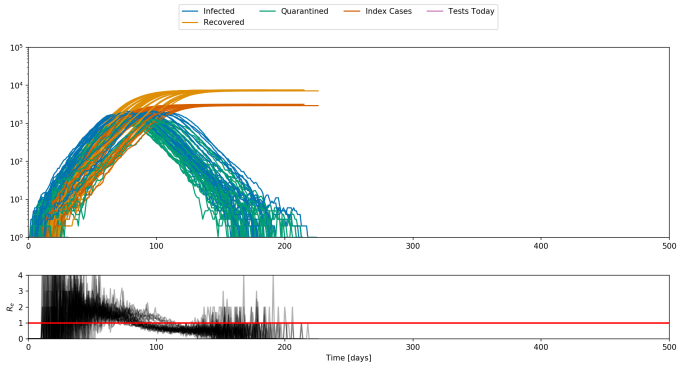


Figure 19: Outbreak dynamics when someone with zero days latency period **cannot** infect others on the same day he or she became infected.

such as two exposed people meeting each other. Nevertheless we show the full outbreak curves here.

The ODE model matches quite well with the Python and C++ models, but it has the advantage of having been tuned to achieve as good a match as possible. The age-structure model has a time offset compared to the others, but the time when the outbreak takes off in the IB models is very variable so such offsets are not relevant.

Even though both individual-based models implement the same disease parameters and intervention protocols, there are a number of implementations details that are treated differently and can lead to slightly different outcomes. Examples of this are whether nor not people with a 0 days latency period can infect others on the same day they themselves became infected, or whether we allow someone to still infect others on the day they become symptomatic and are quarantined.

Figs. 19 and 20 compare what happens for two implementations of a latency period of zero days. In one implementation, it is possible for the carrier to infect others on the same day he or she became exposed. In the other implementation, a carrier can infect others only on days after he or she became exposed.

One big difference between the IB models is that the Python model accounts for the case where a contact is traced from an index case who did not infect the contact, but in the time between when the contact took place and when the person was traced, the contact was infected by someone else. In the C++ model, this special case is not considered. The effect on the outcomes is small, especially when taking a look-back time of only 7 days as done for most of the main scan.

Fig. 21 shows how the population structure affects the course of the epidemic. Social graph populations lead to long-term changes in the outbreak dynamic that are not accurately captured by just the reproductive number reached after interventions.

Fig. 22 is similar to Fig. 6 but for different parameters. In this scenario, the reproductive number changes both due to the interventions and naturally because a large enough fraction of the population becomes exposed early on that non-linear effects are important.

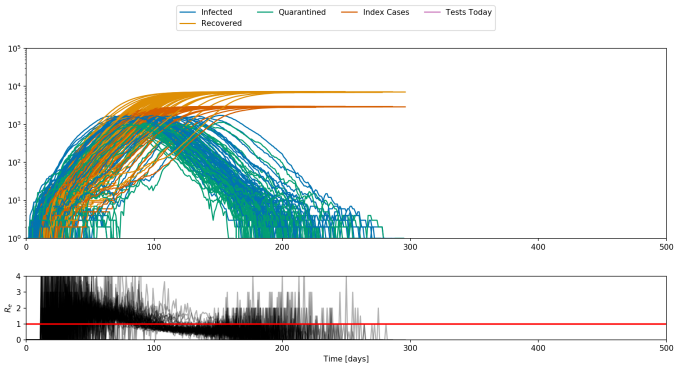


Figure 20: Outbreak dynamics when someone with zero days latency period **can** infect others on the day he or she became infected.

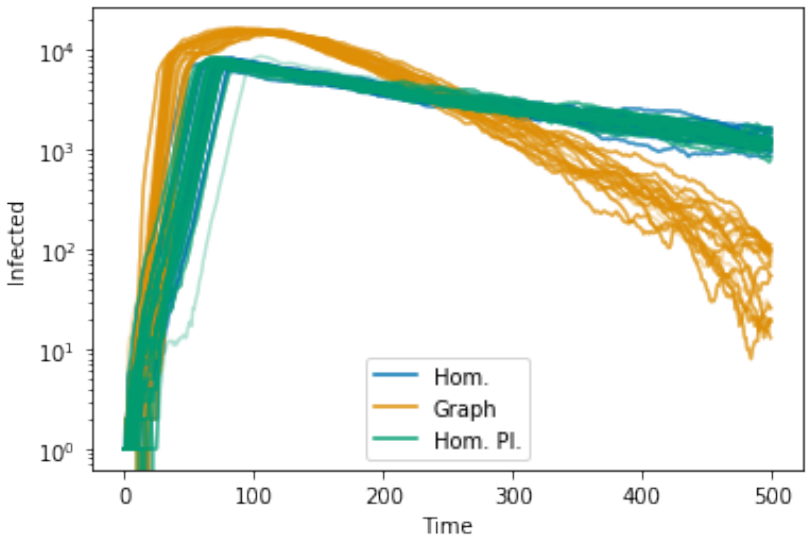


Figure 21: The number if infectious people each day for a social graph and a homogeneous population is compared.

