COVID-19 INDIVIDUAL-BASED MODEL WITH INSTANTANEOUS CONTRACT TRACING

FROM THE PATHOGEN DYNAMICS GROUP, BIG DATA INSTITUTE, NUFFIELD DEPARTMENT OF MEDICINE, UNIVERSITY OF OXFORD (ROBERT HINCH, WILL PROBERT, ANEL NURTAY, MICHELLE KENDALL, CHRIS WYMANT, MATTHEW HALL, KATRINA LYTHGOE, ANA BULAS CRUZ, LELE ZHAO, ANDREA STEWART, LUCA FERRETTI, LUCIE ABELER-DÖRNER, DAVID BONSALL, CHRISTOPHE FRASER); IN COLLABORATION WITH ANTHONY FINKELSTEIN; IBM UK; FACULTY.AI

Contents

Values of model parameters reported in this document match those in a separate master spreadsheet on April 9, 2020. In the event of mismatches at later dates, the spreadsheet values reflect updated values.

1. Overview

Our individual-based model (IBM) was developed to simulate the spread of COVID-19 in a city, and to analyse the effect of both passive and active intervention strategies. Its primary purpose at this stage is to assist the design and evaluation of approaches to instantaneous contact tracing using a mobile phone app that measures proximity events between phones with the app. The IBM explores the effectiveness of such a contact-tracing app intervention by modelling five alternative scenarios which differ in how far the contact tracing extends through the contact network, and the conditions on which individuals are released from app-instructed quarantine.

In the absence of therapeutical interventions, a key intervention for preventing infectious disease spread is detecting cases and tracing their contacts, so that those at risk of infecting others can be isolated. An IBM is well-suited to quantifying the effects of combinations of non-pharmaceutical interventions in an epidemic because, unlike a simple epidemiological model, it records a history of previous events. The IBM was developed using the most relevant representation of the human contact network for a directly transmitted short-generation-time pathogen. It includes the three major domains of interaction: the home, the workplace (or school for children, or a regular social environment for older individuals), and the random interactions of daily life and travel. The IBM is age-stratified. It is not spatially stratified at this stage. The contact processes are currently parameterised based previous work interviewing participants; they will be updated based on contact data collected by phone. The model of infection spread via interactions between individuals is parameterised based on the current literature of COVID-19 epidemiology and can be updated as more data emerges. The disease process is currently modelled to ensure that the time-scales of disease progression are correct. Since the impact of COVID-19 on hospitals is large, and the clinical outcome of infection depends on access to good hospital care, a more detailed model of patient flows and transmission within hospitals is planned.

2. Demographics

The demographics of the IBM are based upon UK national data for 2018 from the Office of National Statistics (ONS). Individuals are categorised into nine age groups by decade, from age group (0-9 years) to (80+ years). The simulation is run on a static population and the proportion of individuals in each age group is the same as that specified by the population level statistics in Table 1. Every individual is part of a household, which forms an important part of their daily interactions. Household size data is shown in Table 1. Since the duration of the simulated epidemic is less than a year, we do not consider changes in the population due to births, deaths due to other causes, and migration.

3. Interaction Network

Every individual in the population is represented by a node in the simulation. Interactions between individuals are modelled via connections between nodes. The connections form networks, which are generated to represent different types of daily interactions. In this work we consider three types of networks that represent household, workplace, and miscellaneous interactions for the studied population. Some of these networks are static and recur daily (e.g. household), whilst others are transient and are regenerated daily (e.g. miscellaneous).

TABLE 1. Age-stratified population of the UK and number of households containing n people, with $n = 1, 2, \ldots 6$, provided by the ONS. Parameter values match the Fraser group IBM model parameters, April 9, 2020

Demographic Parameters		
Name	Description	Value
n_total	Total population simulated	100,000
population_0_9	population 0-9 years old	8,054,000
population_10_19	population 10-19 years old	7,528,000
population_20_29	population 20-29 years old	8,712,000
population_30_39	population 30-39 years old	8,835,000
population_40_49	population 40-49 years old	8,500,000
population_50_59	population 50-59 years old	8,968,000
population_60_69	population 60-69 years old	7,069,000
population_70_79	population 70-79 years old	5,488,000
population_80	population 80+ years old	3,281,000
household_size_1	households with 1 person (thousands)	7,452
household_size_2	households with 2 people (thousands)	9,936
household_size_3	households with 3 people (thousands)	4,416
household_size_4	households with 4 people (thousands)	4,140
household_size_5	households with 5 people (thousands)	1,104
household_size_6	households with 6 people (thousands)	552

TABLE 2. Average number of interactions for an individual in each age group acquired from empirical estimates (12). Parameter values match the Fraser group IBM model parameters, April 9, 2020

Mean daily interactions	
Age group	Value
children (0-19 years)	12 + household interactions
adults (20-69 years)	11 + household interactions
elderly (70+ years)	6 + household interactions

The interaction networks have two roles in the IBM. First, they model the transmission of infection between individuals on each day that a connection is made. Second, when we model contact tracing via the phone app, the phone sees a subset of the network which is then used for contact tracing.

The membership of different networks leads to age-group assortativity in the interactions. A previous study of social contacts for infectious disease modelling has estimated the mean number of interactions that individuals have by age group (12). This study is based on participants being asked to recall their interactions over the past day. We estimate mean interactions by age group by aggregating data (see Table 2). It is possible for an individual not to be connected to anybody (*e.g.* a person living alone in a household).

3.1. **Household Network.** Each individual is assigned to live in a single household. Each day every person has an interaction with everybody within their household. The proportion of people living in households of different sizes is taken from ONS data (see Demographics). There are two important population-level aggregate statistics that we match: the household-size structure and the population age-structure. When individuals are assigned to households, we need to match these aggregate statistics and also require that our households reflect typical age mixes. To achieve this we use a reference panel 10,000 households containing the household composition by age, produced by down-sampling UK-wide household composition data from the 2011 Census produced by the ONS. The household structure was then generated by sampling from the reference household panel with replacement. To match the aggregate statistics for the population, we used a rejection-sampling method. Households were sampled sequentially from the reference panel and accepted if, when combined with the existing samples, the discrepancy with the target aggregate statistics was reduced. If the discrepancy with the target aggregate statistics increased by less than a threshold, the new sample was retained, otherwise it was rejected. The threshold changes dynamically with each sample depending on whether it was accepted or rejected to keep a constant rejection rate throughout the sampling. The discrepancy between the sample aggregate statistics and the target aggregate statistics was calculated by taking the sum of the square of the differences. Finally, we checked the discrepancy of the total sampled set of households against the aggregate statistics and rejected the whole sample if it was above a threshold.

