medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Impact of Superspreaders on dissemination and mitigation of COVID-19

Kim Sneppen

University of Copenhagen, Niels Bohr Institute, 2100 København Ø, Denmark

Lone Simonsen

University of Roskilde, Department of Science and Environment, 4000 Roskilde, Denmark

Abstract

Background: The severity of the Covid-19 pandemic has led to the use of extreme control measures which has halted the spread, but at enormous socioeconomic cost. We set out to explore whether a model that includes 'superspreader' could explain aspects of mitigation strategies that explain the dramatic effect.

Methods: We developed an age structured agent-based model that includes persons who spread the disease far more widely than others, acting in a society in which transmission occurs in three sectors: home, school/work and a third category representing other social contacts. We employ the model to study the impact of lock-down when superspreaders are present.

Findings: As expected, without mitigation imposed, the inclusion of superspreaders in the population does not change the epidemic trajectory. However in a structured society we find that superspreaders made a substantial difference. Our simulations demonstrate that workplaces and schools may open without much effect on the epidemic, as long as the other social contacts are drastically limited.

Interpretation: The recent observations of profound pandemic control that cannot be captured by standard SEIR models, require disease transmission models that consider superspreaders. We have found that transmission can be controlled simply by limiting contacts such as public transportation and

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

large events. Indeed, eliminating superspreader opportunities can uniquely explain the success of Sweden's relaxed approach and the moderate lockdown used in Denmark.

Funding: KS received funding from the EU Horizon 2020 research program under ERC grant No. 740704, and LS from the Carlsberg Foundation in Denmark.

Keywords: SARS-CoV-2, coronavirus, COVID-19, Agent based model, superspreader.

---

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Impact of Superspreaders on dissemination and mitigation of COVID-19

Kim Sneppen

University of Copenhagen, Niels Bohr Institute, 2100 København Ø, Denmark

Lone Simonsen

University of Roskilde, Department of Science and Environment, 4000 Roskilde, Denmark

Introduction

The emerging SARS-CoV-2 virus that is causing the COVID-19 pandemic which emerged in late 2019 has lead countries around the world to an unprecedented lock-down strategy. We have seen that these lock-downs work extremely well in terms of halting the spread, at enormous socioeconomic cost. However, we do not know what aspects of the mitigation efforts are causing the effect. As many countries are already in the process of opening up, the relative contribution of different aspects – reducing contacts at home, schools, workplaces and other sectors of society – should be better understood.

For this pandemic, a pattern of "superspreader" events has been documented to occur [1]. For example, as South Korea reopened the country carefully recently, a single infected person visiting a night club led to at least 50 new infections. In the US, a 2,5 hour long choir rehearsal tragically led to 75% of the participants getting infected and several deaths. These, as well as outbreaks in prisons and hospitals and following carnival parties are vivid examples of SARS-CoV-2 superspreader events that illustrate the importance of these events in the spread of this novel virus. It is also reminiscent of documented superspreader events in the dissemination of SARS-CoV in the 2003 outbreak in Asia and Canada. One recent study estimated that 80% of infections may be caused by 10% of the infected population, pointing to transmissibility (R_0) being highly heterogeneous [2]. Thus, the practice of

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

relying on an average R_0 in dynamic disease models can obscure considerable individual variation in infectiousness when extreme transmitters are included [3, 4, 5, 6]. It is also understood that such heterogeneity will not affect the epidemic spread by much in standard dynamic SEIR models. One might ask why are we often seeing superspreading events as a precursor to outbreaks? Also, why do we see marked heterogeneity in the geographical distribution and intensity of outbreaks across countries and regions?

Such heterogeneity in transmission risk is a well understood phenomenon in infectious diseases. In 2003 Riley et al. pointed out that 'superspreading events' could explain 80% of transmission events for a wide variety of diseases, thus being a defining characteristic of many epidemics, including vector-borne diseases due to large variability in parasite density in blood in infected individuals [7]. In the 2014 Ebola outbreak in west Africa, superspreader events are thought to have played a key role in sustaining onward transmission of the epidemy [8]. Overall, this means that certain individuals infect far greater numbers of persons than the average.

For SARS-CoV-2 the observed patient-to-patient variation in virus load appears to be several orders of magnitude [9]. Similarly, observed infectivity varies greatly [2]. Therefore we choose to focus on the variation in personal specific infectivity, while assuming that there is no variation in individual exposure.

We ask here whether the phenomenon of superspreading might play an important role in our ability to control the ongoing COVID-19/SARS-CoV-2 pandemic with non-pharmaceutical mitigation strategies. To investigate the effect of heterogeneity in various spaces we needed to modify our disease transmission model. One can easily do this in a agent-based model in which each person can be assigned individual infection properties. Our findings suggest a rating of the importance of limiting spread in the home, the school- or work place versus in other parts of the social environment where many new contacts are encountered and superspreaders events may occur.

Methods

We developed an age stratified version of the agent-based model previously described in [14], see Fig. 1. The model is simulated in 6 hour timesteps, during which each infected agent has a probability of infecting another agent. In the structured model infections attempts are weighted equally between the three social sectors outlined in Fig. 1. When simulating mitigation