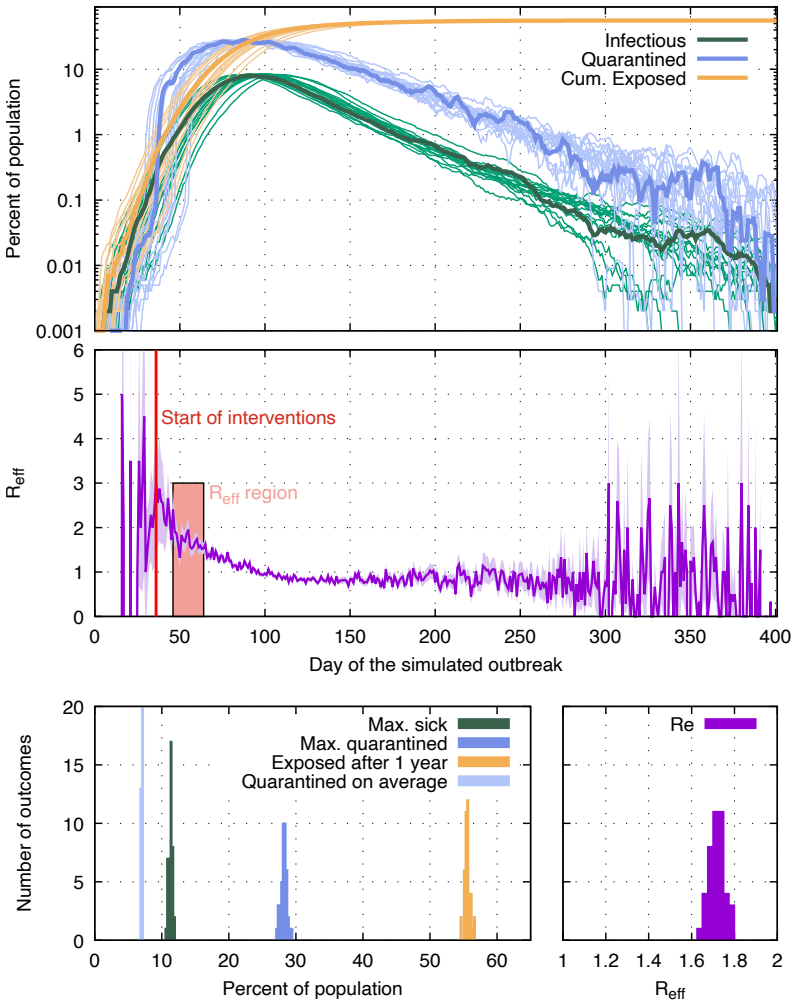


Figure 22: Same as Fig. 6 changing the following parameters: p_{app} =0.6, $\alpha \cdot f_m$ =0.6.

D Results from all scenarios

The results for the main parameter scan are shown in Figs. 23 through 28. Results for the scan with different timing and delays are shown in Figs. 29 through 34

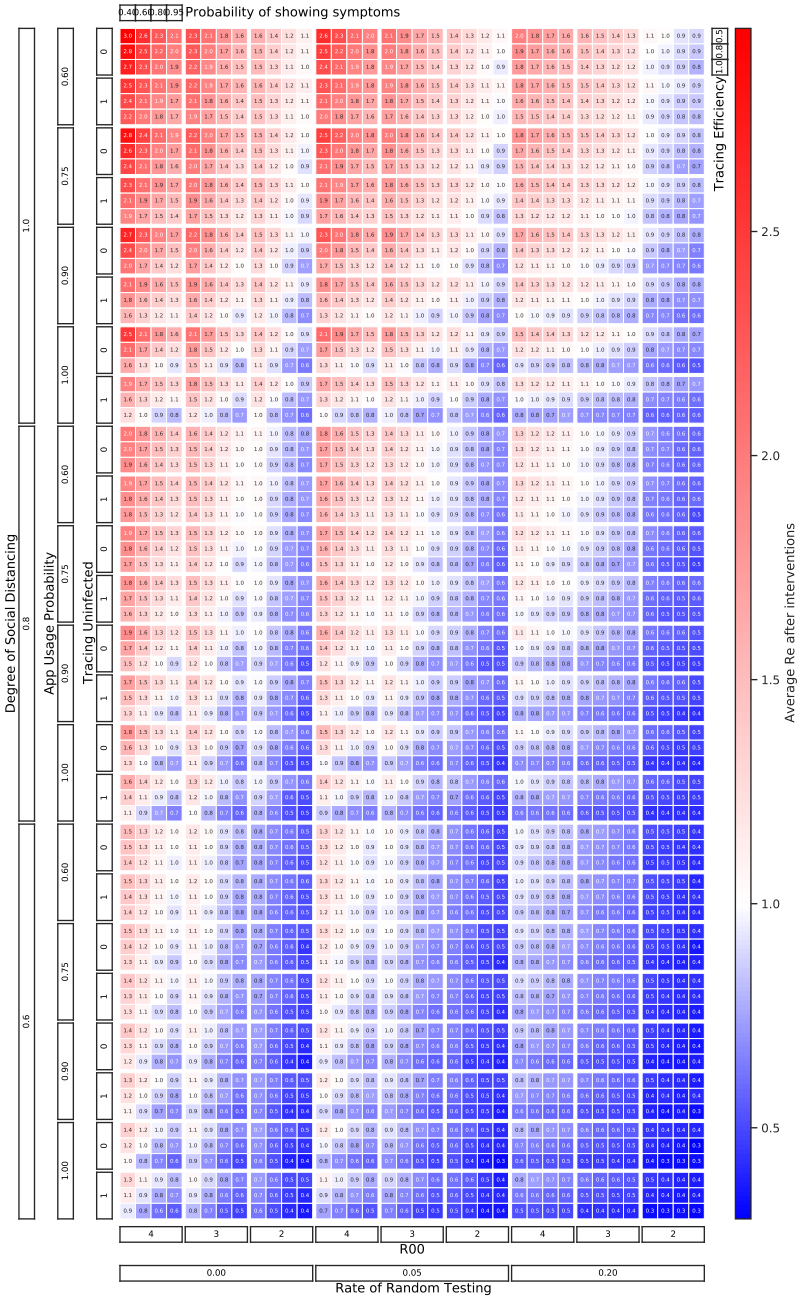


Figure 23: The effective reproductive number reached once interventions are turned on is shown for all combinations of parameters in the default parameter scan (black font color in Tab. 1).

