Societal heterogeneity contributes to complex dynamic patterns of the COVID-19 pandemics: insights from a novel Stochastic Heterogeneous Epidemic Model (SHEM)

Authors: Alexander V. Maltsev[1] and Michael D. Stern[2]*

Affiliations:

[1]Translational Gerontology Branch, NIA/NIH, Baltimore, MD, USA

[2]Laboratory of Cardiovascular Science, NIA/NIH, Baltimore, MD, USA

**Corresponding Author*:

Michael D. Stern,

Email: SternMi@mail.nih.gov

ORCID information:
Michael D. Stern: 0000-0003-4476-7425
Alexander V. Maltsev: 0000-0002-2219-1114

Abstract:

After months of COVID-19 quarantine, businesses are reopening their doors and people are reentering society. Within the current COVID-19 data, after a slow-down of infection, a new peak in active cases is already forming. Here we developed a new Stochastic Heterogeneous Epidemic Model (SHEM) to investigate genesis of complex pandemic patterns with the focus on the role of heterogenous societal structure. Using this model with R_0 of COVID-19, we simulated viral infection in different scenarios where isolated, communities surround the main cluster of the population. Depending on the parameters of heterogeneity and isolation period, our simulations generated a multimodal growth periods with multiple peaks, an extended plateau, a prolonged tail, or a delayed second wave of infection. We show that timing and magnitude of infection for previously unaffected isolated clusters of people, such as suburban neighborhoods, are critical aspects of these patterns. Our model can be applied to communities at any given scale of population described as a fractal-like structure, i.e. from the entire human population, to a country, down to a province or city levels with relevant societal heterogeneity structure. The current COVID-19 pandemic development worldwide and in the US follows a bimodal rise pattern, qualitatively similar to that in our simulations. We interpret our data to indicate that the secondary peak in the pattern is contributed by the states and counties in the US with late infection surges that are analogous to isolated suburbs in our model. Furthermore, the on-going early reopening, i.e. premature partial reopening in our model, is further accelerating the peak rise. This peak now reaches new heights, forcing many states in the US to reverse their policies and reestablish their quarantine measures. Our results support these timely and effective measures: we show that longer quarantine periods can reduce the number of deaths and transform the current trend into a substantially delayed (>1 year) second wave. If this scenario

becomes a reality, it is important (i) to develop the vaccine and/or effective treatment before the second wave; (ii) to warn people living in suburbs that it is these isolated areas that may hold a false sense of security, but they should continue to take extra care for their public health.

Significance Statement:

We developed a new multiscale Stochastic Heterogeneous Epidemic Model (SHEM) and demonstrated a major role of heterogeneity of societal structure in pandemic development. Using this model with R_0 of COVID-19, we simulated viral infection in communities surrounding the main cluster of the population. Depending on the model parameters, our simulations of COVID-19 infection generated a large variety of dynamic patterns, including a multimodal growth period with multiple peaks observed now in the US and worldwide, as well as a delayed second wave of infection. The timing and magnitude of infection for these previously unaffected isolated clusters of people are critical aspects of these patterns. Our abstract multiscale model can be applied at given levels of human societal organization with relevant societal heterogeneity specified at those levels.

Introduction

Coronaviruses represent one of the major pathogens that primarily target the human respiratory system. Previous outbreaks of coronaviruses (CoVs) that affected humans include the severe acute respiratory syndrome (SARS)-CoV and the Middle East respiratory syndrome (MERS)-CoV (1). COVID-19 is a disease caused by the novel coronavirus SARS-COV2 virus that is both fatal and has a high transmission rate (R_0), almost twice that of the 2017-2018 common influenza (2, 3). The World Health Organization stated that this combination of high health risk and susceptibility is of great global public health concern, and efforts must be directed to prevent further infection while vaccines are still being developed (4). As of July 2020, there are more than ten million confirmed COVID-19 cases with more than a half of million deaths. Older adults seem to be at higher risk for developing more serious complications from COVID-19 illness (5, 6). In today's absence of a vaccine and impactful treatments, the most effective way to combat the virus is to find and implement mitigation strategies for susceptible populations. An invaluable resource in this difficult task is numerical modeling studies that can reveal key factors in pandemic development.

What models could be useful? Direct study of the available data of COVID-19 is complicated because many cases and deaths are underrepresented. However, a simple model that correctly captures large-scale behaviors, but gets some details wrong, is useful, whereas a complicated model that gets some details correct but mischaracterizes the large-scale behaviors is misleading (7). During the H1N1 pandemic, generic (i.e. non-specific) stochastic influenza models were an important tool to understand and quantify the full effects of the virus in simulations of important scenarios (8). Stochastic simulation models such as FluTE (2010) or GLEaM (2011) can simulate the spatial interaction and clusterization of millions of people (9,

10). The FluTE model includes fine details of the societal structure, such as demography, host movement, and social contact rates in workplaces, schools, and homes. Simulation results of viral propagation can deduce specific parameter measures that would mitigate worst case scenarios. With respect to COVID-19, the FluTE model has been recently used to offer interventions to mitigate early spread of SARS-CoV2 in Singapore (11). The international outbreak of COVID-19 has also recently been modeled using GLEaM by Chinazzi et al. (12). As a result, the effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak was assessed and a mitigation strategy was suggested.

The bimodal growth pattern is already visible in real data going into July 2020, distinguishing itself from the curve generated by simple models such as SEIR (13). The GLEaM model predicts a steady decrease in disease cases, however it is possible that COVID-19 generates a distinct second wave of infection that may be more severe than the first, as seen in the influenza pandemics of 1918 (H1H1), 1957 (H2H2), 1968 (H3N2), and 2009 (H1N1). Recently updated forecasts from a popular University of Washington model known as IHME (Institute for Health Metrics and Evaluation) predicts a second wave of the spread of COVID-19 just as 2020 summer ends. A novel multi-region pandemic model (LIST) of Friston et al. (14) which incorporates regional factors, does display in some cases, a second wave of infection. The model suggests that a second wave of infection typically occurs when a population is re-exposed to a virus by an influx of infected people from another region. Furthermore, in a separate study by Althouse et al (15), the transmission of disease is shown to be stochastic, and is commonly dominated by a small number of individuals, driven by super-spreading events.

