Halting SARS-CoV-2 by Targeting High-Contact Individuals

Gianluca Manzo[1] and Arnout van de Rijt[2]

[1] *GEMASS CNRS and Sorbonne University, 59/61 Rue Pouchet, Paris 75017, France*
[2] *European University Institute, Via dei Roccettini 9, San Domenico di Fiesole (Florence) 50014, Italy*
Correspondence should be addressed to gianluca.manzo@cnrs.fr

Abstract: Network scientists have proposed that infectious diseases involving person-to-person transmission could be effectively halted by interventions targeting a minority of highly connected individuals. Could this strategy be effective in combating a virus partly transmitted in close-range contact, as many believe SARS-CoV-2 to be? Effectiveness critically depends on high between-person variability in the number of close-range contacts. We analyzed population survey data showing that the distribution of close-range contacts across individuals is indeed characterized by a small proportion of individuals reporting very high frequency contacts. Strikingly, we found that the average duration of contact is mostly invariant in the number of contacts, reinforcing the criticality of hubs. We simulated a population embedded in a network with empirically observed contact frequencies. Simulations showed that targeting hubs robustly improves containment.

Introduction

1.1 Most policy measures that are currently used or considered to contain the novel coronavirus SARS-CoV-2 are aimed at broad groups of citizens (children, elderly, contact professions) or categories of meeting places (schools, restaurants, airports) (Zhang et al. 2020). As such they leave large chunks of the workforce idling or operating below capacity for extensive periods (Meidan et al. 2020). Such mass measures are widely viewed as necessary but are costly. They have been shown to have a negative impact on national economic growth for several countries (Pichler et al. 2020, pp. 17–20) through both forced industrial inactivity and consumer behavior change (see Goolsbee & Syverson 2020).

1.2 At the same time, a fair amount of evidence now suggests that the spread of many person-to-person viruses is driven by a small fraction of individuals, sometimes referred to as "super-spreaders", who are responsible for the vast majority of secondary infections (James et al. 2007, Figure 1; Stein 2011; Wong et al. 2015; Sun et al. 2014. Many infected people appear to affect no one else. SARS-CoV-2 follows the same pattern. Estimates of the over-dispersion parameter K — which, unlike population-level estimates of the basic reproductive number, R_0, quantifies heterogeneity across individuals in their capacity to generate secondary cases (Lloyd-Smith et al. 2005), consistently suggest that between 10% and 20% of cases are responsible for between 80% and 90% of secondary infections (Endo et al. 2020; Bi et al. 2020; Adam et al. 2020; Miller et al. 2020). Individuals generating an unusually high number of secondary infections are thought to have played a pivotal role in SARS-CoV-2 outbreak in many countries (for an overview, see Kay 2020; for a case study, see Hamner et al. 2020. This suggests that if one could identify and protect super-spreaders, the virus may be controlled through focused interventions at lower overall cost.

1.3 The sources of high dispersion in individuals' capacity to generate secondary infections are not well-known. Some emphasize individual-level heterogeneity in infectiousness, such as differences in viral load, length of infection, and asymptomatic infection (Woolhouse et al. 1997; Galvani & May 2005; Cho et al. 2016). Others

relate superspreading to specific contextual settings in which infectious individuals infect many others at once — so-called "superspreading events" (James et al. 2007; Hodcroft 2020). One way this may happen is that the buildings in which events take place facilitate airborne transmission by dispersing small droplets from any one source to many targets (Morawska & Milton 2020).

1.4 Here we consider a third possibility, namely that the phenomenon of superspreading in SARS-CoV-2 has a network-structural basis. Some individuals may have jobs, living conditions, or social behavior that generate many more close-range contacts than others. Their status as "hubs" in the network of close-range contacts could render them disproportionately instrumental in viral propagation, as they are both more likely to contract the virus, and once they have it, pass it on to more others. In some cases, these high levels of contact derive from specific roles that these individuals play in an event, e.g., when a waitress or priest transmits a virus through serial dyadic contact. Without consideration of network structure, we may be inclined to blame the event and label it a super-spreader event post hoc. Yet, an appreciation of the network structure of close-range interactions within these events would suggest a targeted policy protecting high-contact individuals, where undifferentiated event-level policies would impose high costs on large groups (Manzo 2020).