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

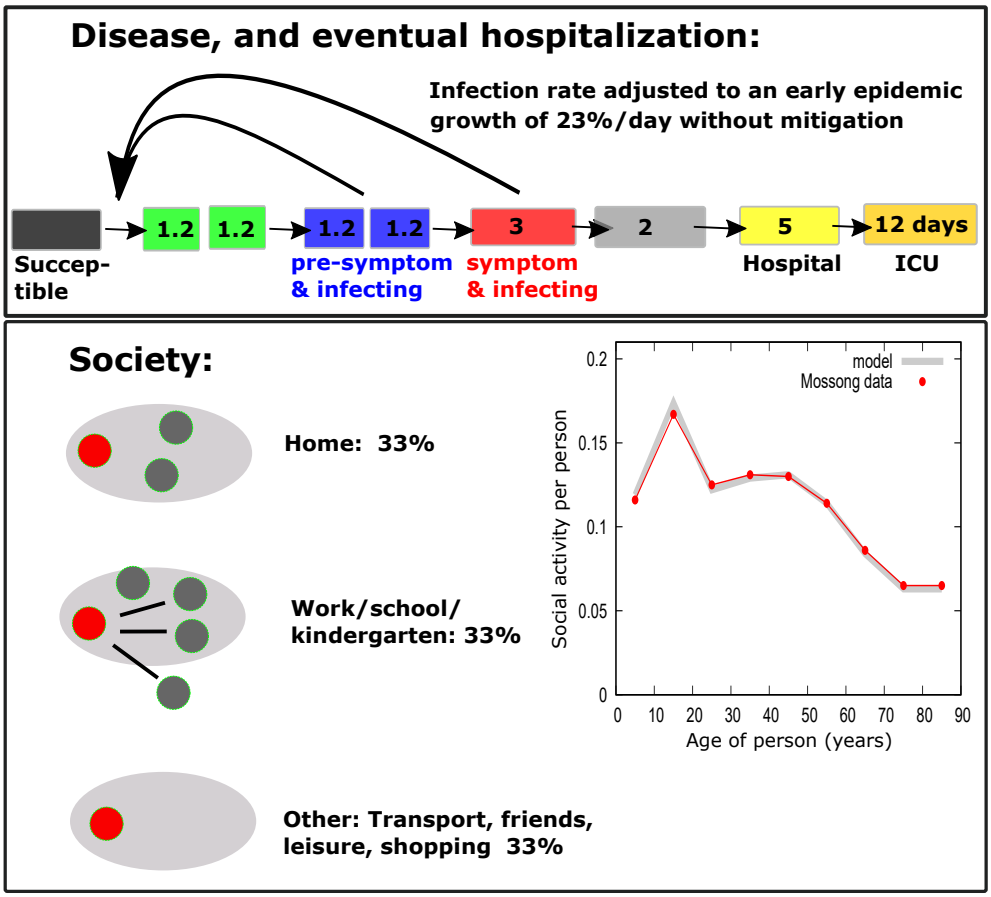


Figure 1: **Agent based model with social structure**. The top panel shows various stages of the disease, each simulated with a corresponding rate. People can infect from blue and red stages, and their relative duration corresponds to 45% pre-symptomatic infection [10]. The bottom panel illustrates the schematic society considered in our model, where interactions of each individual are related to three sectors: contacts in "Home", "Work/school" and "Other" is adjusted to occur equally often (idealized simplification from ref. [11, 12]). Home is the smallest unit (average cluster of size 2.1), while Work/school is larger (each person has an average of 10 connections at work and 26 at school). All ages between 20 and 70 are mixed in the work places. For each agent the Home and Work/school interactions are fixed throughout the epidemic. The "Other" group is treated as consisting of random contacts. are further assigned a social activity as function of age, as shown in the insert, adjusted to fit the data from ref. [13].

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

scenarios, the infections in one or more sectors are reduced. Infections from superspreaders are simulated by repeating their infectious activity 50 times within the 6 hour time window. When there is 10% of these superspreaders they cause between 75 and 85% of all infections.

In Figs. 2-4 we simulate 200.000 agents, while Fig. 5 considers one million. Simulations are started with 50 infected people. Throughout this work we calibrate the infection rate to fit an exponential growth of the unconstrained epidemics of 23%/day, as reported by [15]. A more complete description, that also includes the age dependent hospitalization and intensive care unit (ICU) probabilities, can be found in the supplement.

We finally explore the real-life situations in Denmark and Sweden (Fig. 5), now fitting the model to daily patients using data from the Danish [16] and Swedish [17] health authorities. The Swedish data were averaged over a 5 days time window.

Role of the funding source: The funders of the study had no role whatsoever in the study design, data collection, data interpretation, or writing of this paper. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

The agent-based perspective on the spread of an epidemic in a population allows us to address different types of questions and assess different mitigation strategies than what can be done with the more commonly used SEIR models. Naturally an agent based model resembles its corresponding SEIR variant when the agents are simulated without any social structure.

Model without social structure

First we modeled a COVID-19 like epidemic without "superspreaders" (SS) and social structure. Figure 2a) at 100% contact rate corresponds to such an idealization. This closely resembles the corresponding predictions for a SEIR model without social structure. When we reduced the contact rate so that only 50% and 25% of contacts remain, then the Covid-19 epidemic was strongly mitigated or halted, respectively. The largest reduction was enough to bring the effective R_0 (Re) below 1.

Next we introduced "superspreaders" into the model to examine the effect of this phenomenon. The simulations in Fig. 2b assume that 10% of the agents are 50 times more infectious than the rest. Using the model we find

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

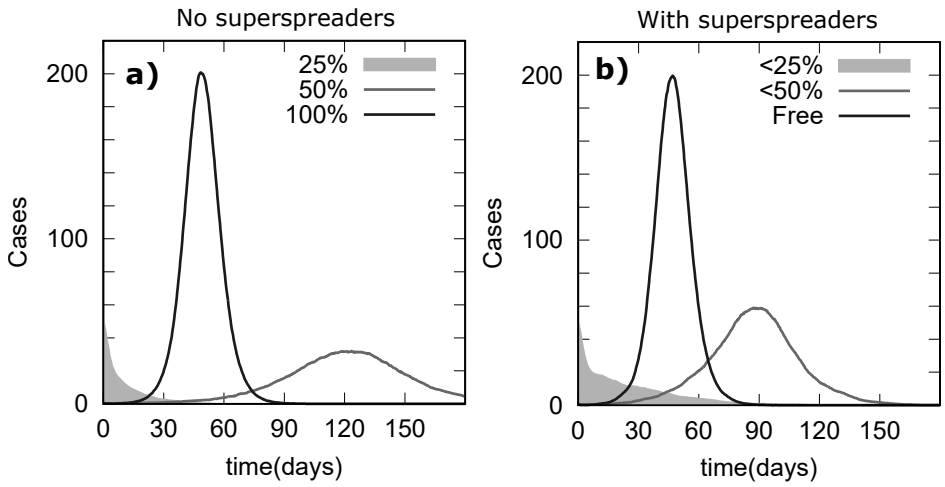


Figure 2: **Effects on epidemic trajectory of the proportion of normal contacts.** Panel **a**: Classic (SEIR) type model where all agents infect at same rate ($R_0 = 2.7$). The 100% curve marks the unmitigated epidemic, the other two curves mark two levels of lock-down in which 50% and 25% of contacts, respectively, remain. Panel **b**: Same random contacts as in a), but now assuming that 10% of infected agents are superspreaders, that transmit 50 times more effectively than others. Here the lock-down only imposes a limit on the maximum number of contacts, rather than a restriction in the number of contacts for all agents.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

that 85% of all transmission events are from these highly infectious agents. Although the unmitigated epidemic progresses in a qualitative way similar to the classical model in panel a), one is now able to halt the epidemic by only reducing the maximal number of encounters. In fact, if only 25% of all contacts remain for the superspreader, one can halt the epidemic. Thus, when they are prevented from acting as such, a dramatic level of mitigation can be achieved while everybody else can live their normal lives. This is because, as superspreaders infect many times more than the other agents, a reduction in the maximum number of contacts per agent only affects them. Of course, this scenario is difficult to achieve in real life since we do not know the identify of the superspreaders.

Model with social structure

Next we considered a structured society, by introducing into the model three sectors where interactions can take place (Fig. 1). This will allow us to introduce an indirect cap on the maximal social activity without knowing the identity of the superspreaders.