Despite extensive efforts to understand and predict the COVID-19 spread, the key factors that determine the bimodal rise pattern, asymmetry of the recovery phase, and the emergence of

a second wave remain unclear. In principle, agent-based models like FluTE that track the fate of every individual ought to be the most precise. However, in addition to requiring vast computational resources, these models require detailed statistical information about the social interactions and grouping of individuals. This information is well characterized for seasonal influenza, but has been completely disrupted by mitigations for COVID-19 and continues to be in flux, dynamic and unpredictable in ways that were not anticipated even weeks ago. Therefore, instead of a data-based predictive model like those described above, we chose to develop a scenario model that can be used as a tool to study the consequences of an arbitrary set of hypothesized conditions in a user-specified network of populations, which could be not only geographical but also social subgroups. Indeed, one underexplored but important factor of pandemic spread is regional heterogeneity which is characterized by various subpopulation clusterization, societal interaction, and disease mitigation strategies. Our hypothesis is that the multiple peaks and waves being observed now are the consequence of asynchronous propagation of infection among groups with widely varying degrees of intra-group interaction and isolation from main hubs (a metapopulation of infections as the term is used in ecology). We also attempt to take account of over-dispersed stochasticity (15), which is usually not incorporated compartmental models but can be critical in small or virgin populations.

To approach this problem, we developed a novel Stochastic Heterogeneous Epidemic Model (dubbed SHEM) which for the first time incorporates heterogeneous aspects of society. The model was inspired by stochastic models of local calcium release dynamics inside heart cells, driven by explosive calcium-induced-calcium-release (16, 17). As a demonstration of this tool, we examine several key scenarios of heterogeneity where separate communities of various clusterization and transmission levels (R_0) surround the bulk of the population. For purposes of

this model, R_0 is defined as the basic reproduction number, i.e. the average number of secondary cases (that would be) generated by one infectious individual over the course of his infection if all of his contacts were susceptible. This is a property of the virus and of the social behavior of individuals, but is distinct from $R(t)$ the actual realized reproduction number at time t subject to the degree of immunity of the population at that time. The bulk of our population was assigned R_0=2.5 which is within the range of SARS-CoV-2 basic reproduction number based on the early phase of COVID-19 outbreak in Italy (18). Interplay of various degree of heterogeneity and isolation periods in our model generated various dynamic patterns of the pandemics, including a multi-modal growth periods, an extended plateau, prolonged tail, or a delayed second wave of infection.

Results

The aim of our simulations was not to closely mimic COVID-19 development globally or in any particular country or city; rather we wanted to generate and investigate a variety of patterns of infection dynamics for various lockdown periods in different scenarios of societal heterogeneity. Thus, our simulations differed in clusterization and in R_0 and the time when R_0 is changed. Our model is composed of a large cluster of individuals weakly connected to smaller much more susceptible clusters of individuals that have their own internal R_0. In our topology the small clusters are not connected to one another. Our simulations have three stages, a fully opened level, a closed level, and a reopen level. We performed two different sets of simulations (described below); each set had different scenarios characterized by a specific set of parameters; and because our model is stochastic, each scenario was run 10 times.

Simulations with equal urban and suburban populations

In the first set of simulations we examined the virus spread in simple hypothetical scenarios with equal urban and suburbs populations (Fig. 1). The large (urban) cluster was composed of 1 million individuals set to R_0=2.5 (open level). The suburban population consists of 250 clusters, each with 4000 +/- 500 people and with the same internal R_0=2.5 that remained constant throughout all simulation stages. The urban large cluster is weakly connected with 0.001% movement into the suburban clusters (*alphinpop*) while suburban clusters have 0.1% movement into the urban cluster (*alphoutpop*). We investigated the 4 scenarios, specified below. In each scenario except #1, the large cluster closed to R_0=1.25 at t=40 days (closed level, e.g. This New York City under lockdown, based on 21% antibody positive at the peak (19).

1) No mitigation, i.e. freely expending pandemic: The large cluster of individuals stays always open.
2) Early, partial (premature) reopening to R_0 = 1.9 at 100 days.
3) Moderate lockdown period with full reopening at 225 days to R_0 = 2.5.
4) Late full reopening at 365 days to R_0 = 2.5.

A general tendency throughout all 4 scenarios was that as the lockdown period increases, the magnitude of the infection decreases but its span increases. At the same time the interplay of urban cluster and suburban clusters generated a variety of specific patterns in virus spread dynamics.

In the first “no mitigation” scenario (Fig. 1A) the suburban areas generate a strong second peak at the time when infection in the urban cluster had gone through its peak and is decaying. The second peak substantially extends the overall span of the pattern, nearly twice as much.

The epidemic rise in the “premature reopening” scenario (Fig. 1B) is multi-modal. The rise of infection substantially slowed in the main cluster during closed period, but a notable contribution from the suburban areas becomes notable closer to the reopening at day 100 (Fig. 1B, inset). In the reopen period, the infection surges in both the suburban and urban clusters. The peak of the suburban clusters happens later than the urban cluster, creating an apparent plateau in active infection cases from day 175 to 225.

In the “moderate reopening” scenario (Fig. 1C), a complex, multi-modal rise is also observed and is characterized by two distinct peaks. During the closed stage, the first peak in total infections emerges from isolated suburban clusters, while the second peak in total infection in the reopen stage is composed mainly by the urban cluster.

The “late reopening” (Fig. 1D), decreased infection during the first wave in both urban and suburban clusters, but resulted in a delayed second wave of infection. This second wave pattern reflects the most effective mitigation, because the respective second wave of deaths can be prevented with the timely development of a vaccine or effective treatment (for example a year from the time of infection onset in Fig. 1E.

Simulations of smaller suburbs population highly susceptible to COVID-19

Another set of simulations was performed to give insight into possible future waves of infection with regard to super-spreading events in highly susceptible, highly isolated communities such as nursing homes, religious places of worship, prisons, compact university cities, or shelters that are very socially distanced from the main community, but are very dense and vulnerable internally. The large (urban) cluster was composed of 1 million individuals set to R_0 = 2.5 (open level). The highly susceptible population consists of 300 clusters, each with 1200

+/- 500 people and with the same internal R_0 = 5.0 that remained constant throughout all simulation stages. The urban large cluster is weakly connected with 0.0006% movement into the susceptible clusters (*alphinpop*) while susceptible clusters have 0.06% movement into the urban cluster (*alphoutpop*). We investigated 2 scenarios as specified below. In each scenario the large cluster closed to R_0 = 1.25 at t=70 days (closed level).

1) Early reopening: The large cluster of individuals partially reopens to R_0 = 1.9 at t=160 days during the decay of the first wave.
2) Late reopening: The large cluster of individuals fully reopens to R_0 = 2.5 at t=350 days after the first wave is over.