3.2. **Workplace Networks.** Each individual is a member of one workplace network (including e.g. schools for children and social activities for older adults). The workplace networks are modelled as Watts-Strogatz small-world networks (14). When constructing the workplace networks, we ensure the absence of overlaps between the household interactions and the local interactions on the small-world network. Every day each person interacts with a random subset (50%) of their connections on their workplace network.

For children, there are separate workplace networks for the both the 0y-9y age group (i.e. primary schools) and 10y-19y age groups (i.e secondary schools). On each of these networks we introduce a small number of adults (1 adult per 5 children) to represent teaching and other school staff. Similarly for the 70y-79y age group and the 80y+ age group we have separate networks representing day-time social activities among elderly people (again with 1 adult per 5 elderly people). All remaining adults (the vast majority) are part of the 20y-69y network. Due to the difference in total number of daily interactions, each age group has a different number of interactions in their workplace network. Parameters and values corresponding to the workplace network are shown in Table 3. Note the parameters are for the mean daily connections, so the number of interactions on the network are higher due to the daily sampling of connections.

3.3. **Random Network.** In addition to the recurring structured networks of households and workplaces, we include a number of random interactions as well. These interactions occur once and are independent of the previous day's connections. The number of random connections an individual makes is the same each day (without interventions) and is drawn at the start of the simulation from a negative-binomial distribution, an over-dispersed skew distribution. This variation in the number of interactions introduces "super-spreaders" into the network who have many more interactions than average.

TABLE 3. Mean numbers of daily connections for members of each age group, fraction of adults involved in occupational networks for children and for elderly people, and rewiring parameters for randomisation of daily interactions (12). Parameter values match the Fraser group IBM model parameters, April 9, 2020

Workplace Network Parameters		
Name	Description	Value
mean_work_interaction_child	mean daily connections on 0y-9y and 10y-19y networks	10
mean_work_interaction_adult	mean daily connections on the 20y-69y network	7
mean_work_interaction_elderly	mean daily connections on the 70y-79y and 80y+ networks	3
child_network_adults	fraction of adults in child network	0.2
elderly_network_adults	fraction of adults in elderly network	0.2
daily_fraction_work	fraction of daily work connections made	0.5
prob_network_rewire	rewiring probability of the Watts-Stogatz small-world network	0.1

TABLE 4. Parameters for numbers of random connections that members of each age group have per day (12). Parameter values match the Fraser group IBM model parameters, April 9, 2020

Random Network Parameters		
Name	Description	Value
mean_random_interaction_child	mean number of connections for children (0y-19y)	2
mean_random_interaction_adult	mean number of connections for adults (20y-69y)	4
mean_random_interaction_elderly	mean number of connections for elderly (70y+)	3
sd_random_interaction_child	s.d. number of connections for children (0y-19y)	2
sd_random_interaction_adult	s.d. number of connections for adults (20y-69y)	4
sd_random_interaction_elderly	s.d. number of connections for elderly (70y+)	3

The mean numbers of connections were chosen so that the total number of daily interactions matched that from a previous study of social interaction (12). The number of random interactions was chosen to be lower in children in comparison to other age groups. In the simulation, each day a list is created containing all individuals who make random interactions and each person is repeated by the number of interactions they make. This list is shuffled and interactions are made between adjacent pairs on the shuffled list.

4. Infection Dynamics

The infection is spread by interactions between infected and susceptible individuals. The rate of transmission is determined by three factors: the status of the infector; the susceptibility of the infectee to infection, according to age; and the type of interaction (i.e. on which network it occurred). We present the details of each of these rates below. We currently do not have data on the distribution of the duration of interactions, so the effect of this on transmission is not modelled.

To model **the status of the infector** we note that infectiousness varies over the natural course of an infection, i.e. as a function of the amount of time a person has been infected, τ. Infectiousness starts at zero at the precise moment someone is infected ($\tau = 0$), reaches a peak at some intermediate time, and tends to zero a long time after infection (large

τ). Following (6), we take the functional form of infectiousness to be a (scaled) gamma distribution. We chose the mean and standard deviation as intermediate values between different reports (6; 7; 10).

We define *asymptomatic* individuals as those who never develop symptoms during their infection, and *pre-symptomatic* individuals as those who do not currently have symptoms but will do so later. Both types of individuals may infect others. The overall infectiousness of asymptomatic individuals has been estimated to be lower (9), reflected in our parameter `asymptomatic_infectious_factor`. Individuals who show symptoms but only mildly have also been estimated as less infectious than those with more severe symptoms (9), reflected in our parameter `mild_infectious_factor`. For each adult age category, we took the fraction of symptomatic infections that are mild as the fraction of confirmed cases with with either no pneumonia (which was rare – always less than 6.5%) or mild pneumonia, i.e. the fraction without severe pneumonia, reported in (16). For the age categories 0-9 and 10-19, we took the fraction of infections that are mild as the fraction clinically defined as 'mild' in the paediatric meta-analysis (3), excluding asymptomatic infections from the denominator. These fractions of symptomatic infections that are mild by age are listed in Table 7.

To model the **susceptibility to infection of a contact according to their age** we referred to the literature (9; 17; 1) where close contacts of confirmed cases were monitored and tested. The number tested and the number of positive results was reported within each age group, with the ratio of the latter to the former defining the per-age attack rate. We assumed the proportion testing positive was constant within each age bin in each study, and merged the counts into a single set of bins 0-9, 10-19, ... 80+, with contributions from all three studies. The largest age bin in each study (which differed in its lower bound, with no upper bound) was assumed to be ten years wide in order to define the merging; the counts in the resulting 80+ category were excluded for the following fitting step as they were sensitive to the assumed upper bound in each study. We fit the polynomial form $A + B \times (\text{age})^k$ to the proportions in each age category, using the bin's midpoint, minimising the sum of squared differences from the observed values. The best fitting values of A, B and k were 0.0300, 5.53×10^{-6}, and 2.00 respectively. We took the values predicted by this fit at each age bin's mid-point to define the relative susceptibility of each age group. A final normalisation factor was defined so that the average susceptibility for an individual in a population with the age distribution considered here was 1, defining our parameters `relative_susceptibility_0_9` etc. The merged data and fit are shown in Table 5 and Fig. 1.

Finally, we model the **type of interaction**, i.e. on which network it took place. Whilst we do not have data on the length of interactions, interactions which take place within a person's home are likely to be closer than other types of interactions leading to higher rates of transmission. This is modelled using a multiplicative factor.