In Fig. 3 we show simulated epidemics for a socially structured model without superspreaders. The infection rate is again adjusted to fit the overall growth rate of 23%, and we observe an epidemic that resembles the basic model in Fig 2a). We examined, one at a time, the effect of lock-downs on each of the three contact sectors ("Home", "Work/school", and "Other") (see Fig. 3b-d). In panel b), when disallowing any social contact in the "Home" sector, we find that the epidemic peak size and ICU usage are reduced to 60% of their non-mitigated value. Next, when closing down all contacts in the "Work/schools" sector, the peak reduced to 40% of the unmitigated size; ICU usage is lower due the fact that the elderly are not being directly affected by this measure. Mitigating "Work/school" contacts has a larger impact than reducing "Home" contacts because the associated network is greater. Finally, when closing the "Other" sector the epidemic peak again reduces to 40% of its original size but we found an even larger reduction in ICU utilization (Fig 3d). This reflects the fact that the elderly are largely connected to the society through "Other" contacts.

We found that the model predicts final epidemic attack rates of between 80% and 100% of the population, with or without mitigation in the "Home" and "Work/school" sectors (Figure 3e). However, when closing down the "Other" sector, the attack rate is further reduced and contagion to the elderly is nearly eliminated.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

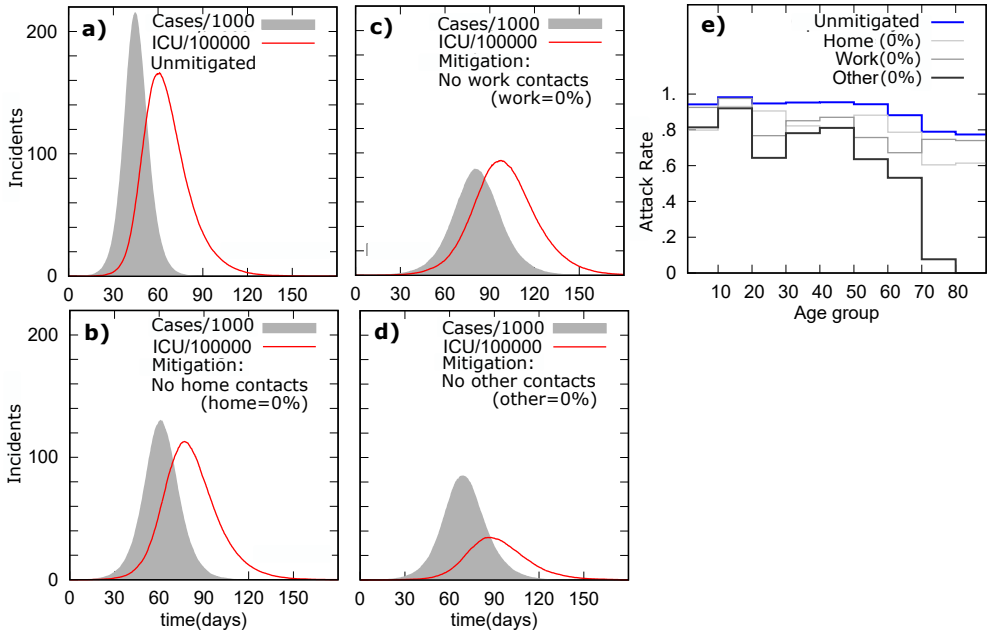


Figure 3: **Agent based model with three sectors of contacts (social structure) but no superspreaders. a)** Unmitigated epidemic scenario (gray curve) using Norwegian age-specific data for ICU utilization (red curve). **b)** Model simulation without Home contacts, and **c)** without Work/school contacts, and **d)** without "Other" contacts, that is, contacts outside Home and Work/school. A three quarters reduction in peak daily ICU usage is achieved when limiting the "Other" contacts, that is, contacts outside home and work/school. **e)** Final population attack rate as a function of age for each of the four simulations in panel a,b,c,d. We find that mitigating "Other" contacts results in dramatic reductions in ICU. This is explained by the fact that the older generations having contact to the remaining society largely through contacts of the "Other" type.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

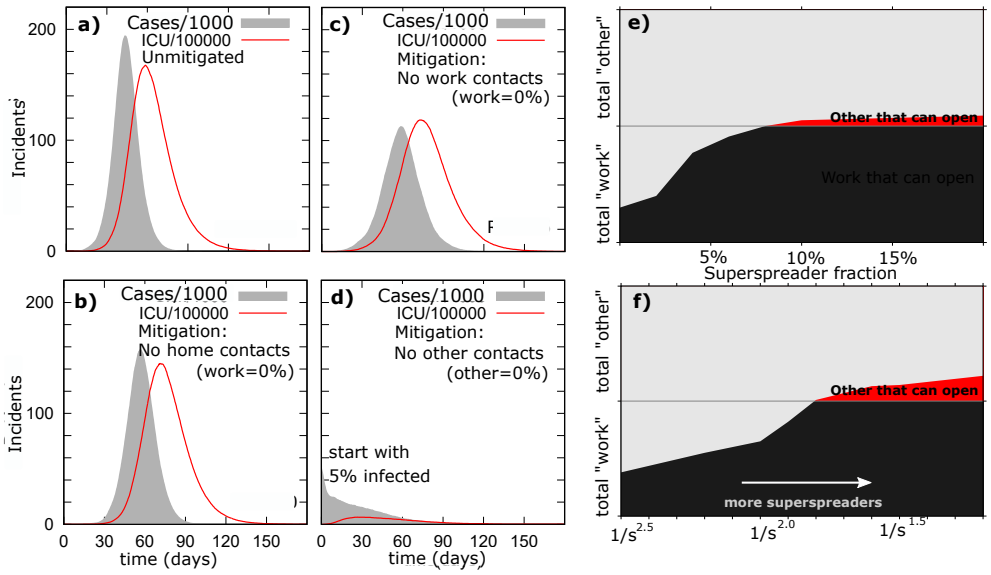


Figure 4: **Agent based model with social structure and superspreaders.** Panels a-d) recapitulate the simulations done for the social-structured model of Fig.3. **a)** Unmitigated epidemic scenario (gray curve) using Norwegian age-specific data for ICU utilization (red curve). **b)** Model simulation without Home contacts, and **c)** without Work/school contacts, and **d)** without "Other" contacts. **e)** Sensitivity analysis where the light blue area marks the fraction of "Work/school" and "Other" sectors that need to be isolated to end the Covid-19 epidemics. Here the superspreaders still infect a factor of 50 times more than other agents but their number is varied between 0 and 20%. **f)** Sensitivity analysis of model with a distribution of individual infectivity drawn from $p(s) \propto 1/s^{\gamma}$, $s < 1000$ with γ varying between 1.2 and 2.5. In the sensitivity analysis we used a population size of one million.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Finally we introduced superspreaders into the socially structured model. Fig. 4 recapitulates the simulations shown in Fig. 3, but this time with 10% of the population being superspreaders that are 50 times more infective. These 10% superspreaders caused 75% of all infections.