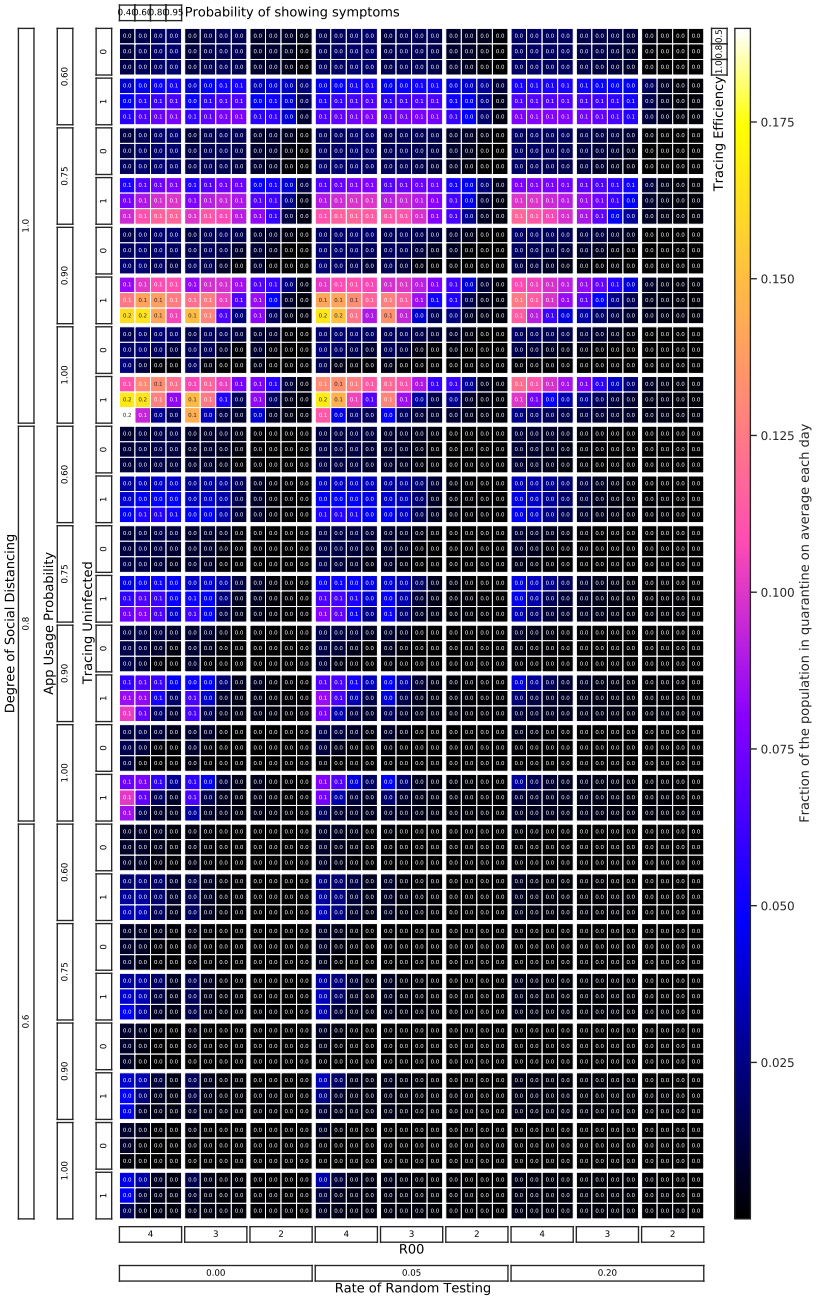


Figure 24: The fraction of the population quarantined on average over one year of the outbreak is shown for all combinations of parameters in the default parameter scan (black font color in Tab. 1).