Combining all effects, we model the rate at which the virus is transmitted in a single interaction by

$$\lambda(t, s_i, a_s, n) = \frac{R S_{a_s} A_{s_i} B_n}{\bar{I}_{a_s}} \int_{t-1}^{t} f_\Gamma(u; \mu_i, \sigma_i^2) \mathrm{d}u, \tag{1}$$

where t is the time since infection; s_i indicates the infector's symptom status (asymtomatic, mild, moderate/severe); a_s is the age of the susceptible; n is the type of network where the

TABLE 5. Steps in our calculation of susceptibility by age. The attack rate – the fraction of close contacts of a confirmed case infected – was merged from references (9; 17; 1). We fit to these values, and then scale them all identically to give a normalised susceptibility to infection. Parameter values match the Fraser group IBM model parameters, April 9, 2020

age group	number of contacts infected	number of contacts	attack rate	attack rate predicted from fit	normalised susceptibility
0-9	25.7	788	0.0326	0.0301	0.71
10-19	27.9	927	0.0301	0.0312	0.74
20-29	47.5	1422	0.0334	0.0334	0.79
30-39	56.5	1576	0.0358	0.0368	0.87
40-49	47.7	1350	0.0353	0.0412	0.98
50-59	50.9	1217	0.0419	0.0467	1.11
60-69	54.5	824	0.0662	0.0533	1.26
70-79	31.6	562	0.0563	0.0610	1.45
80+	3.64	125	0.0292 (excluded)	0.0699	1.66

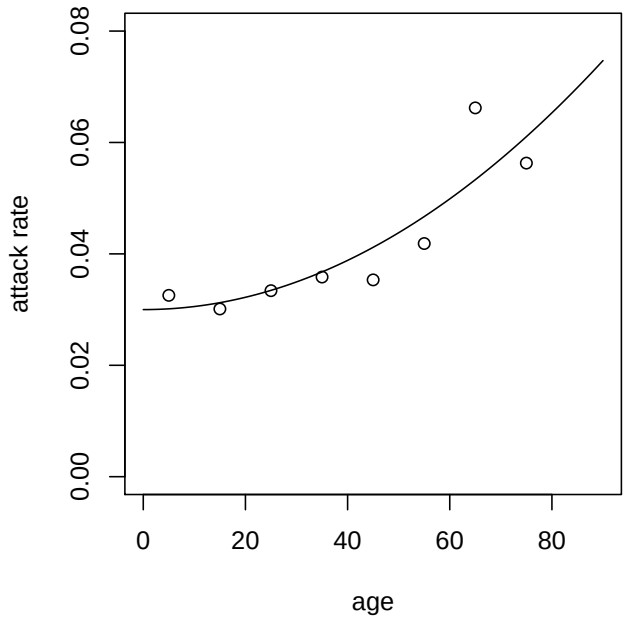

FIGURE 1. The attack rate, defined as the fraction of close contacts of a confirmed case infected, as a function of the contacts' age. Open circles show data merged from references (9; 17; 1); the line shows the fit $A + B \times (\text{age})^k$.

interaction occurred; $\bar{I}_{a_s}$ is the mean number of daily interactions for somebody of the age of the susceptible; $f_\Gamma(u; \mu, \sigma^2)$ is the probability density function of a gamma distribution; μ_i and σ_i are the mean and width of the infectiousness curve; R scales the overall infection rate

TABLE 6. Description of infection parameters and their values. The infection rate$^{\times}$ was determined by fitting the simulation to have an epidemic doubling time of 3 days. Relative susceptibility values* were derived by merging and fitting to data from sources (9; 17; 1) as explained in the text. Parameter values match the Fraser group IBM model parameters, April 9, 2020

Infection Parameters			
Symbol	Name	Description	Value
R	infectious_rate	mean number of people infected by each moderately/severely symptomatic individual	$6.75^{\times}$
A_{asym}	asymptomatic_infectious_factor	relative infection rate of asymptomatics	0.29
A_{mild}	mild_infectious_factor	relative infection rate of mild symptomatic	0.48
A_{sym}	-	relative infection rate of moderate/severe symptomatics	1
B_{home}	relative_transmission_household	relative infection rate of household interaction	2
B_{work}	relative_transmission_workplace	relative infection rate of work interaction	1
B_{random}	relative_transmission_random	relative infection rate of random interaction	1
S_{0-9}	relative_susceptibility_0_9	relative susceptibility of age group 0 - 9 to an average person	0.71^{*}
S_{10-19}	relative_susceptibility_10_19	relative susceptibility of age group 10 - 19 to an average person	0.74^{*}
S_{20-29}	relative_susceptibility_20_29	relative susceptibility of age group 20 - 29 to an average person	0.79^{*}
S_{30-39}	relative_susceptibility_30_39	relative susceptibility of age group 30 - 39 to an average person	0.87^{*}
S_{40-49}	relative_susceptibility_40_49	relative susceptibility of age group 40 - 49 to an average person	0.98^{*}
S_{50-59}	relative_susceptibility_50_59	relative susceptibility of age group 50 - 59 to an average person	1.11^{*}
S_{60-69}	relative_susceptibility_60_69	relative susceptibility of age group 60 - 69 to an average person	1.26^{*}
S_{70-79}	relative_susceptibility_70_79	relative susceptibility of age group 70 - 79 to an average person	1.45^{*}
S_{80}	relative_susceptibility_80	relative susceptibility of age group 80+ to an average person	1.66^{*}
μ_i	mean_infectious_period	the mean of the gamma probability density function for infectiousness	6 days
σ_i	sd_infectious_period	the standard deviation of the gamma probability density function for infectiousness	2.5 days
-	n_seed_infection	number of individual randomly infected at start of simulation	5

(under some simplifying assumptions it is mean number of people infected by each moderately/severely symptomatic individual); S_{a_s} is the scale-factor for the age of the susceptible; A_{s_i} is the scale-factor for the infector being asymptomatic; B_n is the scale-factor for the network on which the interaction occurred. Table 6 contains the values of the parameters used in simulations. The rate of virus transmission is converted to a probability of transmission

$$P(t, s_i, a_s, n) = 1 - e^{-\lambda(t, s_i, a_s, n)}. \tag{2}$$

The epidemic is seeded by infecting a number of individuals at the start of the infection. The infection was assumed to take place immediately before the simulation starts.

5. Natural history of infection

Upon infection, an individual enters a disease progression cascade where the outcome and rates of progression depend on the age of the infected person. The disease state transitions are shown in Figure 2 and the model parameters are in the table Disease Dynamics Parameters.

A fraction $\phi_{\text{asym}}(\text{age})$ of individuals are asymptomatic and do not develop symptoms, a fraction $\phi_{\text{mild}}(\text{age})$ will eventually develop mild symptoms, and the remainder develop moderate/severe symptoms. Each of these proportions depend on the age of the infected individual. Those who are asymptomatic are infectious (at a lower level, see Infection Dynamics section) and will move to a recovered state after a time $\tau_{\text{a,rec}}$ drawn from a gamma distribution. Once an individual is recovered we assume that they have immunity and cannot be reinfected.