Figure 4 panels b) and c) show that in this superspreader model the elimination of Home or Work/school contacts, respectively, had less influence on the epidemic peak size than in the model without superspreaders (see Figure 3). Even if one eliminated all contacts in both "Home" and "Work/school" sectors, the peak epidemic size would be substantial (not shown), corresponding to an R_e value of about 1.7. In contrast, Fig. 4 shows that the epidemic is halted by eliminating "Other" contacts.

We thus conclude that superspreaders matter greatly in terms of the effect of mitigation strategies. Epidemics driven by a small proportion of highly infectious people are less sensitive to reductions in close contacts at Home or Work/school, but are highly affected by changes in their random contacts represented by the "Other" sector.

Sensitivity analyses of infection heterogeneity

Next, we asked ourselves whether this striking prediction was sensitive to assumptions about the frequency and intensity of superspreaders. To simplify our analysis, we considered lockdown in each sector in prioritized order: first "Other", then "Work/ school" until the epidemic was contained. We do not consider "home" lockdown as a realistic option.

We first allowed the superspreader fraction to vary between 0% and 20% while retaining their 50x greater infection intensity. In Figure 4e), the reduction in contacts needed to halt the epidemic is indicated in light grey. Thus, the boundary of the light grey region marks epidemic control ($R_e = 1$). It can be seen that when the fraction of superspreaders decreases below 8% the epidemic cannot be stopped by closing down the "Other" sector alone. Furthermore, the required lockdown increases substantially when the superspreader fraction is small. Below ∼4%; the superspreaders only contribute little and are no longer important.

We next explored the effect of considering the infection intensity distribution rather than an either-or scenario (Panel f). Each person was at the start of the simulation assigned a particular infectivity drawn from a power law distribution, limited by an upper limit of 1000x of the minimum. When this distribution is wider than about $1/s^{1.8}$, corresponding to 25% of superspreaders causing 80% of the infections, the Covid-19 epidemic can be mitigated

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

by a full lockdown of the "Other" sector. If the distribution is wider, for example $1/s^{1.5}$, corresponding to 16% of infected causing 80% of infections, one can stop the epidemic while still allowing 14% of the "Other" contacts.

Simulations of real life lockdown in Denmark

To further explore the importance of superspreaders, we applied our model to the real life situation that took place in Denmark after a lock-down on March 13. Here, locking down a population of 5,8M was achieved early in the epidemic, and was followed one month later by a limited reopening in the Schools/work sector. After an additional 40 days, a second more extensive reopening took place. Throughout this whole period, the population continued general social distancing and hygiene measures.

First, we considered the model without superspreaders (Fig. 5a) and assume that the first lockdown imposed a reduction in the Work/school sector of about 25% of the pre-pandemic level, and that a month later this is eased to about 50% of the pre-pandemic level. Then in the second reopening we assumed all restrictions on the Work/school sector were dropped. We furthermore assumed that the "Other" sector would be reduced and reinstated in same proportions first down to 25%, then up to 50% and finally back to the original 100%). Even so these reductions are not enough to reproduce the observation that the epidemic virtually came to a halt after the initial lockdown. In the simulations in figure 5a), in order to fit the observed inpatient data we introduced an additional social distancing factor of 1/3 in both the Work/school and "Other" sectors. This additional factor 1/3 can be understood as the effect of the general hygiene measures and social distancing. Now the model traced the observed inpatient data (from early March to mid-May). Importantly, this was not enough to halt the epidemic in a scenario of complete reopening, in which case we predict a large epidemic surge (Figure 5a). These predictions are very sensitive to the fudge factor of 1/3 employed to account for the effects of social distancing and hygienic measures.

Next we introduced superspreaders into the model. We found that the trajectory is more stochastic (see 3 simulation with same parameters in Figure 5b)). With the presence of superspreaders it is sufficient to assume that the general distance/hygiene factor only applies to the "Other" sector with a factor 1/6. Thus one can explain the trajectory of the epidemic without imposing social distancing in the Work/school sector. This emphasizes that the "Other" sector reduction could be enough to control the epidemic i.e.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

making Re < 1. More realistic simulations should of course include a combination of these factors. Nevertheless, we have demonstrated a clear effect of having super spreaders in the model.

Focusing the reduction on the "Other" sector only allows us to interpret these results as the reduction that could be accomplished by closing down venues where many strangers meet (festivals, public transport, soccer games). Panel c) illustrates this effect of opening the "Other" sector to half the pre-pandemic level after the 2nd reopening: the epidemic takes off again.

The case of Sweden

Finally, we considered the unusual case of Sweden, a country famous for never having locked down their country, but rather depended on encouraging social distancing and hygiene procedures in the general population, as well as a ban on large gatherings (>50 persons), and recommendations to avoid non-essential travel within the country. Interestingly, Sweden has with this strategy been able to mitigate the epidemics so that the effective Re has been a little below 1 since mid-April [17].

In Fig. 5d) we introduced a reduction of contacts in the "Other" sector by a factor 1/6 early on (see red band in panel d). We found that this assumption in principle allows us to contain the epidemics and bring Re sufficiently down to start seeing effects of herd immunity. We emphasize that this herd-immunity can, according to our model, only function in an environment where the "Other" sector is small. Therefore, an eventual increase in this type of contact will probably lead to a second epidemics wave.

Our model can fit the observed epidemic in Sweden with a factor 1/6 reduction in "Other", or, alternatively, by a factor 1/4 reduction in both the "Work/school" and "Other" sectors (See Figure S1). This second alternative is rather unappealing, given that schools in Sweden are largely open.

Discussion

SARS-CoV-2 is an intriguing pathogen that has challenged our understanding of disease transmission already from its beginning, in Wuhan. How could the epidemic be so slow to spread in the epicenter in Wuhan, with an estimated growth of 10% a day [18], while it later expanded at a rate of 30% per day. Why did it spread so slowly in Japan, and why did it not yet cause a major epidemic in Africa, while causing disaster with growth in death rates of up to 40%/day in northern Italy and Spain. Are we missing something?