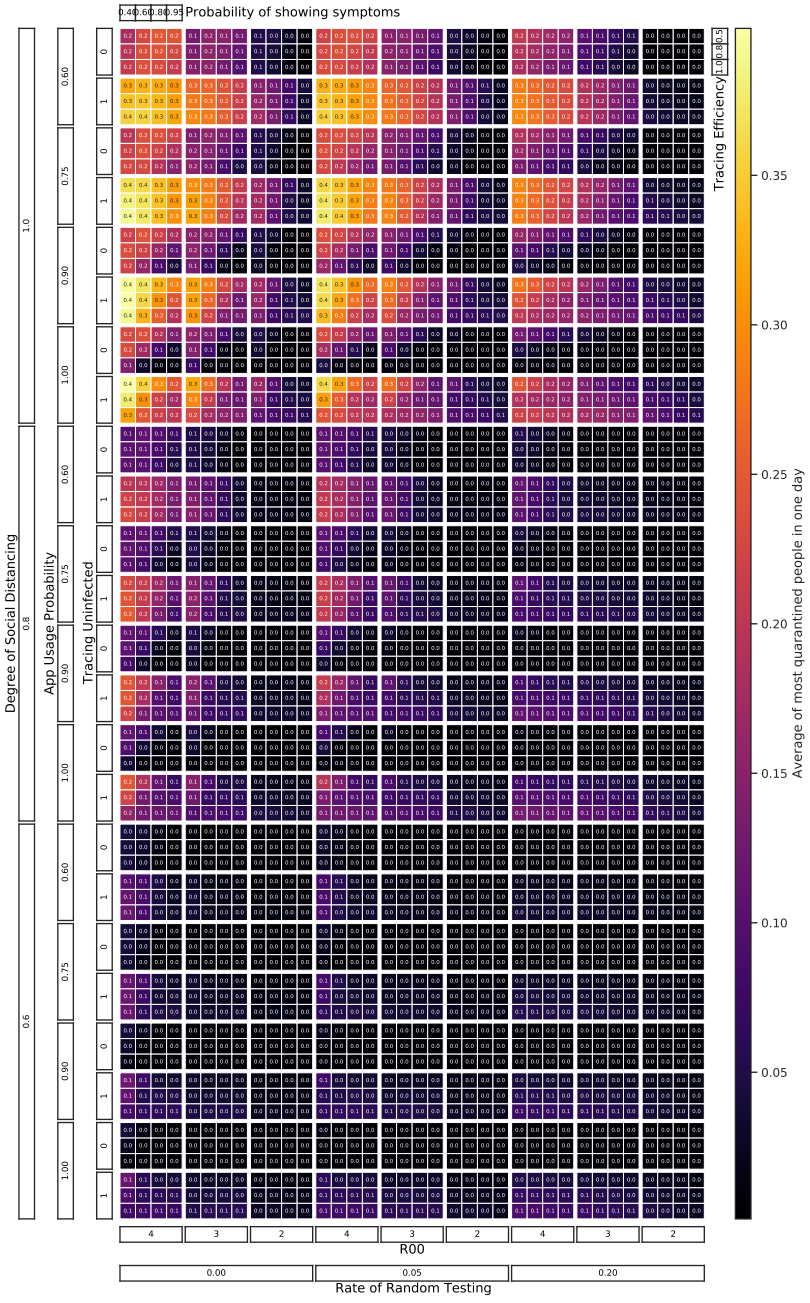


Figure 25: The fraction of the population that is quarantined on the day when most people are quarantined is shown for all combinations of parameters in the default parameter scan (black font color in Tab. 1).

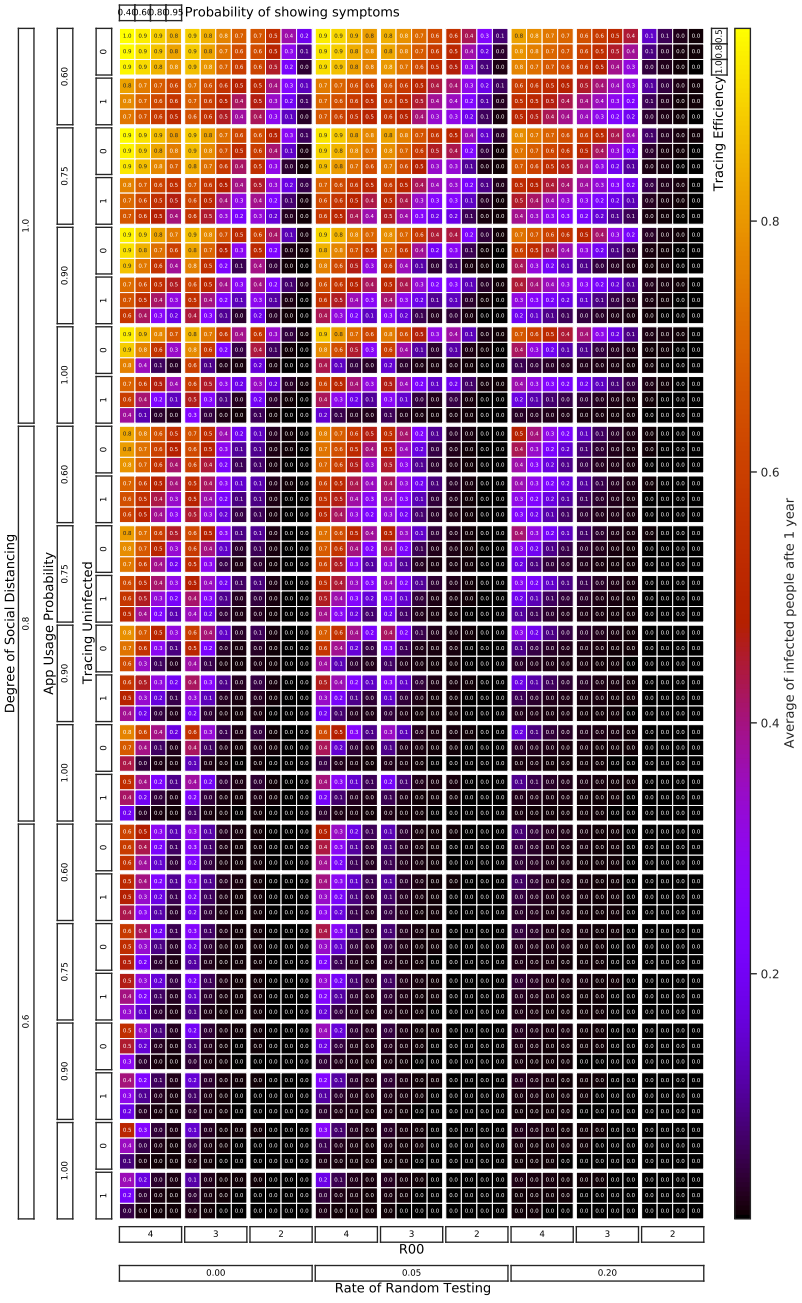


Figure 26: The total fraction of the population that has been exposed one year into the outbreak is shown for all combinations of parameters in the default parameter scan (black font color in Tab. 1).