The first scenario simulates viral infection featuring a premature reopening phase starting on day 160. The major result of scenario 1 is that a distinct second wave does not form following the premature opening; rather a prolonged tail of infection emerges that spans for about a year (Fig. 2A). Also a bimodal shape of total infection is formed and is caused by the delayed peak of infections in isolated clusters (Fig. 2B). Note that the current state of infection in the US (Fig. 2C) resembles the bimodal shape in Fig. 2B. A second wave of infection is demonstrated in Scenario 2 (Fig. 3). It is also important to note than the reopen phase is substantially longer than the open phase (nearly three years, compared to one). Furthermore, the onset of this delayed second wave substantially varied in each stochastic simulation run of our simulations (see ten runs overlapped in Fig. 3C). The prolonged quarantine mitigates the most deaths (Fig. 3D), but the second wave pattern can also save more lives if a vaccine or effective treatment is timely developed. Note that some simulations do not generate a second wave, reflecting the stochastic nature of viral propagation.

Discussion

In recent days (as of July, 2020) it has become abundantly clear that the time course of the epidemic in various locations has been very different from a standard smooth bell curve. We hypothesized that the multiple, asynchronous waves and plateaus are in part due to excess heterogeneity, as well as changing efforts at mitigation by social distancing. Geographic heterogeneity is included in many forecasting models (12, 14, 20) which use extensive, public databases of population characteristics and travel patterns, but these do not fully account for the stratification of the social behavior that controls the spread of the virus. Efforts at national or global forecasting are also limited by the patchy quality of the data. On the basis of an early antibody survey that estimated that there are ~50 asymptomatic infections for every one detected; one sophisticated forecasting model predicted in April that California would achieve herd immunity within a few weeks (21), an expectation that was alarmingly refuted by infection numbers on the first week of July.

For these reasons, instead of building a data-based forecasting and estimation model, we chose to develop a scenario-modeling tool that can be used to explore mechanisms in idealized models of social stratification constructed by the investigator. The model is built as a network of "populations" which could, in fact, be social strata of geographic populations. The definition of a metapopulation in ecology is a population of a single species which is spread among different environments that determine the local survival or extinction. Our model could therefore be considered a metapopulation of SARS-Cov2.

An epidemic has been likened to a forest fire, which spreads by diffusion along a front, but can also jump by embers that may or may not start a new blaze. Such spread to virgin areas, with a virus as with a fire, is intrinsically stochastic and such stochasticity, which is not explicitly

included in mean-field models, may contribute to the remarkable patchiness of the COVID-19 epidemic. This has caused the epidemic to appear entirely different to observers in different locations, leading to politicization of the response, which is, itself, a form of social heterogeneity. For rare spread to small, isolated subgroups (embers) this stochasticity is crucial. Patchiness is aggravated by the over-dispersion (super-spreading) of secondary cases of COVID-19, where the majority of infected individuals do not spread the virus, but some can cause up to a hundred secondary infections. We therefore built a model that is explicitly stochastic, with a mechanism to account for over-dispersion, which is not straightforward in a time-dependent evolution model.

Using this model with R_0 of COVID-19, we simulated viral infection in different scenarios where isolated, communities surround the main cluster of the population. We found that the contribution of these communities can be substantial and its timing will determine whether there will be (i) a multimodal growth (Figs 1C, 2B) or (ii) an apparent plateau (Fig. 1B) when isolated clusters contribute later. Furthermore, our simulations can generate a distinct, separate second wave of infection when society has a late full reopening (Fig. 3B). The onset of this second wave can be delayed by a stochastically variable amount (Fig. 3C), appearing seemingly "out of nothing" when the infection has almost disappeared and this second wave can last three times longer than the first wave.

Our simulation results update the current knowledge on virus mitigation strategies that were suggested by recent COVID-19 mitigation studies by Chinazzi et al, Chao et al, and Friston et al. (12, 14, 22). The work of Chinazzi et al. used GLEaM to demonstrate that travel restrictions introduced in Wuhan in January 2020 only delayed epidemic progression by 3 to 5 days within China, and international travel restrictions only helped slow infectious spread until

mid-February. Their results suggest that early detection, hand washing, self-isolation, and household quarantine are more effective than travel restrictions at mitigating this COVID-19 pandemic. Thus, with regards to our results, travel restrictions in preparation of the second wave or additional growth period is not feasible for disease prevention, but other basic precautions, including self-isolation and social distancing remain extremely important. The work of Chao et al. shows that most infections in their model actually originate from pre-symptomatic people or from those who never exhibit symptoms, and it may be impossible to prevent most infections with test-and-isolation policies unless non-symptomatic people are also isolated. In our model, we make no attempt to distinguish between symptomatic and asymptomatic cases, since the relative transmission of the virus is not clearly established. Such a distinction (included in a number of other models) could easily be added by subdividing the 5 compartments, at the risk of added complexity and more parameters needed in a scenario.

While our general recommendations are in accord with those in previous studies, we have discovered that previously unaffected isolated clusters of people have a delayed contribution to total infection cases, and therefore in these areas, mitigation strategies should be most closely followed until effective treatment or vaccination is developed. It is these communities that have the least herd immunity and are at risk for explosive scenarios if precautions are not taken. Mitigation strategies dictate the dynamics of the infection curve and the ideal mitigation strategy would be to have a long enough quarantine so that the second wave apparent in Fig. 1D and Fig. 3 would come late enough (more than one year) allowing a vaccination or effective treatment to be timely developed and implemented, preventing the second wave from emerging.

With respect to the current COVID-19 pandemic development worldwide and in the US, it follows a bimodal rise pattern, qualitatively similar to that in our simulations (Figs 2C vs. 2B).

Our important finding was that the second peak can originate from suburban clusters, and such peaks can emerge independent of lockdown period. For example, they occur within lockdown periods in our simulations in Figs 1B (inset), 1C, 1D, and 2B. Furthermore, an additional peak on the infection rising curve can be linked to the reopening, as shown in Fig. 1C (after day 225, i.e. after the lockdown period ends). Thus, the ongoing early reopening in many states in US is contributing to additional growth of the additional peak. This peak now marches into new heights, with a lot of uncertainty and anxiety about the future development of the pandemic, forcing many states to reverse their policies and reestablishing their quarantine measures. Our results support these timely and effective measures: our simulations in Fig. 3 demonstrate that longer quarantine periods can transform the current bimodal rise into a pattern with a substantially delayed (>1 year) second wave that can be prematurely terminated by vaccination (Figs 1E, 3D).

While we simulated the effects of communities surrounding the main cluster of the population and discussed our simulations with respect to the current COVID-19 situation in the US, this model can be translated to the entire heterogeneous society of the human population in the form of a fractal-like structure, i.e. large countries surrounded by small isolated countries, large highly-populated states (or provinces) surrounded by smaller isolated states, or big counties or cities surrounded by suburbs. In other words, while the COVID-19 pandemic is a complex multiscale problem, our model can be applied at any given scale and incorporate specific heterogeneous structure relevant at that scale.