1.5 Theoretical studies have shown that when networks are characterized by high interpersonal variability in the number of contacts and thus the existence of hubs, epidemics may occur at a much lower per-contact transmission probability (Barrat et al. 2008, ch. 9). Under these circumstances, targeting hubs with transmission-reducing interventions (e.g., protective measures, behavioral restrictions, testing and quarantining if positive, treatment, and eventually vaccination) may effectively control viral spread in the population at large (Dezső & Barabási 2002; Pastor-Satorras & Vespignani 2002).

1.6 The feasibility of this approach critically depends on the actual interpersonal variability in transmission-relevant contact. Early mathematical models of hub-targeting (Dezső & Barabási 2002; Pastor-Satorras & Vespignani 2002; Cohen et al. 2003) as well as recent applications of this approach to SARS-CoV-2 (Herrmann & Schwartz 2020a) assume a scale-free spreading network, while empirical networks often deviate from this assumption (Jones & Handcock 2003; Clauset et al. 2009; Stumpf & Porter 2012; Broido & Clauset 2019). Nevertheless, degree-targeting may still be an effective strategy in the fight against SARS-CoV-2, if close-range contact are highly skewed, with the majority of close-range contacts in society involving a small minority of individuals, as has been found for online contacts (Barabási & Albert 1999; Adamic & Huberman 2002; Vázquez et al. 2002) and sexual contacts (Liljeros et al. 2001; Trewick et al. 2013; Little et al. 2014).

1.7 The approach to network intervention through preferential targeting of hubs has been elaborated over the years, both with respect on how to measure hubs (see Kitsak et al. 2010; Montes et al. 2020) and how to reach them (see Rosenblatt et al. 2020) and it has also recently been applied to SARS-CoV-2 (Herrmann & Schwartz 2020a). However, this literature overwhelmingly relies on observed or simulated networks that are of questionable relevance for the diffusion of a virus such as SARS-CoV-2 for which direct close-range contacts aid droplet transmission (Mittal et al. 2020). Studies based on short-range Bluetooth data, showing high interpersonal variability in the volume of face-to-face interactions (Mones et al. 2018; Sapiezynski et al. 2019) are promising but for now, they only concern social encounters within small populations in single and specific social settings (such as primary schools, hospitals, academic meetings or university) (see Cencetti et al. 2020).

1.8 The objective of this paper is to assess the effectiveness of hub targeting versus undifferentiated interventions for controlling SARS-CoV-2 spread in networks with empirically calibrated frequencies of close-range contact. For this reason, we draw on nationally representative datasets containing information on close-range contacts in various meeting locations and the duration of each contact. Studies have shown that the spreading capacity of seeding hubs may be reduced when networks exhibit high clustering (see, in particular, Montes et al. 2020, Figure 3, panel 3). As a result, we also aim to assess whether the effectiveness of hub targeting vis-à-vis undifferentiated intervention on networks with empirically-calibrated degree is stable across different network features for which lack of appropriate data impedes calibration.

1.9 The paper is organized as follows: From the survey data we derive degree distributions for close-range contact on a country scale (Section 2). We then impose this empirical degree distribution on a synthetic social network with a tunable level of clustering (Sections 3.2-3.12). In this network, we introduce a virus with the main empirical features of SARS-CoV-2, and by an agent-based implementation of a SEIR model, we allow the virus to spread through the network under various transmission conditions (Sections 3.13-3.17). We have designed various ways of reaching the best-connected nodes (Sections 3.18-3.21), and calculate how the trajectory of the epidemic varies under these interventions (Section 4. From our simulation model, we have derived the hypothesis that interventions — such as vaccinations, medical testing, quarantining-if-positive, protections in high-risk professions, and informational campaigns — would be more effective when targeting hubs rather than random individuals (Section 5). We conclude by discussing implications and limitations of the study (Section 6).