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

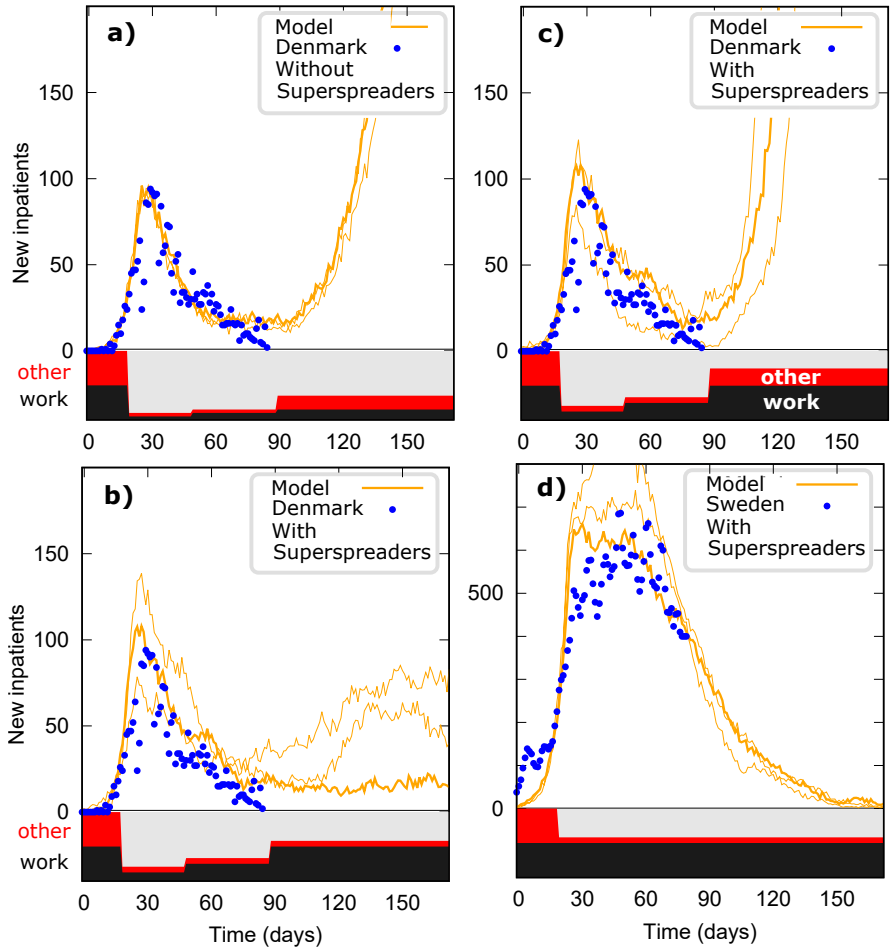


Figure 5: **Applying our social structure model with and without superspreaders, to epidemic data from Denmark and Sweden**. Using 1000000 agents and seeding with 100 infected persons, we fitted our model to Danish daily time series of new hospital patients (blue circles). The initial lock down is assumed to reduce contacts in the Work/school sector to 25%, and then the lock down is partially lifted one month later, to 50%. **a)** Simulation without superspreaders where the "Work/school" and "Other" sectors are reduced as indicated by the width of the black and red regions, respectively, shown at the bottom of the graph. **b)** Trajectory including superspreaders in the model, where "Other" is a factor 1/6 of pre lockdown value. **c)** Alternative outlook of simulation from b) if social distancing in the "Other" sector is only partly maintained after the lock down is fully lifted. **d)** Trajectory of the model with superspreaders compared to Swedish data. We model the Swedish case by setting the "Other" sector to 1/6 of pre-mitigation value. In each scenario we show 3 different stochastic trajectories with same parameters and initial infections. This illustrates that the epidemic when superspreaders are present can be stochastic, even when considering populations sizes of 1 million.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Superspreading is a phenomenon known to occur for coronavira outbreaks of SARS, MERS and now has also been described for COVID-19. Bursts of infections after a choir rehearsal in the United States and South Korean religious meetings are examples of such events. Was it the elimination of superspreading opportunities that ended the emerging SARS-CoV outbreak in 2003? Such questions are not easily addressable with the standard SEIR-type dynamic disease models.

In the present work we have presented an agent-based model to investigate the effect of superspreaders in in the evolution of the epidemics and in various proposed mitigation schemes [1, 2]. Since we do not know how to identify a superspreader individual, we chose to model superspreading in terms of a fixed 10% proportion of individuals transmitting 50-fold more effectively than others. Using this model we could demonstrate that the success of a lockdown and the effects of subsequent re-openings are highly dependent on this heterogeneity of infectivity in the population. The prevalence of 10% superspreaders in our base model causing about 75% of infections is in agreement with recent findings by others [2] where they estimate that 10% of individuals cause 80% of infections. In sensitivity analyses, we explored other possibilities of infection heterogeneity and found that at least half of all transmissions must be from superspreaders.

Assuming arbitrarily that interactions are evenly distributed in the three sectors – Home, Work/school and "Other" – we investigated the effect of superspreaders following a lockdown of such a socially structured environment. We found that in a mitigated epidemic superspreaders matter tremendously. Our agent-based model allowed us to compare the effect of mitigation in each sector separately. Most importantly we found that mitigating the "Other" sector can have a profound impact, even halt the epidemics. This we interpret to mean that limiting random social contacts at large gatherings and in public transportation is what effectively can control the COVID-19 epidemic.

As a caveat, in our simulations we relegated all non-familiar contacts to an "Other" sector, and assumed all other contacts (with known persons) occurred through fixed social networks at home and work/school. In reality part of the interactions in the "Other" sector is of familiar persons, for example friends and extended family. On the other hand, part of the interactions in the "Work/school" sector would involve unknown persons, such as casual contacts at cafeterias, conferences and workshops. Importantly, all of this is not too concerning, because what really matters is how many different people a superspreader agent meets during the period where it infects excessively.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

When we fit our model to actual observed Danish COVID-19 new inpatient data from early March to mid-May we found that our one-month forecast of the epidemic trajectory following the reopening of the country looked far more promising with than without superspreading, providing continued limits on contacts in the "Other" sector continued being applied (see Fig 5, b,c). It is mysterious that the epidemic did not begin again after the rather extensive phase one-reopening in mid-April; we suggest this is due to the continued ban on large gatherings and reduced use of public transportation. If this factor is relaxed in the future, we forecast a return to epidemic growth. In our simulations of the Swedish scenario (Fig. 5d), our superspreader model fits well with the observed data under realistic assumptions. We predict that if the current strategy is maintained, the epidemic will come to a halt within a few months, owing to few superspreader opportunities. Importantly, we can only explain the control of the Swedish epidemic with superspreaders in the model; without them one has to assume unrealistically large levels of general social distancing.

About 4 months into the COVID-19 pandemic we see a heterogeneous epidemic patterns across the world, in which some cities are devastated while other regions are nearly free of disease. This could be due to the action of superspreaders. An epidemic take-off may well require a superspreader because the R_0 is below 1 for the rest of the population. But if a superspreader is infected, the disease may spread to other superspreaders. Although hard to prove, it would not surprise us if the observed major differences across cities, regions and countries could be understood as a stochastic phenomenon modulated by a larger chance for the epidemic to percolate among superspreaders in large cities.