Methods

Model purpose

In view of the constantly changing behavioral environment for COVID-19 in the United States, data-based predictive modeling of the future of the epidemic is difficult. Our model is intended to be a general tool with which to explore scenarios and determine the expected consequences of various interventions. The model is specifically intended to examine the effect of heterogeneity, including not only geographic but also social heterogeneity, i.e. the existence of groups within one geographic location that have different social interaction patterns and may be partially isolated from neighboring groups, *e.g.* nursing homes. The model is fully stochastic and, unlike most compartmental models, incorporates the effect of over dispersion of secondary infections (super spreading).

Structure of the Model

The general model consists of a number of subpopulations ("villages") whose number is limited only by memory. The simulation is based on a generalization of the SEIRD representation. The state of each village is represented by the numbers of individuals in each of 5 states: Susceptible, Exposed (destined to become infected), Infected/infectious, Recovered (immune) and Dead (however, see below under Super-spreading for additional state-dependence). Each village is, by definition, homogeneous and mixed. Villages could represent actual geographic units, but could also be groups or sub-regions that have different social interactions or behavior. At a given time, each village is characterized by its population, the expected mortality of virus infections, and its local value of the basic reproduction number R_0. R_0 is defined as the mean expected number of secondary infections spawned by one average infected individual over the duration of their illness, if the population were totally susceptible. It

is a property of both the virus and the behavior of individuals in the population. Villages are connected by a user-specified network of unidirectional links along which infection or individuals can travel at user-specified rates, including links from each village to itself to represent internal infection/recovery processes.

Simulation Method

The entire collection of populations is simulated as a single, continuous-time Markov chain (birth-death process). There are 14 types of possible events associated with each link:

1) Infection from source to target by transient contact
2) Infection from target to source by transient contact
3) Infected individual moves from source to target
4) Exposed individual moves from source to target
5) Susceptible individual moves from source to target
6) Infected individual moves from target to source
7) Exposed individual moves from target to source
8) Susceptible individual moves from target to source
9) Susceptible gets exposed inside village (self-link only)
10) Exposed converts to infected inside village (self-link only)
11) Infected recovers inside village (self-link only)
12) Infected dies inside village (self-link only)
13) Recovered moves from source to target
14) Recovered moves from target to source

Infection transmitted by transient contact is modeled as though members of one village spend some (small) fraction of their time (*i.e.* of their inter-personal contacts) "visiting" the opposite village at the other end of the link (an example would be staff working at a nursing home) adjusted for any mitigations. The spread of infection in each direction of the link has two components: (1) exposure of susceptibles by visiting infectious individuals and (2) exposure of visiting susceptibles in the visited village, who then carry the infection back home.

At each step, rates of these 14 possible events (computed as in classic SEIRD differential equations, except as below under Super-spreading) are summed over all links in the network to give a total transition rate R_{tot}. A uniformly distributed random number rn is generated and $-log(rn)/R_{tot}$ is taken to be the waiting time until the next event, considered as a Poisson point process. Time t is advanced by that amount and then a particular transition is selected by a second random number rn_2 by partitioning R_{tot} into segments proportional to their relative rates and finding which segment contains $rn_2 * R_{tot}$. The action associated with the prescribed event (*e.g.* increment/decrement numbers in states SEIRD of the villages connected by the link) is taken, and then the above is repeated for the next step. This is continued until t reaches t_{max} specified in the input parameter file. Output is produced whenever t crosses *tout*, which is then incremented to t plus a user-specified interval (usually one day). This time-binning is only for output and plays no role in the continuous-time evolution of the epidemic.

Super-spreading

It is known that the distribution of secondary COVID-19 infections generated by a single, infected individual is over-dispersed (*i.e.* has a long tail compared to the Poisson

distribution of infections expected if transmission were random). Although the average R_0 is estimated to be 2.5-4 in the absence of social distancing mitigations, contact tracing has shown that single individuals have infected up to a hundred others. This is known as super-spreading events, and can occur by several possible mechanisms, involving either a predilection of an individual (*e.g.* a celebrity who travels widely and contacts many other people) or a situation in which individuals were placed in unusually close contact (*e.g.* a church choir in an indoor location). On the other hand, the majority of infected individuals do not appear to spread the infection to anyone. It has been shown (15) that this over-dispersed distribution can be approximated by a negative binomial distribution, with mean R_0 (by definition) and dispersion parameter $r<<1$, for example 3 and 0.16. By iterating this distribution for several generations of viral spread, it is found that the eventual distribution of epidemic size is predicted to be quite different than found for a hypothetical stochastic transmission by Poison-distributed secondary infections with the same R_0.

Unfortunately, viral generations do not remain synchronous in time, so it is not straightforward to incorporate super-spreading in a time-dependent epidemic evolution model except by following the interactions and infections of each individual in the population, as done for example in the FLuTE simulation for influenza (8). This is very compute-intensive, but a more significant objection from our point of view is that it depends on knowing (statistically) the social interaction groups and travel behavior of the population at a fine-grained scale, and these have been severely disrupted by mitigation efforts during the current pandemic. Rather than speculate on these variables, we have developed a modified Markov scheme that tries to reproduce the observed distribution of secondary infections by replacing R_0 in the event-rate calculations by an infectivity that is itself stochastic.

Supposing that an individual produces, *a posteriori*, K secondary infections over the lifetime t_{dur} of his infection, the required rate is K/t_{dur} . However, t_{dur} is itself stochastic: in our model, as in simple SIR models, recovery is treated as a Poisson point process with a rate $1/trec$ per infected individual where *trec* is the observed mean duration of infections (it is not presently known whether that is true of asymptomatic infections). When there are multiple infected individuals present, if recovery events in the population are random with a rate $k_i/trec$, where k_i is the number of infected individuals in the village, a super-spreader is likely to be "recovered" before (or after) doing his "job". To avoid this we have adopted the following scheme:

- In each village j, at each step, an infectivity *inf(*j*)* is maintained that takes the place of k_i*R_0 in the SEIRD rate equations.
- Whenever a new infection is created (by conversion of an exposed individual), a random number K is drawn from a negative binomial distribution of mean R_0 and dispersion r_{eff} , the latter to be determined. *Inf* is incremented by K and the individual infectivity K is placed on the top of a linked list.
- Whenever a random recovery event is generated at the above-mentioned rate, the oldest individual infectivity is removed from the bottom of the list and subtracted from *inf*.