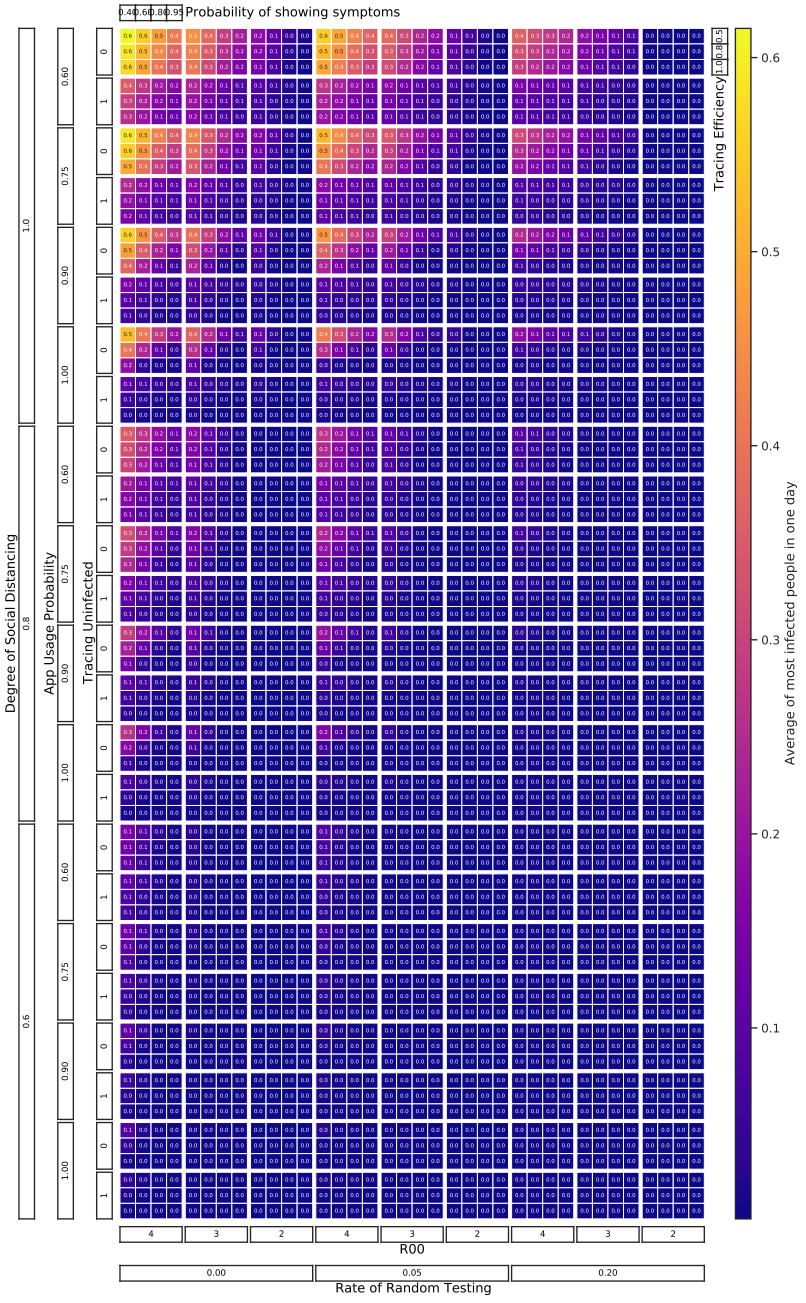


Figure 27: The fraction of the population that is sick on the day when most people are sick is shown for all combinations of parameters in the default parameter scan (black font color in Tab. 1).