Individuals who will develop symptoms start by being in a pre-symptomatic state, in which they are infectious but have no symptoms. The pre-symptomatic state is important for modelling interventions because individuals in this state do not realise they are infectious, therefore they cannot self-isolate to prevent infecting others. Individuals who develop mild

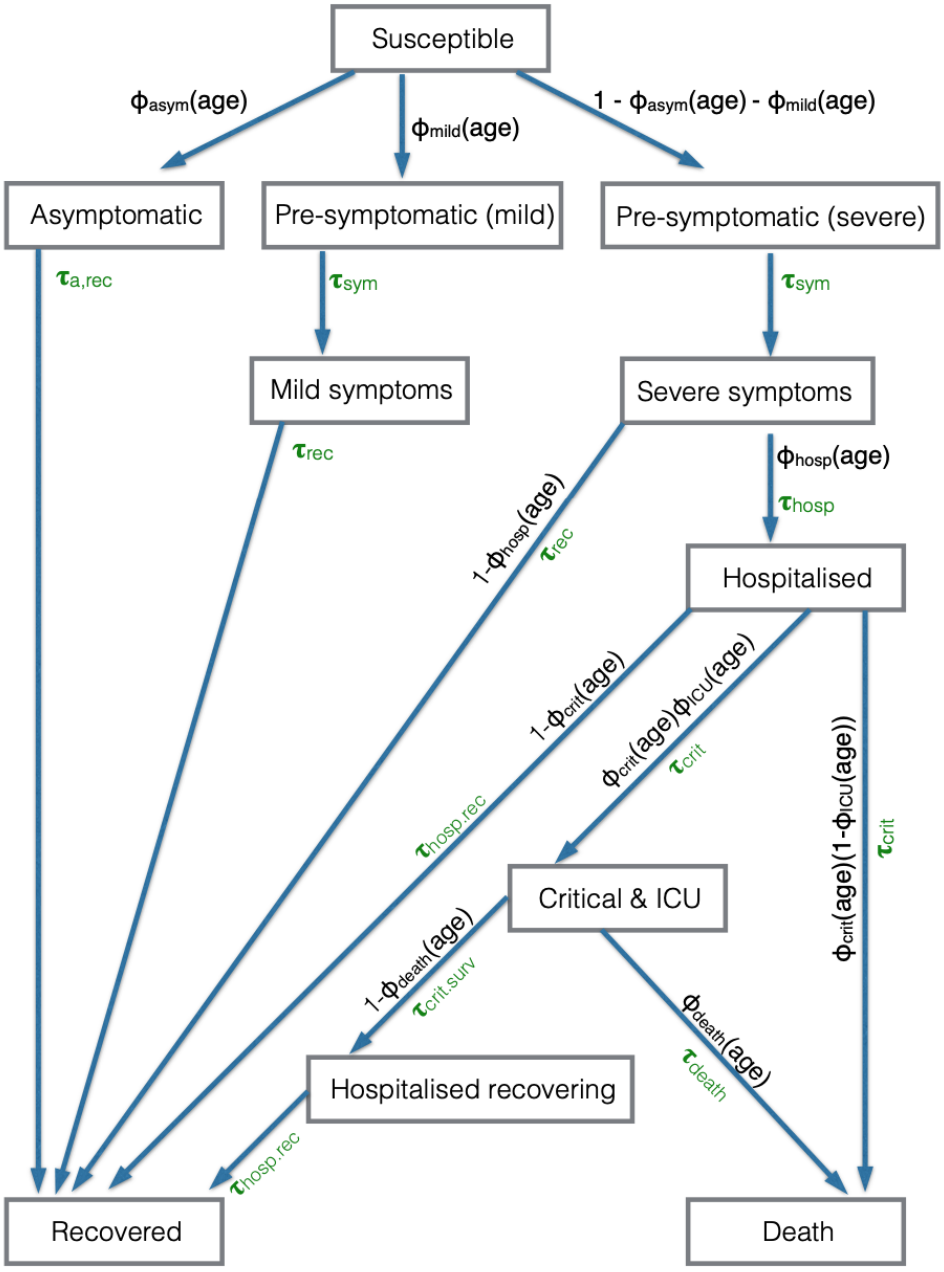

FIGURE 2. The disease status of an individual and the probability and time distribution of transitions. The ϕ_{xxx}(age) variables are the probability of transition to a particular state when there is a choice, where the probability depends upon the age of the individual. The τ_{xxx} are the gamma distributed variables of the time taken to make the transition.

symptoms do so after time τ_{sym} and then recover after time τ_{rec} (both drawn from gamma distributions). The rest of the individuals develop moderate/severe symptoms after a time τ_{sym} drawn from the gamma distributed.

Whilst most individuals recover without requiring hospitalisation, a fraction $\phi_{\text{hosp}}(\text{age})$ of those with moderate/severe symptoms will require hospitalisation. This fraction is age-dependent. Those who do not require hospitalisation recover after a time τ_{rec} drawn from a gamma distribution, whilst those who require hospitalisation are admitted to hospital after a time τ_{hosp}, which is drawn from a shifted Bernoulli distribution (either 5 or 6 days).

Among all hospitalised individuals, a fraction $\phi_{\text{crit}}(\text{age})$ develop critical symptoms and require intensive care treatment, with the remainder recovering after a time $\tau_{\text{hosp,rec}}$ drawn from a gamma distribution. The time from hospitalisation to developing critical symptoms, τ_{crit}, is drawn from a shifted Bernoulli distribution (either 2 or 3 days). Of those who develop critical symptoms, a fraction $\phi_{\text{ICU}}(\text{age})$ will receive intensive care treatment (note for most age groups this will be everybody, however, for the most elderly age groups it may be deemed inappropriate). For patients receiving intensive care treatment, a fraction $\phi_{\text{death}}(\text{age})$ die after a time τ_{death} drawn from a gamma distribution, with the remainder leaving intensive care after a time $\tau_{\text{crit,surv}}$. Patients who require critical care and do not receive intensive care treatment are assumed to die upon developing critical symptoms. Patients who survive critical symptoms remain in hospital for $\tau_{\text{hosp,rec}}$ before recovering.

6. Passive Interventions

The IBM has the ability to model both passive and active non-pharmaceutical interventions. Interventions are designed to reduce the rate of transmission but have the potential to quarantine significant numbers of people. We define passive interventions to be those which do not involve testing or contact tracing. Here we provide details on each of the passive interventions and their impact on the contact network.

6.1. **Hospitalisation.** Upon hospitalisation, a patient immediately stops interacting with both their household and workplace networks. We also reduce the number of random interactions that they have. We do not currently model the interactions within hospitals (this is planned for future work).

6.2. **Self-quarantine upon symptoms.** Upon experiencing symptoms, a proportion of patients in our model self-quarantine immediately as per government advice. There is a daily dropout rate for the duration of the quarantine to model individuals failing to comply. We also allow for the development of symptoms which cause an individual to self-quarantine when in fact they do not have covid-19, for example they may have seasonal flu which has similar symptoms. Quarantine is modelled by stopping interactions on the individual's workplace network and greatly reducing their number of interactions on the random network. The IBM also contains the option that everybody in the household of a person with symptoms will be asked to self-quarantine.