Data Analysis

2.1 We draw on data from COMES-F, a survey conducted in 2012. In the survey a representative sample of about two thousand French residents report their close-range contacts (Bèraud et al. 2015). The COMES-F survey data have several features that make it attractive for our current purpose. First, compared with sensor-based digital data, representative survey data allow a comprehensive picture of contacts across social settings in the target population, thus generating a representative degree distribution. Second, among the major general population contact surveys conducted in Europe during the last decades (for a detailed comparative overview, see Hoang et al. 2019, pp. 730–733), COMES-F is the most recent, with the largest representative sample, based on paper diary, allowed respondents to report up to 40 contacts in their contact diaries, and, for each self-reported contact, recorded location, duration, and frequency. In addition, specific care was given to collect high-quality contact information for respondents aged 0-15. Finally, a recently conducted survey in six countries that we analyze in the Appendix A does not use the more thorough contact diary method, instead asking respondents for an estimate of the number of contacts (Belot et al. 2020).

2.2 Contact survey data are routinely employed by epidemiologists to build social contact matrices, i.e., average contacts between age groups by places like school, public transportation, or home (Prem et al. 2017). COMES-F have been frequently exploited in this way within compartmental models of SARS-CoV-2 spread in France (see, for instance, Di Domenico et al.; Roux et al. 2020; Salje et al. 2020a; Walker et al. 2020. In contrast to prior use of the data, we rely on the entire cross-individual heterogeneity of the observed distribution of close-range contacts. We implement this distribution in a social network model of disease propagation (Section 3), so that we can evaluate the effectiveness of interventions targeting high-contact individuals (Section 4).

2.3 The COMES-F survey was conducted in France during the first half of 2012. An initial sample of 24,250 was drawn from the French population excluding overseas territories through random-digit dialing of landline and mobile numbers. Using quotas for age, gender, days of the week and school holidays, 3,977 subjects who accepted to participate, were sent a contact diary to complete. 2,033 (51%) contact diaries were returned (participants' age and household size were used as sampling weights to maintain representativeness). In these diaries, participants were asked to keep track of all short-range contacts over the course of two full days, and report on sex and age of these contacts, meeting context, and contact duration. Respondents were explicitly instructed to consider as a short-range contact someone they talked to at less than two meters, possibly including physical contact. To relieve the reporting burden, respondents were asked to record in the contact diary no more than 40 close-range contacts. For respondents aged less than 15, an adult member of the household completed the diary.

2.4 Specific questions involved respondents currently in employment. In particular, they were asked whether they regarded their occupation as especially exposed to short-range contacts. This turned out to concern 257 respondents. These respondents had to indicate the average number of persons they estimated to meet each day because of their job. Should this number be higher than 20, those specific respondents were asked to enumerate only non-professional contacts when filling in the contact diary.

2.5 Throughout the paper we will refer to diary-based contacts and job-related extra contacts respectively to differentiate the two types of measurement processes. This is an important distinction. It draws attention to one limitation of contact-diary based data collection. For the vast majority of respondents, contact diaries only allow accurate estimates of the total volume of close-range contacts, but could not precisely distinguish the specific type of each of these contacts (family members, friends, co-workers, clients, unknown persons and so on). In this respect, COMES-F is not unlike other epidemiological diary-based contact surveys where only where a contact occurs is recorded, but not the precise nature of the contact (see, for instance, Mossong et al. 2008; Danon et al. 2013. Following previous analyses of COMES-F data Bèraud et al. 2015, we investigated separately diary-based contacts and job-related extra contacts and explain later how we combined them to calibrate our simulations.