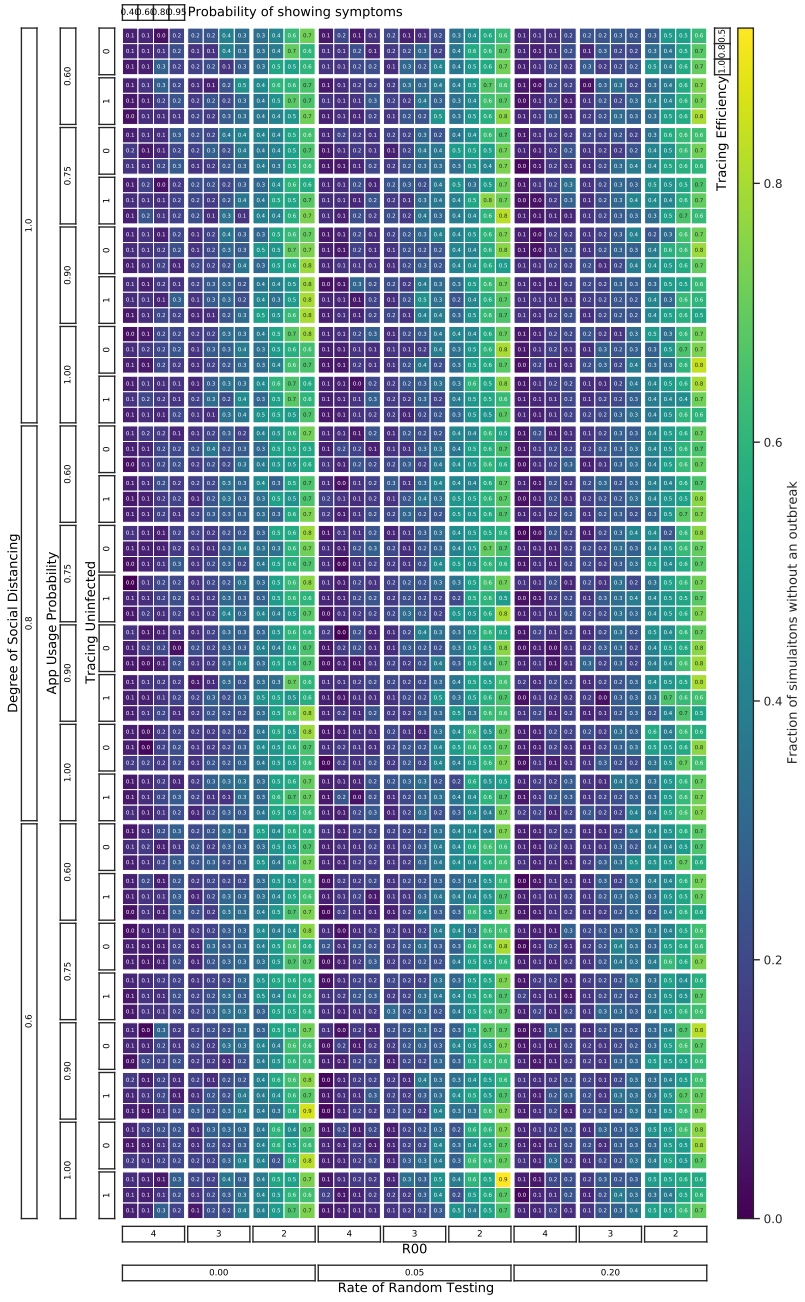


Figure 28: The fraction of MC runs where the outbreak stopped (by chance) before the threshold number of infected people to start interventions was reached.

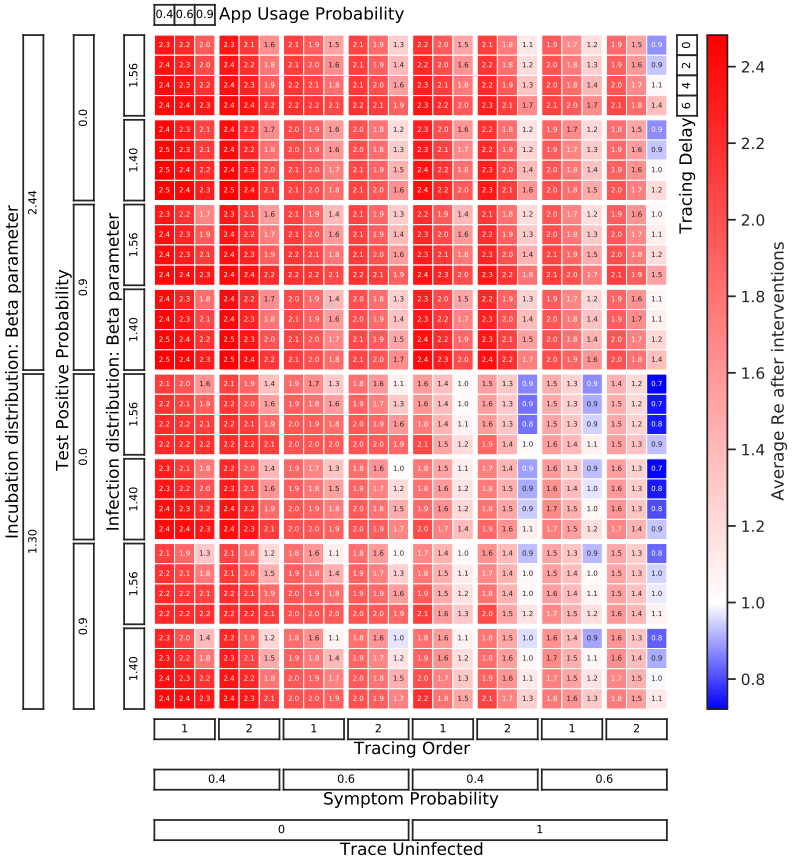


Figure 29: The effective reproductive number reached once interventions are turned on is shown for all combinations of parameters in the timing and delay parameter scan