6.3. **Lock-down.** The IBM can simulate the effect of a country-wide lock-down by reducing the number of contact that people have. In a lock-down, we reduce the number of interactions that people have by 80% on both their work-place and random networks. Additionally, given that during a lock-down people stay at home more, we increase transmission rate for interactions on the household network by a factor.

TABLE 7. Proportion of people in each stage of illness whose disease progresses further; mean and standard deviation for density functions of the times that each transition – disease progression or recovery – takes. Parameter values match the Fraser group IBM model parameters, April 9, 2020

Disease Dynamics Parameters				
Symbol	Name	Description	Value	Source
$\phi_{\text{asym}}(0-9)$	fraction_asymptomatic_0_9	fraction of infected who are asymptomatic age group 0 - 9	0.18	(11)
$\phi_{\text{asym}}(10-19)$	fraction_asymptomatic_10_19	fraction of infected who are asymptomatic age group 10 - 19	0.18	(11)
$\phi_{\text{asym}}(20-29)$	fraction_asymptomatic_20_29	fraction of infected who are asymptomatic age group 20 - 29	0.18	(11)
$\phi_{\text{asym}}(30-39)$	fraction_asymptomatic_30_39	fraction of infected who are asymptomatic age group 30 - 39	0.18	(11)
$\phi_{\text{asym}}(40-49)$	fraction_asymptomatic_40_49	fraction of infected who are asymptomatic age group 40 - 49	0.18	(11)
$\phi_{\text{asym}}(50-59)$	fraction_asymptomatic_50_59	fraction of infected who are asymptomatic age group 50 - 59	0.18	(11)
$\phi_{\text{asym}}(60-69)$	fraction_asymptomatic_60_69	fraction of infected who are asymptomatic age group 60 - 69	0.18	(11)
$\phi_{\text{asym}}(70-79)$	fraction_asymptomatic_70_99	fraction of infected who are asymptomatic age group 70 - 79	0.18	(11)
$\phi_{\text{asym}}(80)$	fraction_asymptomatic_80	fraction of infected who are asymptomatic age group 80+	0.18	(11)
$\phi_{\text{mild}}(0-9)$	mild_fraction_0_9	fraction of infected who will develop mild symptoms in age group 0 - 9	0.79	(3)
$\phi_{\text{mild}}(10-19)$	mild_fraction_10_19	fraction of infected who will develop mild symptoms in age group 10 - 19	0.79	(3)
$\phi_{\text{mild}}(20-29)$	mild_fraction_20_29	fraction of infected who will develop mild symptoms in age group 20 - 29	0.73	(3)
$\phi_{\text{mild}}(30-39)$	mild_fraction_30_39	fraction of infected who will develop mild symptoms in age group 30 - 39	0.68	(3)
$\phi_{\text{mild}}(40-49)$	mild_fraction_40_49	fraction of infected who will develop mild symptoms in age group 40 - 49	0.65	(3)
$\phi_{\text{mild}}(50-59)$	mild_fraction_50_59	fraction of infected who will develop mild symptoms in age group 50 - 59	0.59	(3)
$\phi_{\text{mild}}(60-69)$	mild_fraction_60_69	fraction of infected who will develop mild symptoms in age group 60 - 69	0.53	(3)
$\phi_{\text{mild}}(70-79)$	mild_fraction_70_79	fraction of infected who will develop mild symptoms in age group 70 - 79	0.41	(3)
$\phi_{\text{mild}}(80)$	mild_fraction_80	fraction of infected who will develop mild symptoms in age group 80+	0.27	(3)
$\phi_{\text{hosp}}(0-9)$	hospitalised_fraction_0_9	fraction of hospitalisations among symptomatic age group 0 - 9	0.001	(5)
$\phi_{\text{hosp}}(10-19)$	hospitalised_fraction_10_19	fraction of hospitalisations among symptomatic age group 10 - 19	0.003	(5)
$\phi_{\text{hosp}}(20-29)$	hospitalised_fraction_20_29	fraction of hospitalisations among symptomatic age group 20 - 29	0.012	(5)
$\phi_{\text{hosp}}(30-39)$	hospitalised_fraction_30_39	fraction of hospitalisations among symptomatic age group 30 - 39	0.032	(5)
$\phi_{\text{hosp}}(40-49)$	hospitalised_fraction_40_49	fraction of hospitalisations among symptomatic age group 40 - 49	0.049	(5)
$\phi_{\text{hosp}}(50-59)$	hospitalised_fraction_50_59	fraction of hospitalisations among symptomatic age group 50 - 59	0.102	(5)
$\phi_{\text{hosp}}(60-69)$	hospitalised_fraction_60_69	fraction of hospitalisations among symptomatic age group 60 - 69	0.166	(5)
$\phi_{\text{hosp}}(70-79)$	hospitalised_fraction_70_79	fraction of hospitalisations among symptomatic age group 70 - 79	0.243	(5)
$\phi_{\text{hosp}}(80)$	hospitalised_fraction_80	fraction of hospitalisations among symptomatic age group 80+	0.273	(5)
$\phi_{\text{crit}}(0-9)$	critical_fraction_0_9	fraction of hospitalised people in age group 0 - 9 in critical condition	0.05	(5)
$\phi_{\text{crit}}(10-19)$	critical_fraction_10_19	fraction of hospitalised people in age group 10 - 19 in critical condition	0.05	(5)
$\phi_{\text{crit}}(20-29)$	critical_fraction_20_29	fraction of hospitalised people in age group 20 - 29 in critical condition	0.05	(5)
$\phi_{\text{crit}}(30-39)$	critical_fraction_30_39	fraction of hospitalised people in age group 30 - 39 in critical condition	0.05	(5)
$\phi_{\text{crit}}(40-49)$	critical_fraction_40_49	fraction of hospitalised people in age group 40 - 49 in critical condition	0.063	(5)
$\phi_{\text{crit}}(50-59)$	critical_fraction_50_59	fraction of hospitalised people in age group 50 - 59 in critical condition	0.122	(5)
$\phi_{\text{crit}}(60-69)$	critical_fraction_60_69	fraction of hospitalised people in age group 60 - 69 in critical condition	0.274	(5)
$\phi_{\text{crit}}(70-79)$	critical_fraction_70_79	fraction of hospitalised people in age group 70 - 79 in critical condition	0.432	(5)
$\phi_{\text{crit}}(80)$	fatality_fraction_80	fraction of hospitalised people in age group 80+ in critical condition	0.709	(5)
$\phi_{\text{death}}(0-9)$	fatality_fraction_0_9	fraction of fatalities among people in critical condition in age group 0 - 9	0.33	(8; 4)
$\phi_{\text{death}}(10-19)$	fatality_fraction_10_19	fraction of fatalities among people in critical condition in age group 10 - 19	0.25	(8; 4)
$\phi_{\text{death}}(20-29)$	fatality_fraction_20_29	fraction of fatalities among people in critical condition in age group 20 - 29	0.5	(5)
$\phi_{\text{death}}(30-39)$	fatality_fraction_30_39	fraction of fatalities among people in critical condition in age group 30 - 39	0.5	(5)
$\phi_{\text{death}}(40-49)$	fatality_fraction_40_49	fraction of fatalities among people in critical condition in age group 40 - 49	0.5	(15)
$\phi_{\text{death}}(50-59)$	fatality_fraction_50_59	fraction of fatalities among people in critical condition in age group 50 - 59	0.69	(15)
$\phi_{\text{death}}(60-69)$	fatality_fraction_60_69	fraction of fatalities among people in critical condition in age group 60 - 69	0.65	(15)
$\phi_{\text{death}}(70-79)$	fatality_fraction_70_79	fraction of fatalities among people in critical condition in age group 70 - 79	0.88	(15)
$\phi_{\text{death}}(80)$	fatality_fraction_80	fraction of fatalities among people in critical condition in age group 80+	1	(15)
$\phi_{\text{ICU}}(0-9)$	icu_allocation_0_9	fraction receiving ICU treatment if required in age group 0 - 9	1	
$\phi_{\text{ICU}}(10-19)$	icu_allocation_10_19	fraction receiving ICU treatment if required in age group 10 - 19	1	
$\phi_{\text{ICU}}(20-29)$	icu_allocation_20_29	fraction receiving ICU treatment if required in age group 20 - 29	1	
$\phi_{\text{ICU}}(30-39)$	icu_allocation_30_39	fraction receiving ICU treatment if required in age group 30 - 39	1	
$\phi_{\text{ICU}}(40-49)$	icu_allocation_40_49	fraction receiving ICU treatment if required in age group 40 - 49	1	
$\phi_{\text{ICU}}(50-59)$	icu_allocation_50_59	fraction receiving ICU treatment if required in age group 50 - 59	1	
$\phi_{\text{ICU}}(60-69)$	icu_allocation_60_69	fraction receiving ICU treatment if required in age group 60 - 69	1	
$\phi_{\text{ICU}}(70-79)$	icu_allocation_70_79	fraction receiving ICU treatment if required in age group 70 - 79	1	
$\phi_{\text{ICU}}(80)$	icu_allocation_80	fraction receiving ICU treatment if required in age group 80+	0.5	
μ_{sym}	mean_time_to_symptoms	mean time to symptoms, days	6	(6)
σ_{sym}	sd_time_to_symptoms	s.d. time to symptoms, days	4.39	(10)
μ_{hosp}	mean_time_to_hospital	mean time to hospital after showing symptoms, days	5.14	(13)
μ_{crit}	mean_time_to_critical	mean time to critical after hospitalisation, days	2.5	ISARIC
μ_{death}	mean_time_to_death	mean time to death once critical if in ICU, days	6	(2)
σ_{death}	sd_time_to_death	s.d. time to death once critical if in ICU,days	2	(2)
$\mu_{\text{crit,surv}}$	mean_time_critical_survive	mean time in ICU if survive, days	4	(2)
$\sigma_{\text{crit,surc}}$	sd_time_critical_survive	s.d. time in ICU if survive, days	2	(2)
μ_{rec}	mean_time_to_recover	mean time to recover after symptoms (non-hospitalised), days	12	(15)
σ_{rec}	sd_time_to_recover	s.d. time to recover after symptoms (non-hospitalised), days	5	(15)
$\mu_{\text{hosp,rec}}$	mean_time_to_hospitalised_recover	mean time to recover after hospitalisation (no critical care) days	8	
$\sigma_{\text{hosp,rec}}$	sd_time_to_hospitalised_recover	s.d. time to recover after hospitalisation (no critical care) days	3	
$\mu_{\text{a,rec}}$	mean_asymptomatic_to_recover	mean time to recover after asymptomatic, days	15	(15)
$\sigma_{\text{a,rec}}$	sd_asymptomatic_to_recover	s.d. time to recover after asymptomatic, days	5	(15)