Contact volume and duration

2.6 Contact distributions are shown in Figure 1. Through the contact diaries, the 2,033 individuals reported a total of 19,728 per-day close-range contacts (left panel). The median number of contacts was 8 whereas the average was approximately 9.5. Respondents reporting a number of close-range contacts greater than twice ($n =$175) or even three times ($n =$36) the mean were not rare. The distributional character of the right skew is captured by the distribution of close-range contacts above 19 not significantly deviating from the 95% confidence level from a power law with a scale parameter 5.1 ($n =$175)[1]. Variance and skew were more pronounced among

respondents who declared extra-job related contacts ($n = 257$) (Figure 1's right panel). Overall, they reported 14,971 additional contacts. For those contacts the median was 30 whereas the average was approximately 58. The tail of the distribution in Figure 1 above 17 did not significantly differ from a power law with a scale parameter 2.5 ($n = 190$).

2.7   The central tendencies of both distributions were consistent with those found in other diary-based contact surveys (see Hoang et al. 2019, pp. 727–728). In both cases, averages were clearly driven by a small fraction of individuals reporting high numbers of short-range contact. The feature of high distributional skew is visible in recent smaller-scale contact surveys conducted in China (Zhang et al. 2019, 2020). More particularly, the power law scale parameter estimates were similar to those found for the UK Social Contact Study (Danon et al. 2012, 2013). In Appendix A we show that this variability persisted within major demographic categories. Here, we describe rather the relationship between the volume of contacts and their duration. As we treated high-contact individuals as leverage for effective intervention in viral diffusion dynamics, it is important to examine this relationship: the superspreading potential of hubs could be reduced if contacts were on average much shorter.

2.8   From a social network perspective, one could expect a negative correlation. As time and cognitive resources needed to sustain independent social relationships are limited (see, for instance, Dunbar 2016), individuals with many contacts may on average spend less time per contact. If this were the case, then hubs may expose and be exposed to more people, although per contact could face less risk, reducing the criticality of hubs in the contagion.

2.9   Face-to-face, close-range contacts may partially escape this logic, however. Let us consider, for instance, a dance instructor who spends ten hours a day in a closed room giving private lessons to ten different small groups of four dancers during one hour. Such a respondent would typically declare to experience 40 contacts with skin touch per day for one hour and would probably add to this some contacts at home for more than one hour a day. Thus, this person would combine high contact frequency with high average contact duration. Large-scale surveys of social encounters indeed documented that these situations are frequent. It is common for people to be involved in different types of physically-closed social interactions — e.g., in family, friendship groups, classrooms, dance clubs, choir rehearsals, stadium visits and manual team labor–, sequentially or simultaneously, at different time of the day, sometimes with more than one person at a time (see, in particular, Danon et al. 2012, Figure 1a and Figure S2). When face-to-face interactions are at stake, individuals combine contact time across multiple and possibly simultaneous social interactions rather than experiencing them as independent and mutually exclusive events (like when one needs time and energy to build durable friendship or professional connections).

2.10   In the contact diaries, respondents additionally reported its approximate duration for each contact. Figure 2 shows two scatter-plots with on the y-axis the total duration of contact summed across all respondents' contacts (left panel), and the average duration of contact (right panel) by the number of reported contacts on the x-axis. For each plot, we showed the median y value for each x value and a LOESS curve. If time and cognitive resources were limited for social encounters as they seem to be for more durable social relationships, one would observe a line with zero slope in the left panel and a sharply declining curve inversely proportional to x in the right panel.

2.11   Strikingly, the actual empirical relationships were very different. In the left plot, we find a monotonically increasing relationship with a slope that remains similar in magnitude over the observed interval between 0 and 40, a pattern that matches previous studies in the UK (see Danon et al. 2012, Figure S3c; Danon et al. 2013, Figure 2). Correspondingly, in Figure 2's right plot we found little relationship between the number of close-range contacts and their average duration, with perhaps a slight decline in average contact length at high numbers of per-person contact. These results reinforce the criticality of hubs in spreading processes. The negative impact of higher numbers of contacts is not proportionally counteracted by the brevity of contact. For this reason, in our simulations, the dyadic transmission probability does not depend on the total number of contacts that an agent has.