It is in this perspective that we propose that COVID-19 models used to describe the epidemic trajectory and forecast the effect of mitigation strategies and re-openings are inaccurate unless they include superspreaders. Without this element, such models will easily overestimate the epidemic size as well as the mitigation intensity needed in order to regain control.

We conclude that there is an urgent need of including heterogeneity and superspreading phenomena in our understanding of the present COVID-19 epidemic, and a enormous need for experimentally determining the abundance, mechanism and strength of interaction of the superspreader phenomenon.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Acknowledgements: KS thanks Raul Donangelo and Andreas Eilersen for enlightening discussions and corrections to the manuscript, and LS thanks Viggo Andresen for insights.

References

[1] Y. Liu, R. M. Eggo, A. J. Kucharski, Secondary attack rate and superspreading events for sars-cov-2, The Lancet 395 (2020) e47.

[2] A. Endo, S. Abbott, A. J. Kucharski, S. Funk, et al., Estimating the overdispersion in covid-19 transmission using outbreak sizes outside china, Wellcome Open Research 5 (2020) 67.

[3] R. M. May, R. M. Anderson, Transmission dynamics of hiv infection, Nature 326 (1987) 137–142.

[4] M. E. Woolhouse, C. Dye, J.-F. Etard, T. Smith, J. Charlwood, G. Garnett, P. Hagan, J. Hii, P. Ndhlovu, R. Quinnell, et al., Heterogeneities in the transmission of infectious agents: implications for the design of control programs, Proceedings of the National Academy of Sciences 94 (1997) 338–342.

[5] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Physical review letters 86 (2001) 3200.

[6] R. A. Stein, Super-spreaders in infectious diseases, International Journal of Infectious Diseases 15 (2011) e510–e513.

[7] S. Riley, C. Fraser, C. A. Donnelly, A. C. Ghani, L. J. Abu-Raddad, A. J. Hedley, G. M. Leung, L.-M. Ho, T.-H. Lam, T. Q. Thach, et al., Transmission dynamics of the etiological agent of sars in hong kong: impact of public health interventions, Science 300 (2003) 1961–1966.

[8] M. S. Lau, B. D. Dalziel, S. Funk, A. McClelland, A. Tiffany, S. Riley, C. J. E. Metcalf, B. T. Grenfell, Spatial and temporal dynamics of superspreading events in the 2014–2015 west africa ebola epidemic, Proceedings of the National Academy of Sciences 114 (2017) 2337–2342.

[9] T. C. Jones, B. Mühlemann, T. Veith, M. Zuchowski, J. Hofmann, A. Stein, A. Edelmann, V. M. Corman, C. Drosten, An analysis of sars-cov-2 viral load by patient age, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

[10] X. He, E. H. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, et al., Temporal dynamics in viral shedding and transmissibility of covid-19, Nature Medicine (2020) 1–4.

[11] P. Klepac, S. Kissler, J. Gog, Contagion! the bbc four pandemic–the model behind the documentary, Epidemics 24 (2018) 49–59.

[12] P. Klepac, A. J. Kucharski, A. J. Conlan, S. Kissler, M. Tang, H. Fry, J. R. Gog, Contacts in context: large-scale setting-specific social mixing matrices from the bbc pandemic project, medRxiv (2020).

[13] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, et al., Social contacts and mixing patterns relevant to the spread of infectious diseases, PLoS medicine 5 (2008).

[14] A. Eilersen, K. Sneppen, Estimating cost-benefit of quarantine length for covid-19 mitigation, medRxiv (2020).

[15] A. Remuzzi, G. Remuzzi, Covid-19 and italy: what next?, The Lancet (2020).

[16] D. Statens seruminstitut, Overvågning af covid-19, *https://www.ssi.dk/sygdomme-beredskab-og-forskning/ sygdomsovervaagning /c/covid19-overvaagning* (????).

[17] S. Folkhalsmyndigheten, Riket: skattning av det momentana reproduktionstalet, *https://www.folkhalsomyndigheten.se/contentassets/* 4b4dd8c7e15d48d2be744248794d1438/ (????).

[18] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong, et al., Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia, New England Journal of Medicine (2020).

[19] A. Aleta, D. Martin-Corral, A. P. y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N. E. Dean, M. E. Halloran, I. M. Longini, S. Merler, et al., Modeling the impact of social distancing, testing, contact tracing and household quarantine on second-wave scenarios of the covid-19 epidemic, medRxiv (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

[20] R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. Walker, H. Fu, et al., Estimates of the severity of coronavirus disease 2019: a model-based analysis, The Lancet Infectious Diseases (2020).

[21] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, Superspreading and the effect of individual variation on disease emergence, Nature 438 (2005) 355–359.

Research in context:

Evidence before this study:
Since the emergence of SARS-CoV-2 in China, superspreaders are emerging as a theme, as it also with SARS-CoV in 2003. Analyzing COVID-19 data, substantial individual-level variation in transmissibility has been inferred, and it has been estimated that 10% of infected persons cause 80% all infections.

Added value of this study:
We developed an agent-based socially structured model to simulate the effect of superspreaders on the epidemic, and in the context of country lockdowns and reopenings. Our simulations demonstrate that COVID-19 is effectively mitigated if one limits social activity where many strangers meet, as superspreaders will infect mainly among the many people they rarely meet.

Implications of all available evidence:
Modelling heterogeneity and super spreading is critical in terms of understanding observed COVID-19 epidemic patterns. Indeed, mitigating superspreading opportunities is an effective way to mitigate Covid-19 and prevent 2nd waves because the remaining population cannot sustain the epidemic. The superspreader phenomenon may also explain the observation of huge COVID-19 epidemic variability between different areas in a country.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Suplementary material: Impact of Superspreaders on dissemination and mitigation of COVID-19

Kim Sneppen.
University of Copenhagen, Niels Bohr Institute, 2100 København Ø, Denmark

Lone Simonsen.
University of Roskilde, Department of Science and Environment, 4000 Roskilde, Denmark

Extended Method Section

We use the agent-based model which was previously described in [14], see also Fig. 1 in main text. Other agent based models with similar structure can be found in [19]. Briefly, our model is an age stratified variant of model in [14], a modification that allow us to quantify the epidemics in terms of inpartient data and need for intensive care units (ICU).

The model is simulated in discrete time-steps of 6 hours, during which each infected agent has a probability of infecting another agent, potentially leading to transmission of the disease. Infection opportunities for superspreaders are simulated by repeating their infectious activity multiple times in each 6 hour time window. They do this repeatedly in either the"other" sector, or repeatedly within their fixed social networks.