The overall rate of infection events, based on *inf*, will be a superimposition of the rates of the individual infection events generated by each infected individual. The mean number of secondary infections actually realized by one infected individual over the life of his infection will be $K*t_{dur}/trec$. Since infections recover in the order in which they were created, if there are n infections active, t_{dur} will be the nth waiting time of the Poisson point process at rate $n/trec$ (where $n=k_i(j)$), whose probability distribution is proportional to $poisson(n-1,n/trec)$. The individual infections generated by individual K are a Poisson point process, so the probability

that the individual actually generates *j* secondary infections is *poisson(j, K*t_{dur}/trec*). Integrating this over the distribution of t_{dur}, multiplying by the probability mass distribution of the negative binomial distribution and summing over *K* we find:

$$p(j,n) = \frac{n^n \Gamma(n+j) r^r \sum_{k=0}^{\infty} \frac{r_0^k k^j (n+k)^{-n-j} (r+r_0)^{-r-k} \Gamma(r+k)}{\Gamma(k+1)}}{j\Gamma(j)\Gamma(n)\Gamma(r)}$$

This is a long-tailed probability distribution that can be fit, by an appropriate choice of r_{eff}, to approximate the empirical negate

ve binomial distribution with *r=0.016* over the relevant range. With more than a few active infections present, the distribution converges to:

$$p(j,inf) = \frac{n^n \Gamma(n+j) r^r \sum_{k=0}^{\infty} \frac{r_0^k k^j (n+k)^{-n-j} (r+r_0)^{-r-k} \Gamma(r+k)}{\Gamma(k+1)}}{j\Gamma(j)\Gamma(n)\Gamma(r)}$$

We choose r_{eff} to give the best least-squares fit on a linear scale of the case *n=1*, which is the most important stochastic case since it governs the chance that a single infected individual can start an outbreak, with the chance that an infected individual causes no secondary infections, *p(0,1)=0.62* similar to the empirical distribution. These distributions are all normalized with mean R_0 and differ dramatically from the Poisson distribution (Fig. 4, dashed line) assumed in the classic SEIR model. Larger values of *n* are decreasingly important because the aggregate distribution of the actual infection rate controlled by the sum *inf* behaves similarly to *negbinomial* (R_0 *,n*r)* which converges to Poisson so stochastic effects become less important once there are many active cases.

Super-spreaders vs super-spreading events

Super-spreading can be a property of the individual or of the circumstances. What happens when an individual infected patient migrates to a new village? Does he keep his identity or does he assume the infectiousness typical of the local R_0 of his new environment? In the model we can make the choice, determined by the parameter *spreads*. If this is positive (default) then he keeps his prior *K* value which simply migrates from the top (newest) link to be added to the top of the infection list in the new village, thereby preserving his infectious lifetime in him new home. If *spreads* is negative, then the *K* value is re-randomized using the local R_0 and r_{eff} .

Description of Parameters

We created the program with customizable parameters for a main suburban cluster and two subpopulations. The suffix *"2"* that follows certain parameters denotes the value in the reopen stage (with exception of *amu2*, which is constant for all stages) and the suffix *"pop"* is used to designate subpopulation 1.

General Parameters
pname = embedded name of results
tbin = time step for output purposes
nh = total active clusters from all three populations
tswitch1 = start of the closed period
tswitch = start of the reopen period

Urban Cluster
r00 = R_0 of the urban cluster during the open stage
r01 = R_0 of the urban cluster during the closed stage
r02 = R_0 of the urban cluster during the reopen stage
kstot = population of the urban cluster
i0 = the integer number of infectious individuals at initialization in the urban cluster
kx0 = the integer number of exposed individuals at initialization in the urban cluster
amu0 = mortality rate per infection for the urban cluster and subpopulation 2 (constant)

Subpopulation 1
amu2 = mortality rate per infection for subpopulation 1 (constant)
btinpop = rate of people moving into a given subpopulation 1 cluster (open + closed)
btoutpop = rate of people moving out of a given subpopulation 1 cluster (open + closed)
btinpop2 = rate of people moving into a given subpopulation 1 cluster (reopen)
btoutpop2 = rate of people moving out of a given subpopulation 1 cluster (reopen)
alphinpop = the fraction of the urban cluster that visits a subpopulation 1 cluster on average (open + closed)
alphinpop2 = the fraction of the urban cluster that visits a subpopulation 1 cluster on average (reopen)
alphoutpop = the fraction of a subpopulation 1 cluster that visits the urban cluster on average (open + closed)
alphoutpop2 = the fraction of a subpopulation 1 cluster that visits the urban cluster on average (reopen)
anbar = expected population of each subpopulation 1 cluster
anvar = variability of each subpopulation 1 cluster
rinpop = internal R_0 of subpopulation 1 (open + closed)
rinpop2 = internal R_0 of subpopulation 1 (reopen)
npop = total number of subpopulation 1 clusters

Subpopulation 2
alphin = the fraction of the urban cluster that visits subpopulation 2 on average (open + closed)
alphin2 = the fraction of the urban cluster that visits subpopulation 2 on average (reopen)
alphout = the fraction of the subpopulation 2 that visits the urban cluster on average (open + closed)
alphout2 = the fraction of the subpopulation 2 that visits the urban cluster on average (reopen)
btin = rate of people moving into a given subpopulation 2 cluster (open + closed)
btout = rates of people moving out of a given subpopulation 2 cluster (open + closed)
btin2 = rate of people moving into a given subpopulation 2 cluster (reopen)
btout2 = rates of people moving out of a given subpopulation 2 cluster (reopen)
rin = internal R_0 of subpopulation 2 (open + closed)
rin2 = internal R_0 of subpopulation 2 (reopen)
amu0 = mortality rate per infection for the urban cluster and subpopulation 2 (constant)

Running the Program

The ELF64 executable file is included in Supplementary Materials and is meant to be run on Linux systems. The executable file should be placed into a directory with its corresponding parameter file. The parameter files we used to generate our results are also included in

Supplementary Materials. Before using a parameter file, rename it to *epistoch.dat*. To run the executable file, execute the command:

swnum=1 filenum=1 epinew.exe

The results will be generated in the same directory, with prefix *epiout*. An automatic parameter file with suffix *params* will also be generated. The *epiout* files are formatted as described below.

Output file

Files start with one line in quotes that displays the number of people in subpopulation 1 and 2 that became infected. Then there are 16 columns of numbers as follows:

1) time (not exactly the same for different files)

2) total exposed

3) total infected

4) total susceptible

5) total recovered

6) total dead

7-11) same for subpopulation 1

12-16) same for the urban cluster

Values for an additional subpopulation 2 may be inferred by subtracting the urban data and the subpopulation 2 data from total (columns 2-6).

Acknowledgments

The work was supported by the Intramural Research Program of the National Institute on Aging, National Institute of Health. This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD. (http://biowulf.nih.gov).

Figure legends

Fig. 1: Different dynamic patterns of COVID-19 spread in simulations with equal urban and suburban populations but various lockdown strategies. Main cluster closed at day 40 to R_0=1.25. The full lockdown periods are marked by green shade. Contributions are shown by different colors. **A,** suburban clusters generate a second delayed peak when no intervention is implemented. **B,** an apparent plateau after early reopening and complex rise pattern during close period (inset). **C,** A multimodal rise with additional peak generated by rural cluster after full reopening at day 225. **D,** A delayed second wave emerged after full reopening at day 365. **E,** The dynamics of total number of deaths in each scenario (A-D).