2.12   Unfortunately, COMES-F does not contain information on contact duration for respondents' self-reported estimates of job-related extra contacts. As a consequence, we were unable to test the relationship between frequency and duration for the fraction of respondents reporting very high frequencies of contact (Figure 1's right plot). For this reason, we adopted a conservative approach and based the calibration of the synthetic network underlying our main analysis on diary-based contacts only (see Figure 1's left plot). In the Appendix C, however, we re-ran all analyses under a combination of diary-based contacts and job-related extra contacts in such a way that a fixed transmission probability assumption is still defensible. Results were consistent and robust across these different specifications.
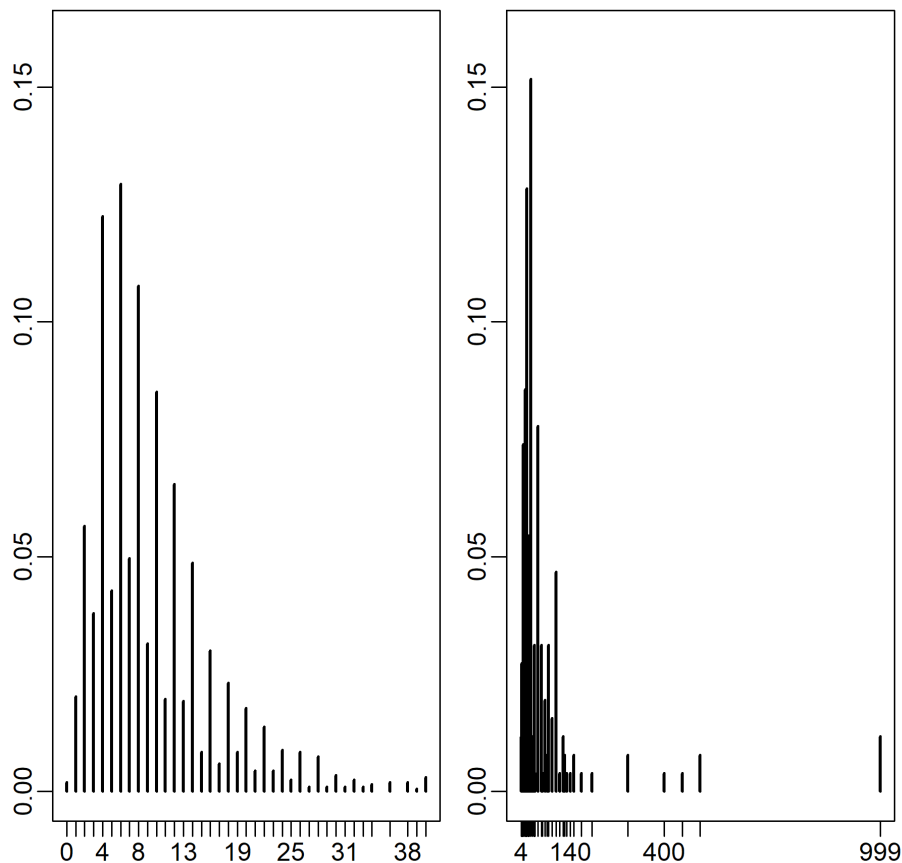
Figure 1: *Left:* Fraction of cases (y-axis) reporting a given number of close-range contacts (averaged over the two days) (x-axis) ($n =$2,033); *Right:* Fraction of cases (y-axis) reporting a given number of daily job-related contacts (x-axis) among respondents regarding their occupation as especially exposed to social contacts ($n =$257).