TABLE 8. Parameters corresponding to passive interventions (hospitalisation and self-quarantine upon symptoms). Parameter values match the Fraser group IBM model parameters, April 9, 2020

Passive Intervention Parameters		
Name	Description	Value
hospitalised_daily_interactions	Daily random interactions of a hospitalised individual	0
quarantined_daily_interactions	Daily random interactions of a quarantined individual	0
self_quarantine_fraction	Proportion of people who self-quarantine upon symptoms	0.8
quarantine_length_self	Maximum number of days quarantine for individuals self-reporting symptoms	7
quarantine_dropout_self	Daily probability of drop out for an individual quarantining after self-reporting symptoms	0.02
daily_non_cov_symptoms_rate	Daily probability of reporting similar symptoms which are not covid-19, including seasonal flu	0.002
lockdown_work_network_multiplier	Multiplier applied to proportion of work contacts interaction in lock-down	0.2
lockdown_random_network_multiplier	Multiplier applied to proportion random contacts interaction in lock-down	0.2
lockdown_house_interaction_multiplier	Multiplier applied to strength of transmission in household contacts in lock-down	1.5

6.4. **Shield Group.** Given that fatality rate is highly skewed towards the over 70s, we have the option of applying a lock-down just to this demographic group. The effect of the lock-down is the same as that described in section 6.3, with the exception that interactions on the household network are not increased.

7. Active Interventions

We define active interventions to be those which involve contact tracing or testing. There are three events in the IBM which can be the initial trigger for an active intervention:

(1) developing symptoms (true covid-19 or not) in the community
(2) testing positive for covid-19
(3) hospitalisation (clinical diagnosis alone, or combined with a positive test result)

The individual who provides this initial trigger is the index-case. There are three types of active intervention which can be triggered:

(1) testing for covid-19 infection
(2) self-quarantining
(3) digital contact tracing

7.1. **Testing for infection.** Currently the test for SARS-COV-2 is not sensitive immediately upon infection, which we model by only returning a positive test if the patient has been infected for three days when the test takes place. There are also delays in the testing procedure between ordering and taking a test, and then getting results. These test delays are modelled in the IBM, adding further delays to interventions and so reducing their efficacy. When an individual has received a positive test they should already be self-quarantining, but we continue to include a dropout rate for poor adherence. A positive test can also trigger digital contact tracing, if this was not triggered already when an individual reported symptoms.