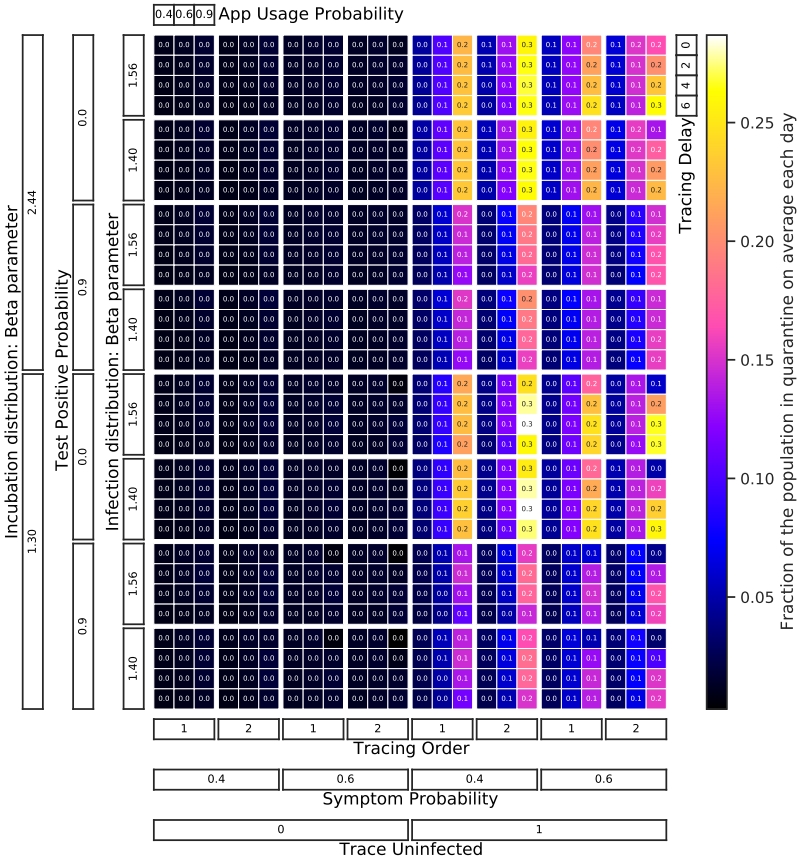


Figure 30: The fraction of the population quarantined on average over one year of the outbreak is shown for all combinations of parameters in the timing and delay parameter scan.

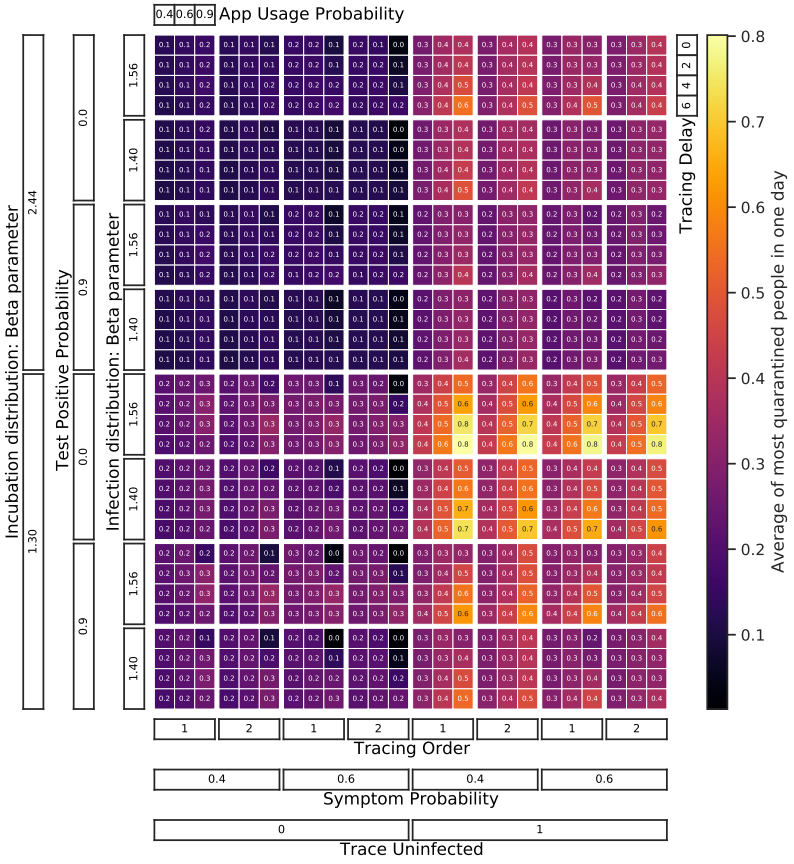


Figure 31: The fraction of the population that is quarantined on the day when most people are quarantined is shown for all combinations of parameters in the timing and delay parameter scan.

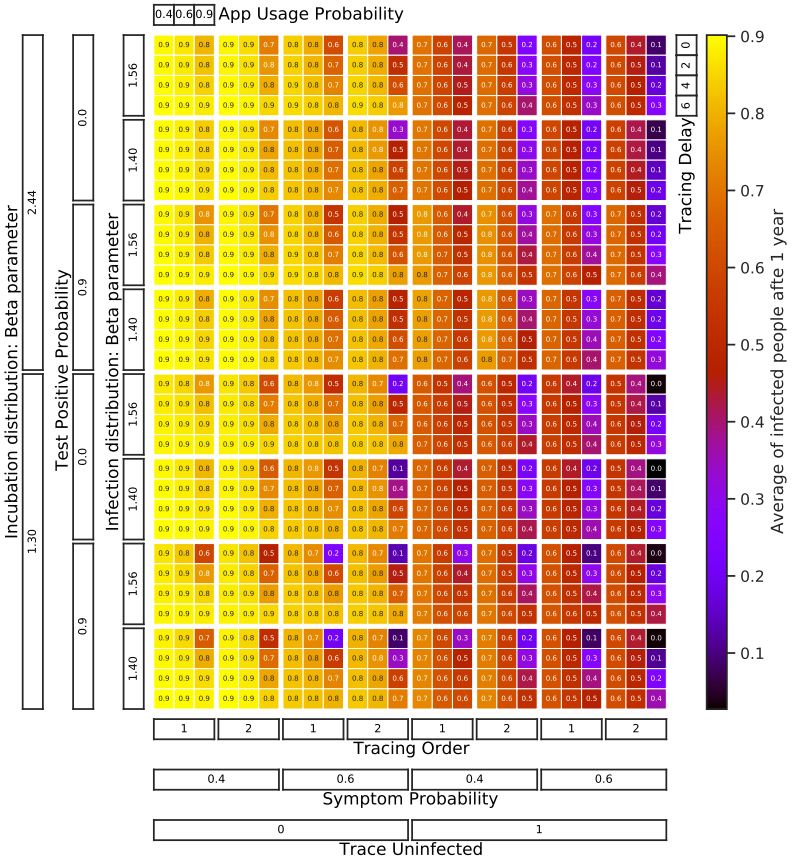


Figure 32: The total fraction of the population that has been exposed one year into the outbreak is shown for all combinations of parameters in the timing and delay parameter scan.