Model

3.1   Using COMES-F survey data we built an agent-based computational model in which the degree distribution of the synthetic network through which the virus diffuses is calibrated on the survey contact data (for other work using empirical network data in agent-based diffusion models, see Smith & Burrow 2018; Manzo et al. 2018. Our aim was to study the macroscopic consequences of cross-individual variability in close-range contact frequencies empirically observed at a country-level and assess whether this variability can be exploited for effective intervention in the ongoing epidemic. As such, we simulated a population the size of the COMES-F sample from which we eliminated four respondents who reported no close-range contacts.

Network construction and features

3.2   We connected 2,029 agents, each representing one respondent, according to two social network models. The first, which we will refer to as the "degree-calibrated" (DC) network, is the focus of our simulation. It is built to match the actual contact distribution and, at the same time, to tune network clustering, i.e., the propensity for two neighbors of a node to also be neighbors of one another. The second, which we will refer to as the "Erdős-Rényi" (ER) network, constitutes a benchmark to compare the dynamics and effects of interventions.

3.3   In the DC network model, agents were first given a degree (number of network ties) precisely equal to the number of close contacts per day reported by each respondent in the contact diary (see Figure 1's left plot). Then, to connect agents to one another, we adapted the configuration model, an algorithm that was proposed to generate random networks with arbitrary degree distributions (e.g. Jackson 2008, pp. 83–85). To avoid duplicate links and self-links while ensuring an exact match between each virtual agent and an empirical respondent, we considered source agents in descending order of the to-be-generated degree rather than in random order, and
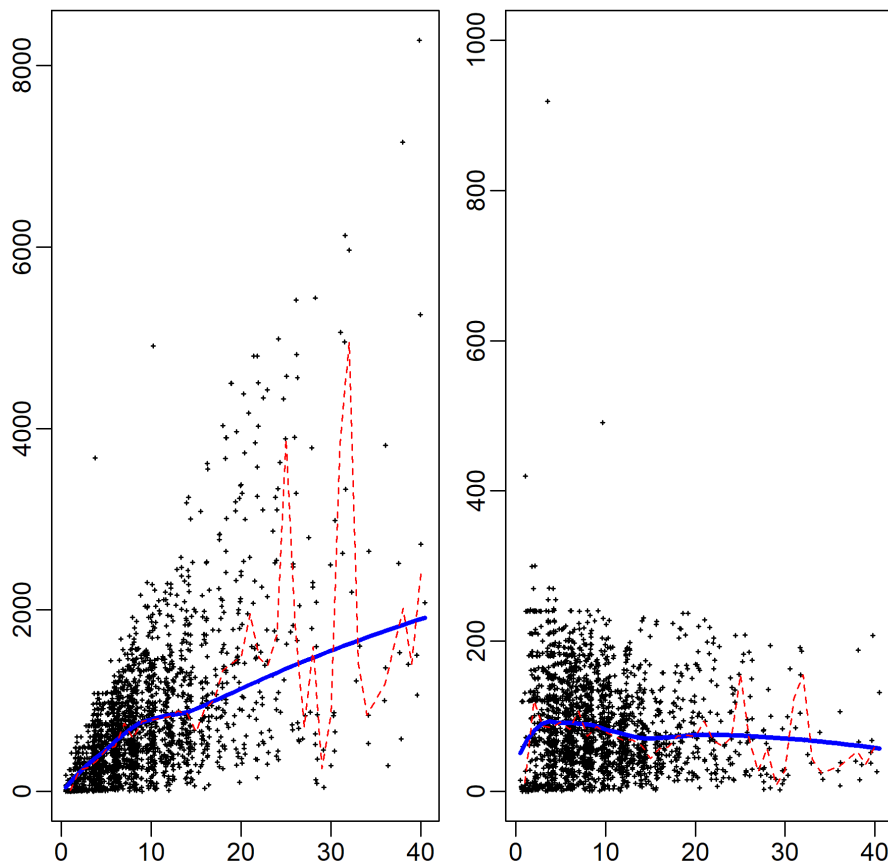
Figure 2: Left panel: Respondents' total daily contact duration (in minutes) (y-axis) as a function of daily # close-range contacts (x-axis). Right panel: Respondents' average contact duration (in minutes) (y-axis) as a function of daily # close-range contacts (x-axis). Points are jittered to avoid overlap. $N =$ 2029. Red dashed line: Median values of y-axis conditional on x-axis. Blue solid line: Local non-parametric regression curve (smoothing alpha parameter $=$ 0.5; polynomial degree $=$ 2) (fitted with R loess function). Total daily contact duration is computed as the sum (over all contacts) of the time the respondent declared having spent in each contact. Average contact duration is computed as total daily contact duration divided by daily # close-range contacts. Contact duration was recorded as a 5-category variable (1 $=$ < 5'; 2 $=$ 5'-15'; 3 $=$ 15'-60'; 4 $=$ 1h-4h; 5 $=$ > 4h): we consider the centroid of the interval (i.e. 2.5', 7.5', 22.5', 120', 240', respectively) to build the variables reported on the y-axis. Nota bene: Total daily duration may exceed 24 hours because many contacts happen simultaneously.