7.2. **Self-quarantining.** Individuals are asked to self-quarantine upon developing symptoms, receiving a positive test and / or being contact-traced. Again, we use a daily random dropout rate for individuals leaving quarantine early, and quarantine is modelled by stopping interactions on the individual's workplace network and greatly reducing their number of interactions on the random network.

7.3. **Digital contact tracing.** We model app-based contact tracing which can initiate the quarantining of infected individuals prior to them showing symptoms or testing positive. For infections with high levels of pre-symptomatic transmission, contact tracing is vital to control epidemics. In the model, contact tracing can only originate from somebody with the app and can only trace others who also have the app. App uptake is age-dependent in our model based on smartphone ownership data. Digital proximity sensing is likely to miss some interactions between individuals even if both individuals have the app, so when contact tracing we randomly drop a number of interactions. To contact trace we go through all the interactions the individual has had for the past seven days where both parties have the app and the interaction has not been missed. The IBM allows us to explore various policy options such as requesting traced individuals to self-quarantine, requesting tests for traced individuals, and requesting that their household self-quarantine too. The IBM also contains the option for recursive tracing of contacts of contacts.

7.4. **Quarantine Release.** In additional to completing the specified quarantine time or dropping-out, there are two other mechanisms for ending quarantine. The first is if the index-case tests negative for covid-19, at which point a message is sent to all contacts who were originally traced telling them they can stop quarantining. Note, if an individual has been traced by multiple index-cases, they only get the release message once all index-cases have tested negative. The second mechanism for the early termination of quarantining uses the tracing network to determine whether it is likely the (self-reported) index-case did not have covid-19. After `quarantine_smart_release_day` days of an index-case reporting symptoms, we check to see if any of the traced individuals have developed symptoms. If none have developed symptoms, we send a release message to all those traced that they can stop quarantining (with the same caveat for those traced by multiple index-cases).

8. Implementation Details

This is a high-level description of how the IBM is implemented. It is coded in C, using an object-orientated pattern. All required memory is pre-allocated at the start of the simulation for efficiency.

8.1. **Events.** We use an event-based system to drive disease progression in individuals and interventions, where at each decision point we calculate when the next event will occur and add it to an event list for that day. For each type of event there is an `eventlist` structure which contains an array of linked lists for each day of the simulation. We use doubly linked lists to allow for efficient deletion as well as insertion of events. Event lists also keep track of the current and total number of events by type (defined as all entries today and in the past).

8.2. **Individuals.** Each person in the population is represented by an `individual` structure and the population is static. The `individual` structure contains the following information:

(1) Demographic - age, house number, network membership
(2) Interaction diary - list of all interactions over a period of days (note this is not only the ones tracked by the app)
(3) Disease - current status (i.e. symptomatic, severity) and pointers to both current and future disease events

TABLE 9. Parameters corresponding to active interventions. Parameter values match the Fraser group IBM model parameters, April 9, 2020

Active Intervention Parameters		
Name	Description	Value
allow_clinical_diagnosis	Commence contact tracing on a hospital clinical diagnosis	1
days_of_interactions	Length of historic interactions traced (days)	10
quarantine_days	The number of previous days' contacts to be traced and contacted	7
test_insensitive_period	Number of days following infection the test is insensitive	3
app_users_fraction_0_9	Maximum fraction of the population with the app in age group 0-9	0
app_users_fraction_10_19	Maximum fraction of the population with the app in age group 10-19	90
app_users_fraction_20_29	Maximum fraction of the population with the app in age group 20-29	96
app_users_fraction_30_39	Maximum fraction of the population with the app in age group 30-39	95
app_users_fraction_40_49	Maximum fraction of the population with the app in age group 40-49	91
app_users_fraction_50_59	Maximum fraction of the population with the app in age group 50-59	81
app_users_fraction_60_69	Maximum fraction of the population with the app in age group 60-69	64
app_users_fraction_70_79	Maximum fraction of the population with the app in age group 70-79	41
app_users_fraction_80	Maximum fraction of the population with the app in age group 80+	27
traceable_interaction_fraction	Fraction of interactions that captured if both users have app	0.8
tracing_network_depth	Depth of interaction network to contact (i.e. second-order contacts)	0
test_on_symptoms	Test individuals who show symptoms and are quarantined	0
test_on_traced	Test individuals who have been contact-traced	0
test_order_wait	Minimum number of days to wait to take a test	0
test_result_wait	Number of days to wait for a test result	0
trace_on_positive	Trace contacts of an individual who tests positive (0=no, 1=yes)	0
trace_on_symptoms	Trace contacts of individuals who show symptoms (0=no, 1=yes)	0
quarantine_on_traced	Quarantine individuals who are traced (0=no, 1=yes)	0
quarantine_smart_release_day	Release a chain of quarantined people if after this number of days nobody has shown symptoms on the chain	0
quarantine_household_on_positive	Quarantine household members of a person with a positive test (0=no, 1=yes)	0
quarantine_household_on_symptoms	Quarantine household members of a person with symptoms (0=no, 1=yes)	0
quarantine_household_on_traced	Quarantine household members of a person who has been traced (0=no, 1=yes)	0
quarantine_household_contacts_on_positive	Quarantine the contacts of each household member of a person who tests positive (0=no, 1=yes)	0
quarantine_household_contacts_on_symptoms	Quarantine the contacts of other household members when someone gets symptoms	0
quarantine_length_traced	Maximum number of days quarantine for individuals who are traced	14
quarantine_dropout_traced	Daily probability of drop out for an individual quarantining after being traced	0.02
quarantine_length_positive	Maximum number of days quarantine for individuals with a positive test result	14
quarantine_dropout_positive	Daily probability of drop out for an individual quarantining after a positive test result	0.02

(4) Quarantine - is the person currently quarantined, and pointers to that event and the release event

8.3. **Network Construction.** Each interaction network has an associated `network` structure which contains an array of edges. Network structures can be static (i.e. household), static but down-sampled (i.e. workplace) or dynamically generated at each time step (i.e. random). Network generation is modular and any network can be added to the IBM as long as it can be represented by an array of edges. Once all the networks have been defined, we add each edge to the individual's interaction diary (which are single-linked lists as deletion is not required).