As the time progress an infected agent may change state, with rates set by the duration of each interval in the top panel of Fig. 1 in the main text. As we further follow inpatients and the occupation of ICU, then the model also simulate the agents through the latter steps of the disease using rates from table 1. That is, each period is given in days, and then translated into a rate.

Throughout the paper we calibrate the infection rates to fit an exponential growth of the unconstrained epidemics to be the 23% per day that was reported by [15]. With parameters from Fig. 1 then this correspond to a R_0

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

of 2.7 in a randomly mixed population. For each simulation we measure R_0 from the number of infection attempts for a person that is sick (not counting attempts to a given contact after a successful infection).

Unless otherwise stated then simulations in Figs. 2-4 in main text use a population of 200000 and the epidemic is initiated at time zero by 50 infected persons.

The paper first consider a "well-mixed" population without any social or age structure. The results in Fig. 2, was thus obtained by selecting pair of agents randomly across the entire population for each infection attempt.

Figs. 3 and 4 simulate a population where each person is assigned 3 different types of contacts as outlined in Fig. 1. This choice of contacts represent a simplification of society into 3 equally weighted interactions: "Home", "Work" and "Other". In reality, the observed social activity have been estimated to be more tilted toward the "Home" sector and less toward to work/school type of contacts [11, 12, 13]. By making the relative weights equal it becomes simpler to compare them in our analysis.

The "Home" interactions are build from assigning each person a random home with an average size of 2.1. In these homes we constrain people above 20 to be within 1 age group from each other. We further assign children to parents that are 20-40 years older.

The "Work" interactions are modeled as clusters of average size 8 (Poisson distributed). Further we assign each person connections to two random persons outside this cluster. All ages between 20 and 70 are mixed in the work places. For persons under age 20, we use school classes of average size 24 and assign two teachers in older age groups ($[20-70]$) to each class.

Social activity per person across home and school is the same, apart from a re-scaling to the fit overall age dependent activity shown in insert of Fig. 1.

The interactions within both "Home" and "Work" are fixed throughout the epidemic, and are highly clustered. Noticeably, the "Home" clusters are by far the smallest and the different homes are only indirectly connected (through "Work" and "Other").

The "Other" group in principle contains all other interactions, including in particular all public and transport interactions. At each infection attempt, the interactions are chosen at random from the entire population.

The probability for selecting "Home", "Work" or "Other" are calibrated such that they occur equally often. Furthermore, the agents are assigned an age dependent social activity that is fitted to reproduce the contact data of

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

[13], see insert of Fig. 1.

Age group	Probability of hospitalisation	Probability of ICU given hospital	Population distribution
0-9	0%	5%	10.9%
10-19	0.013%	5%	11.9%
20-29	0.37%	5%	13.3%
30-39	1.1%	5 %	11.7%
40-49	1.4%	6.3%	13.6%
50-59	2.7%	12.2%	13.6%
60-69	3.9%	27.4%	11.7%
70-79	5.5%	43.2%	8.9%
80-	5.5%	70.9%	4.3%

Table 1: *Distribution of our simulated population, in age groups and with probability to end in Intensive Care Unit given hospitalization. The values for hospitalization and ICU occupancy are from the Norwegian health authorities: https://www.fhi.no/sv/smittsomme-sykdommer/corona/koronavirus-modellering/* or equivalently from [20] calibrated to an estimated fatality rate for SARS-CoV-2 of 0.3%.

Modeling superspreaders: The eventual superspreaders (i) are selected at the beginning of the simulation and assigned an individual activity (s_i). Once selected, a superspreader will engage in s_i infection attempts in the allocated time interval of 6 hours, where the probability of infecting someone is the same in each encounter. All other persons are simply assigned an $s_i = 1$. This way of introducing superspreaders in the model is mimicking the introduction of a person-specific reproductive number as it was done by Lloyd et al. in 2005 [21].

For the base model, and for all dynamic simulations in Fig. 2 and 4 we use $s_i = 50$ for 10% of the population.

In the sensitivity analysis (see Fig. 4e,f in main text) we first consider a range of prevalence of superspreaders (from 0% to 20% of the population). Then we consider a distribution of superspreader intensity, where s_i are drawn from a distribution $p(s) \propto 1/s^{\gamma}$, $s \in [1, 1000]$. The choice of the upper cut-off also influences the results, as the average infections per person increase with the upper limit when $\gamma \leq 2$.

The sensitivity analysis (Fig. 4e,f) was done with a population size of one million. In each case we first adjusted the infection rate such that the

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

unmitigated epidemic grows at a rate of 23%/day. Then we removed first a fraction of the other sector until epidemic was marginally contained. If this was not enough, we reduced the "work/school" sector until the epidemic was marginally mitigated. In the sensitivity analysis we confirmed that the result shown in Fig. 4e,f) did not change when changing the timestep from 6 hours to 2 hours.

Supplementary analysis of Sweden

In main text we present a brief analysis of Swedish Covid-19 epidemic. Sweden had a restrictive testing strategy, and their test of Covid-19 cases are closely associated to being hospitalized. There is also tested cases of Covid-19 among personal in health care institutions. We chose to treat their official count of cases as new hospitalizations in Fig. 5d as well as in the supplementary figure 1 shown below.

Supplementary figure 1 attempt to fit the swedish epidemic with and without superspreaders. This is done by assuming that the "other" and "work/school" sector use the same physical distancing measures. We believe this reflect the strategy employed by sweeden where schools and kindergartens was open, and most workforce worked as usual.

To reproduce the epidemic with the decline in new ceases from around mid-April, one need to restrict both work/school and other sector by about a factor 4. This should be compared with Fig. 5d), where the same level of mitigation is found by only restricting the other sector by a factor 6, and not assuming any reduction in the work/school sector. Thus only a little additional reduction in the other "sector", from 25% to 16% is enough to compensate for opening the "work/school" from 25% to 100%.