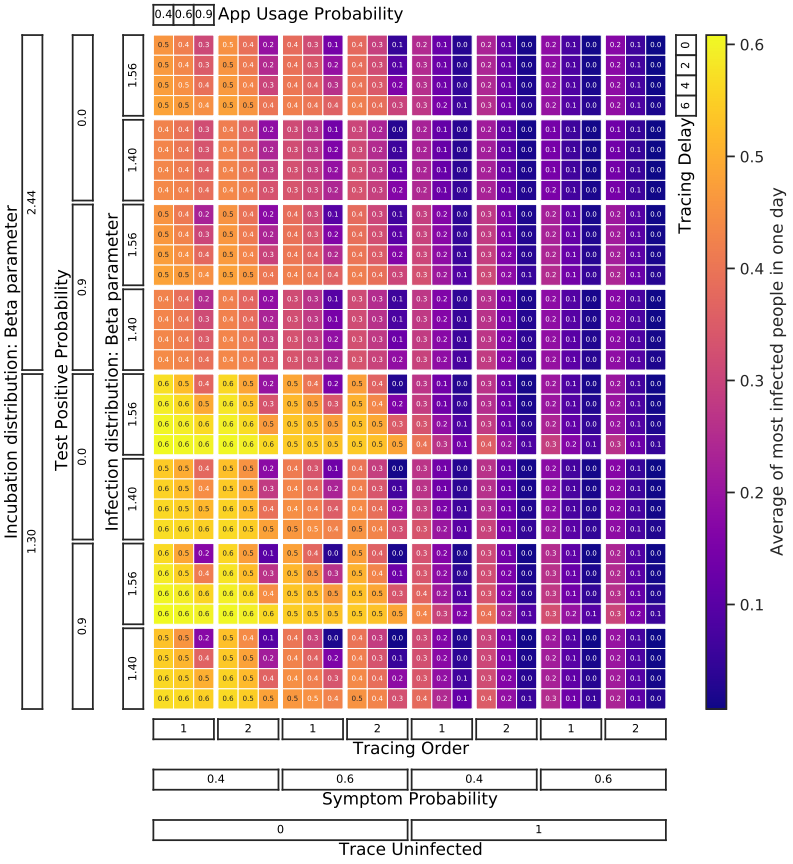


Figure 33: The fraction of the population that is sick on the day when most people are sick is shown for all combinations of parameter in the timing and delay parameter scan.

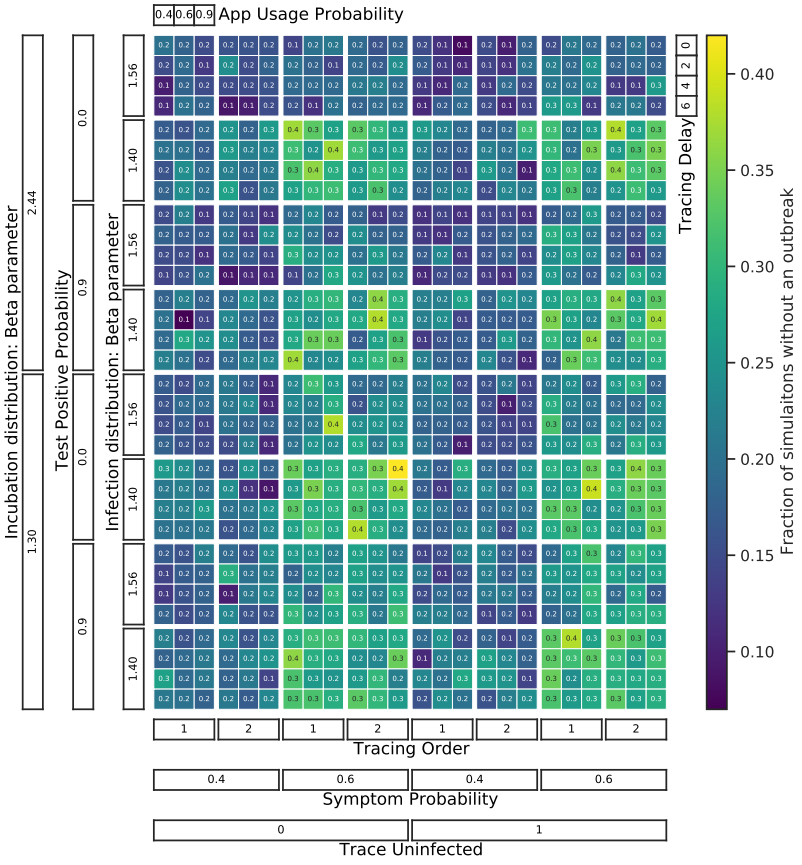


Figure 34: The fraction of MC runs where the outbreak stopped (by chance) before the threshold number of infected people to start interventions was reached (timing and delay parameter scan).