Fig. 2: Bimodal infection rise pattern generated during quarantine in simulations that included highly susceptible isolated clusters. A, A bimodal rise and prolong tail. The full lockdown period is marked by green shade. Contributions are shown by different colors. **B,** Expended bimodal rise pattern in panel A. **C**, Going into July 2020 United States active case curve exhibits a similar bimodal growth pattern (data as of July 8, 2020 from https://www.worldometers.info/coronavirus/country/us/).

Fig. 3: Second wave infection pattern generated by simulations of late reopening that included highly susceptible isolated clusters. A, Example of simulation with no lockdown mitigation in place. **B,** An example of a delayed second wave with contributions of each subpopulation shown by different colors. The full lockdown period is marked by green shade. **C**, Overlapped 10 simulations illustrate the stochastic nature of the second wave onset. **D**,

Lockdown prevents the death of nearly half of the people, as shown by side-by-side mortality curves for 10 runs of each condition.

Fig. 4: The distribution of secondary infections generated by infectious individuals. Black: Observed negative binomial distribution (15); Green, blue, magenta, red: The actual realized number of secondary cases generated over the lifetime of one infection in the presence of n other infections individuals according to our scheme. All distributions have mean R_0 =3.0. Dashed line: Poisson distribution with mean R_0 as implicit in mean-field SEIR models.

References

1. Rothan HA & Byrareddy SN (2020) The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J Autoimmun* 109:102433.
2. Xiaofang F, *et al.* (2019) A Severe Seasonal Influenza Epidemic During 2017–2018 in China After the 2009 Pandemic Influenza: A Modeling Study. *Infectious Microbes & Diseases* 1(1):20-26.
3. Liu Y, Gayle AA, Wilder-Smith A, & Rocklov J (2020) The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J Travel Med* 27(2).
4. Organization WH (2020) Coronavirus disease 2019 (COVID-19) Situation Report – 72. *https://apps.who.int/iris/bitstream/handle/10665/331685/nCoVsitrep01Apr2020-eng.pdf*.
5. Liu K, Chen Y, Lin R, & Han K (2020) Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. *J Infect* 80(6):e14-e18.
6. Li P, *et al.* (2020) Clinical features and short-term outcomes of elderly patients with COVID-19. *Int J Infect Dis* 97:245-250.
7. Siegenfeld AF, Taleb NN, & Bar-Yam Y (2020) Opinion: What models can and cannot tell us about COVID-19. *Proceedings of the National Academy of Sciences*:202011542.
8. Chao DL, Halloran ME, Obenchain VJ, & Longini IM, Jr. (2010) FluTE, a publicly available stochastic influenza epidemic simulation model. *PLoS Comput Biol* 6(1):e1000656.
9. Van den Broeck W, *et al.* (2011) The GLEaMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. *BMC Infect Dis* 11:37.

10. Balcan D, *et al.* (2010) Modeling the spatial spread of infectious diseases: the GLobal Epidemic and Mobility computational model. *J Comput Sci* 1(3):132-145.

11. Koo JR, *et al.* (2020) Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study. *Lancet Infect Dis* 20(6):678-688.

12. Chinazzi M, *et al.* (2020) The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 368(6489):395-400.

13. Weiss H (2013) The SIR model and the Foundations of Public Health. *MATerials MATemàtics* 3:1-17.

14. Friston KJ, *et al.* (2020) Second waves, social distancing,and the spread of COVID-19 across America. *https://arxiv.org/abs/2004.13017*.

15. Althouse BM, *et al.* (2020) Stochasticity and heterogeneity in the transmission dynamics of SARS-CoV-2. *https://arxiv.org/abs/2005.13689*.

16. Stern MD, *et al.* (2014) Hierarchical clustering of ryanodine receptors enables emergence of a calcium clock in sinoatrial node cells. *J Gen Physiol* 143(5):577-604.

17. Maltsev AV, Maltsev VA, & Stern MD (2017) Clusters of calcium release channels harness the Ising phase transition to confine their elementary intracellular signals. *Proc Natl Acad Sci U S A* 114(29):7525-7530.

18. Rahman B, Sadraddin E, & Porreca A (2020) The basic reproduction number of SARS-CoV-2 in Wuhan is about to die out, how about the rest of the World? *Rev Med Virol*:e2111.

19. Schultz C (2020) More than one in five people who were tested for virus antibodies in N.Y.C. had them. *The New York Times: https://www.nytimes.com/2020/04/23/nyregion/coronavirus-new-york-update.html*.

20. Pei S, Kandula S, & Shaman J (2020) Differential Effects of Intervention Timing on COVID-19 Spread in the United States. *medRxiv*. *https://www.ncbi.nlm.nih.gov/pubmed/32511526*.

21. Friston KJ, *et al.* (2020) DYNAMIC CAUSAL MODELLING OF COVID-19. *https://arxiv.org/abs/2004.04463*.

22. Chao DL, Oron AP, Srikrishna D, & Famulare M (2020) MODELING LAYERED NON-PHARMACEUTICAL INTERVENTIONS AGAINST SARS-COV-2 IN THE UNITED STATES WITH CORVID. *https://www.medrxiv.org/content/10.1101/2020.04.08.20058487v1.full.pdf*.