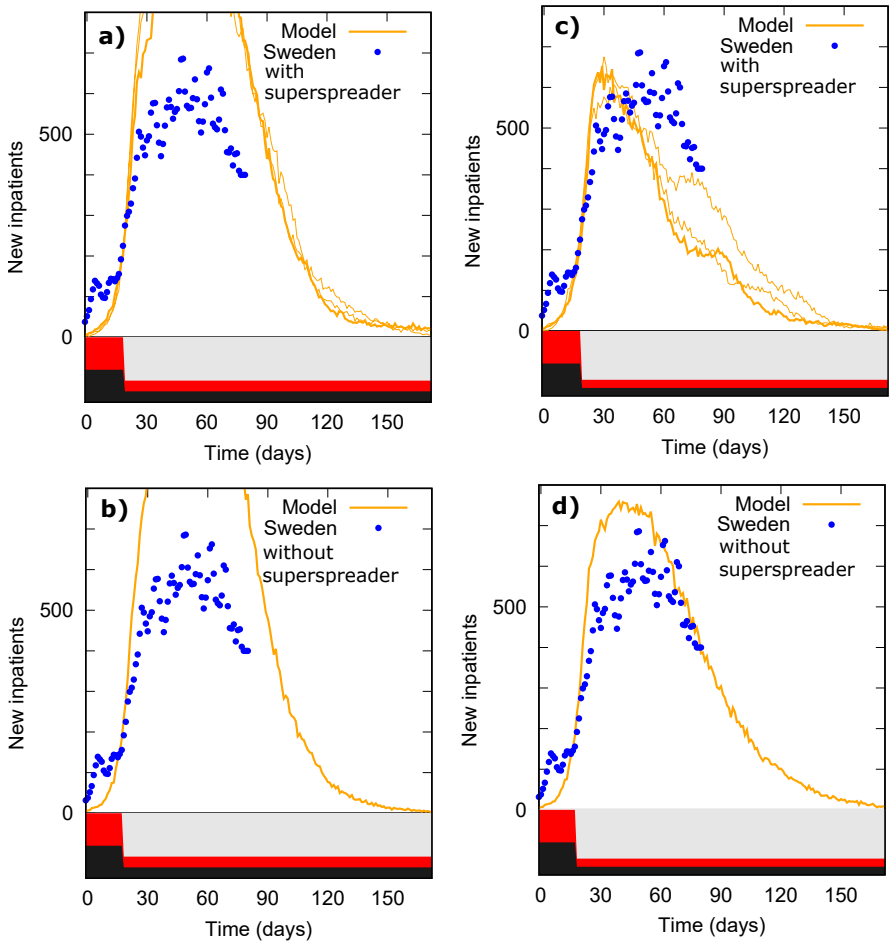
References

[1] Y. Liu, R. M. Eggo, A. J. Kucharski, Secondary attack rate and superspreading events for sars-cov-2, The Lancet 395 (2020) e47.

[2] A. Endo, S. Abbott, A. J. Kucharski, S. Funk, et al., Estimating the overdispersion in covid-19 transmission using outbreak sizes outside china, Wellcome Open Research 5 (2020) 67.

[3] R. M. May, R. M. Anderson, Transmission dynamics of hiv infection, Nature 326 (1987) 137–142.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .



Supplementary Figure 6: **Applying our agent-based social structure model with and without superspreaders, to Swedish COVID-19 data**. Comparison of model to time series of numbers of new inpatients/tested cases in Sweden (blue circles). In all simulations we here assume that the physical distancing in Sweden act equally in work/school sector and in other sector. **a)** Predicted epidemic without superspreaders if "work/school" and "other" is reduced by factor 3. **b)** Epidemic trajectory with superspreaders, $p(s) \propto 1/s^{1.5}$, where "work/school" and other is reduced by factor 3. **c)** Epidemic without superspreaders where both "work/school" and "other" is reduced by factor 4. **d)** Epidemic with superspreaders where "work/school" and "other" is reduced by factor 4. In each simulation we initiate the population with 100 infected individuals, and reduce work/school (black) and other (red) sectors when accumulated attack rate exceeds 1.3%.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

[4] M. E. Woolhouse, C. Dye, J.-F. Etard, T. Smith, J. Charlwood, G. Garnett, P. Hagan, J. Hii, P. Ndhlovu, R. Quinnell, et al., Heterogeneities in the transmission of infectious agents: implications for the design of control programs, Proceedings of the National Academy of Sciences 94 (1997) 338–342.

[5] R. Pastor-Satorras, A. Vespignani, Epidemic spreading in scale-free networks, Physical review letters 86 (2001) 3200.

[6] R. A. Stein, Super-spreaders in infectious diseases, International Journal of Infectious Diseases 15 (2011) e510–e513.

[7] S. Riley, C. Fraser, C. A. Donnelly, A. C. Ghani, L. J. Abu-Raddad, A. J. Hedley, G. M. Leung, L.-M. Ho, T.-H. Lam, T. Q. Thach, et al., Transmission dynamics of the etiological agent of sars in hong kong: impact of public health interventions, Science 300 (2003) 1961–1966.

[8] M. S. Lau, B. D. Dalziel, S. Funk, A. McClelland, A. Tiffany, S. Riley, C. J. E. Metcalf, B. T. Grenfell, Spatial and temporal dynamics of superspreading events in the 2014–2015 west africa ebola epidemic, Proceedings of the National Academy of Sciences 114 (2017) 2337–2342.

[9] T. C. Jones, B. Mühlemann, T. Veith, M. Zuchowski, J. Hofmann, A. Stein, A. Edelmann, V. M. Corman, C. Drosten, An analysis of sars-cov-2 viral load by patient age, 2020.

[10] X. He, E. H. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, et al., Temporal dynamics in viral shedding and transmissibility of covid-19, Nature Medicine (2020) 1–4.

[11] P. Klepac, S. Kissler, J. Gog, Contagion! the bbc four pandemic–the model behind the documentary, Epidemics 24 (2018) 49–59.

[12] P. Klepac, A. J. Kucharski, A. J. Conlan, S. Kissler, M. Tang, H. Fry, J. R. Gog, Contacts in context: large-scale setting-specific social mixing matrices from the bbc pandemic project, medRxiv (2020).

[13] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, et al., Social contacts and mixing patterns relevant to the spread of infectious diseases, PLoS medicine 5 (2008).

medRxiv preprint doi: https://doi.org/10.1101/2020.05.17.20104745; this version posted May 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

[14] A. Eilersen, K. Sneppen, Estimating cost-benefit of quarantine length for covid-19 mitigation, medRxiv (2020).

[15] A. Remuzzi, G. Remuzzi, Covid-19 and italy: what next?, The Lancet (2020).

[16] D. Statens seruminstitut, Overvågning af covid-19, *https://www.ssi.dk/sygdomme-beredskab-og-forskning/ sygdomsovervaagning /c/covid19-overvaagning* (????).

[17] S. Folkhalsmyndigheten, Riket: skattning av det momentana reproduktionstalet, *https://www.folkhalsomyndigheten.se/contentassets/* 4b4dd8c7e15d48d2be744248794d1438/ (????).

[18] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong, et al., Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia, New England Journal of Medicine (2020).

[19] A. Aleta, D. Martin-Corral, A. P. y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N. E. Dean, M. E. Halloran, I. M. Longini, S. Merler, et al., Modeling the impact of social distancing, testing, contact tracing and household quarantine on second-wave scenarios of the covid-19 epidemic, medRxiv (2020).

[20] R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. Walker, H. Fu, et al., Estimates of the severity of coronavirus disease 2019: a model-based analysis, The Lancet Infectious Diseases (2020).

[21] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, Superspreading and the effect of individual variation on disease emergence, Nature 438 (2005) 355–359.