8.4. **Transmission.** The next step is to transmit the pathogen across today's interaction network, which is done as a push from all infected people (by disease status). For every infection status we pre-calculate the transmission rate for someone who has been infected for that length of time. At each time-step we go through all the interactions the infected person had for that day and calculate whether transmission has occurred. Instead of randomly drawing whether transmission has occurred for each interaction, we allocate each individual a quantity of *hazard* (from an exponential distribution) at the start of the simulation. Each interaction with an infected person reduces the persons *hazard* and when a person's *hazard* drops below 0 they become infected. This is mathematically equivalent to randomly drawing individual interactions, which can be seen by calculating the probability of being infected by the N^{th} interaction $P(\text{infected } N^{\text{th}})$, after exposured to interactions with hazard-rates

$\{\lambda_1, .., \lambda_N\}$

$$\begin{aligned} P(\text{infected } N^{\text{th}}) &= P_N \prod_{i=1}^{N-1} (1 - P_i) \\ &= \exp\left(- \sum_{i=1}^{N-1} \lambda_i \right) - \exp\left(- \sum_{i=1}^{N} \lambda_i \right) \\ &= P\left(\sum_{i=1}^{N-1} \lambda_i < T < \sum_{i=1}^{N} \lambda_i \right) \end{aligned}$$

where T is distributed exponentially with mean 1. The fact that different age groups have different susceptibilities (S_{a_s}) is then modelled by allocating different amounts of initial *hazard* to each group. This improves computational efficiency so it is not necessary to draw a random variable for each potential transmission event.

8.5. **Digital Tracing and Release.** When an individual self-reports symptoms or tests positive, they become an index case and are assigned an index token (which form a linked list). Digital tracing is performed by looping through all contacts in the interaction diary and if the app recorded the interaction (i.e. both individuals have the app and the interactions was not randomly missed), a quarantine message is sent along with the index token. Individuals store every index token they have been traced by for 14 days. Upon a negative test of the index case (or other release mechanism), a message is sent to every phone with that index token to remove it. Upon removal of all index tokens, an individual is automatically released from quarantine.

8.6. **Performance.** The IBM for 100k individuals takes approximately 100ms per day to run and requires 250Mb of memory on a 2015 MacBook Pro. Both speed and memory are linear in population size (tested from 100k to 1m). 96% of the CPU usage is spent on rebuilding the daily interaction networks and updating the individual's interaction diaries. About 60% of the memory usage is spent on storing the interaction diaries (for 5 days), another 20% is spent on storing the semi-static networks with the remaining 20% spent on storing individuals and their states in event lists.

References

[1] Qifang Bi, Yongsheng Wu, Shujiang Mei, Chenfei Ye, Xuan Zou, Zhen Zhang, Xiaojian Liu, Lan Wei, Shaun A Truelove, Tong Zhang, Wei Gao, Cong Cheng, Xiujuan Tang, Xiaoliang Wu, Yu Wu, Binbin Sun, Suli Huang, Yu Sun, Juncen Zhang, Ting Ma, Justin Lessler, and Teijian Feng. Epidemiology and transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts. *medRxiv*, 2020.

[2] Intensive care national audit and research centre. Icnarc report on covid019 in critical care 10 april 2020. 2020.

[3] Tiago Henrique de Souza, Jose Antonio Nadal, Roberto Jose Negrao Nogueira, Ricardo Mendes Pereira, and Marcelo Barciela Brandao. Clinical manifestations of children with COVID-19: a systematic review. *medRxiv*, 2020.

[4] Yuanyuan Dong, Xi Mo, Yabin Hu, Xin Qi, Fang Jiang, Zhongyi Jiang, and Shilu Tong. Epidemiological characteristics of 2143 pediatric patients with 2019 coronavirus disease in China. *Pediatrics*, 2020.

[5] Neil M Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. *Report 9*, 2020.

[6] Luca Ferretti, Chris Wymant, Michelle Kendall, Lele Zhao, Anel Nurtay, Lucie Abeler-Dörner, Michael Parker, David Bonsall, and Christophe Fraser. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, 2020.

[7] Tapiwa Ganyani, Cecile Kremer, Dongxuan Chen, Andrea Torneri, Christel Faes, Jacco Wallinga, and Niel Hens. Estimating the generation interval for COVID-19 based on symptom onset data. *medRxiv*, 2020.

[8] Xiaoxia Lu, Liqiong Zhang, Hui Du, Jingjing Zhang, Yuan Y Li, Jingyu Qu, Wenxin Zhang, Youjie Wang, Shuangshuang Bao, Ying Li, et al. SARS-CoV-2 infection in children. *New England Journal of Medicine*, 2020.

[9] Lei Luo, Dan Liu, Xin-long Liao, Xian-bo Wu, Qin-long Jing, Jia-zhen Zheng, Fang-hua Liu, Shi-gui Yang, Bi Bi, Zhi-hao Li, Jian-ping Liu, Wei-qi Song, Wei Zhu, Zheng-he Wang, Xi-ru Zhang, Pei-liang Chen, Hua-min Liu, Xin Cheng, Miao-chun Cai, Qing-mei Huang, Pei Yang, Xin-fen Yang, Zhi-gang Huang, Jin-ling Tang, Yu Ma, and Chen Mao. Modes of contact and risk of transmission in covid-19 among close contacts. *medRxiv*, 2020.

[10] Shujuan Ma, Jiayue Zhang, Minyan Zeng, Qingping Yun, Wei Guo, Yixiang Zheng, Shi Zhao, Maggie H Wang, and Zuyao Yang. Epidemiological parameters of coronavirus disease 2019: a pooled analysis of publicly reported individual data of 1155 cases from seven countries. *medRxiv*, 2020.

[11] Kenji Mizumoto, Katsushi Kagaya, Alexander Zarebski, and Gerardo Chowell. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the diamond princess cruise ship, yokohama, japan, 2020. *Eurosurveillance*, 25(10):2000180, 2020.

[12] Joël Mossong, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS*

medicine, 5(3), 2008.

[13] Lorenzo Pellis, Francesca Scarabel, Helena B Stage, Christopher E Overton, Lauren HK Chappell, Katrina A Lythgoe, Elizabeth Fearon, Emma Bennett, Jacob Curran-Sebastian, Rajenki Das, et al. Challenges in control of Covid-19: short doubling time and long delay to effect of interventions. *arXiv preprint arXiv:2004.00117*, 2020.

[14] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440, 1998.

[15] Xiaobo Yang, Yuan Yu, Jiqian Xu, Huaqing Shu, Hong Liu, Yongran Wu, Lu Zhang, Zhui Yu, Minghao Fang, Ting Yu, et al. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *The Lancet Respiratory Medicine*, 2020.

[16] Yang Yang, Qingbin Lu, Mingjin Liu, Yixing Wang, Anran Zhang, Neda Jalali, Natalie Dean, Ira Longini, M. Elizabeth Halloran, Bo Xu, Xiaoai Zhang, Liping Wang, Wei Liu, and Liqun Fang. Epidemiological and clinical features of the 2019 novel coronavirus outbreak in china. *medRxiv*, 2020.

[17] Juanjuan Zhang, Maria Litvinova, Yuxia Liang, Yan Wang, Wei Wang, Shanlu Zhao, Qianhui Wu, Stefano Merler, Cecile Viboud, Alessandro Vespignani, Marco Ajelli, and Hongjie Yu. Age profile of susceptibility, mixing, and social distancing shape the dynamics of the novel coronavirus disease 2019 outbreak in china. *medRxiv*, 2020.