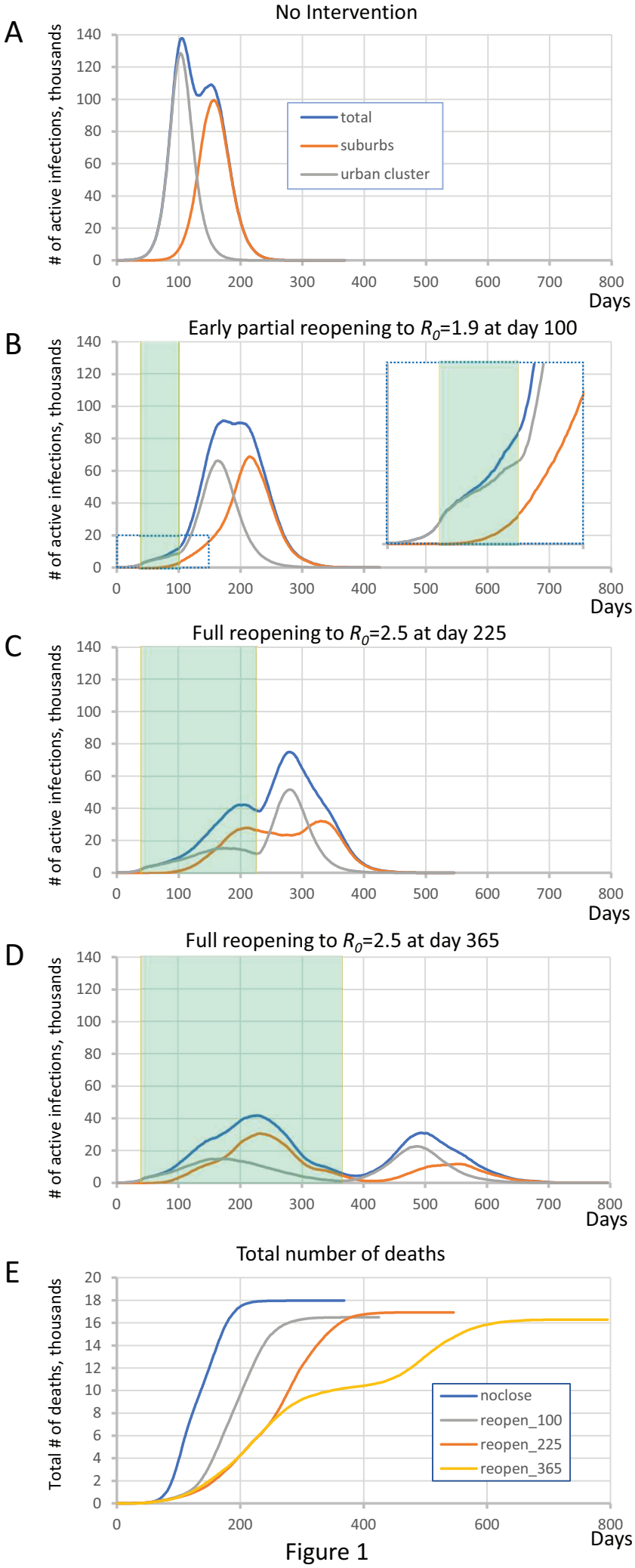
A
No Intervention
# of active infections, thousands
total
suburbs
urban cluster
Days
B
Early partial reopening to R_0=1.9 at day 100
# of active infections, thousands
Days
C
Full reopening to R_0=2.5 at day 225
# of active infections, thousands
Days
D
Full reopening to R_0=2.5 at day 365
# of active infections, thousands
Days
E
Total number of deaths
Total # of deaths, thousands
noclose
reopen_100
reopen_225
reopen_365
Days

Figure 1

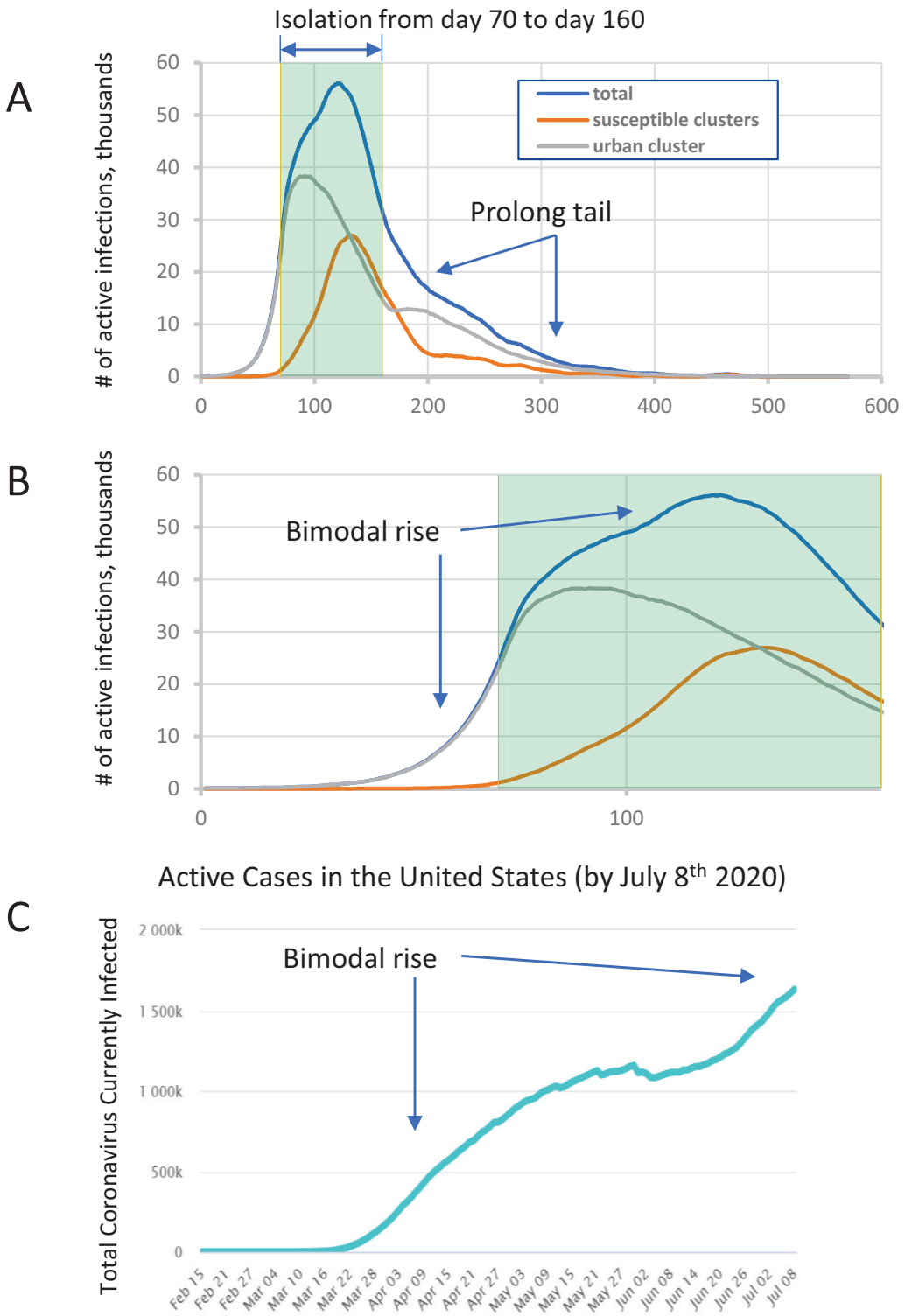
A
Isolation from day 70 to day 160
# of active infections, thousands
total
susceptible clusters
urban cluster
Prolong tail
B
# of active infections, thousands
Bimodal rise
Active Cases in the United States (by July 8th 2020)
C
Total Coronavirus Currently Infected
Bimodal rise

Figure 2

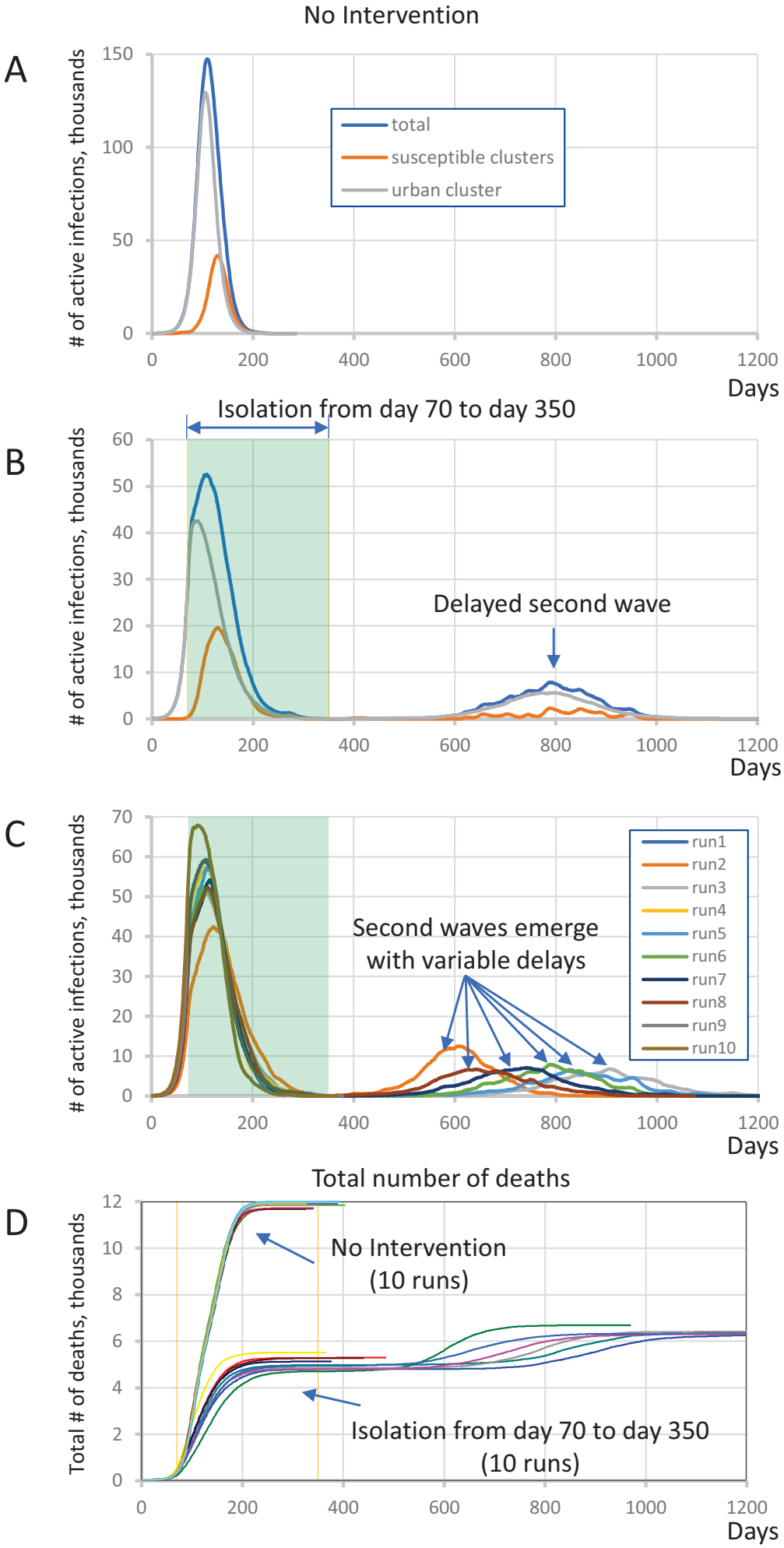
A
No Intervention
# of active infections, thousands
total
susceptible clusters
urban cluster
Days
B
Isolation from day 70 to day 350
# of active infections, thousands
Delayed second wave
Days
C
# of active infections, thousands
Second waves emerge
with variable delays
run1
run2
run3
run4
run5
run6
run7
run8
run9
run10
Days
D
Total number of deaths
Total # of deaths, thousands
No Intervention
(10 runs)
Isolation from day 70 to day 350
(10 runs)
Days

Figure 3

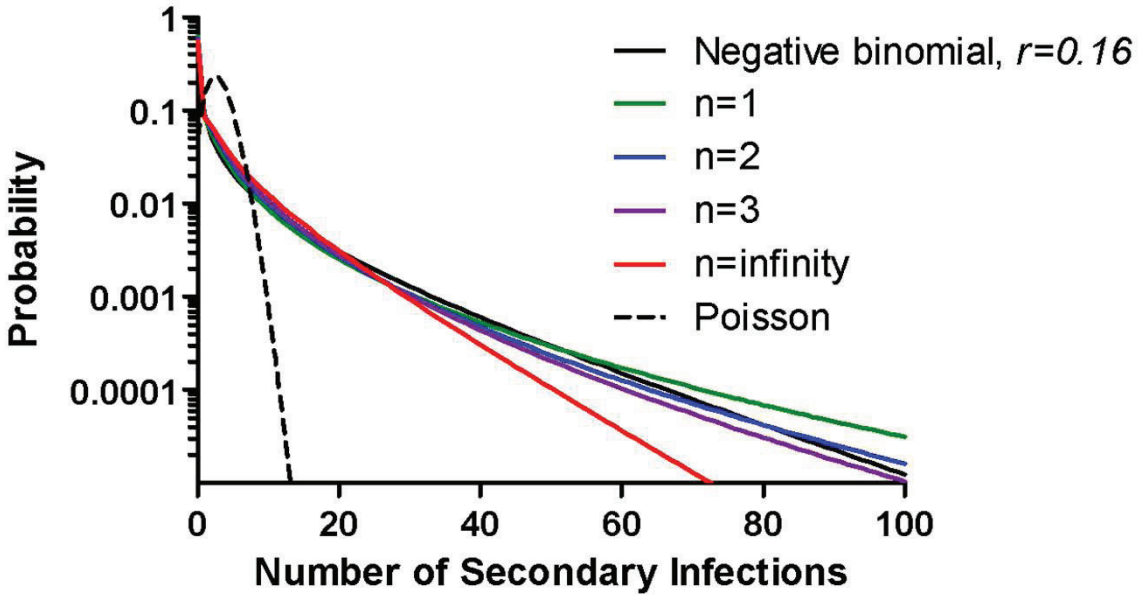
Negative binomial, r=0.16
n=1
n=2
n=3
n=infinity
Poisson
Probability
1
0.1
0.01
0.001
0.0001
0
20
40
60
80
100
Number of Secondary Infections

Figure 4