then randomly picked available destination agents. Each time a connection was made, the degree of the two newly connected agents naturally increased by one. As soon as an agent reached the to-be-generated degree, it was excluded from the search algorithm. We found this procedure to always converge, achieving the intended empirical degree distribution.

3.4 However, the configuration model is known to be able to generate only a limited degree of clustering, which is constrained within this model by the imposed degree distribution and network size (Newman 2003, p. 202). Clustering is a crucial network feature that is known to attenuate the spreading capacity of high-degree nodes (see for instance, Molina & Stone 2012). Recently, with particular reference to the COVID-19 crisis, Block et al. (2020) have presented simulation results showing that increasing local clustering of actors' ego-networks helps to mitigate the epidemic (see, in particular, Box 2).

3.5 Thus, in order to assess the robustness of targeting hubs under various levels of clustering, we modified our generative network algorithm in a simple way. In particular, as soon as a given node reached the required degree, before excluding it from the search algorithm, we went through all its neighbors and connected each of the ego's neighbor pairs with probability p.

3.6 Our second network model is motivated by the fact that in most epidemiological models it is still common to assume random mixing according to which social contacts are assumed to happen at random within certain

categories (Tolles & Luong 2020). From a network perspective, this amounts to postulating a random network where contact probabilities across individuals have little variability and the expected degree of each node is the average degree (Newman 2002; Barthélemy et al. 2005). We therefore also studied an Erdős–Rényi (ER) random graph with the average degree observed in the survey (again, we only considered diary-based contacts), as a benchmark distribution. This random network is characterized by low variability in contact across agents and low clustering.

3.7 Table 1 shows network statistics computed over 100 realizations of the two networks. By construction, the DC network reproduces the features of the actual degree distribution. In particular, the mean was higher than the median, which suggests right-skewness. By contrast, the degree distribution of the ER network had essentially an equal mean and median. The dispersion of the degree also strongly differed between the two networks, with the DC network exhibiting greater variation in the nodal number of links than in the ER network in the order of a doubling of the standard deviation (This difference was much larger when additional professional contacts are considered, see Appendix C).

	Average degree	Median degree	Stdev degree	Clustering coef	Deg-clust corr	Av path length	Diameter
Degree-Calibrated (DC) networks							
$p = 0$	9.72 (0.00)	8 (0.00)	6.56 (0.00)	0.01 (0.00)	-0.06 (0.01)	3.47 (0.00)	6 (0.00)
$p = 0.5$	9.72 (0.00)	8 (0.00)	6.56 (0.00)	0.43 (0.00)	-0.62 (0.01)	4.38 (0.03)	7.45 (0.50)
$p = 1$	9.72 (0.00)	8 (0.00)	6.56 (0.00)	0.57 (0.01)	-0.56 (0.01)	5.52 (0.09)	10.10 (0.59)
Erdős-Rényi (ER) network							
	9.65 (0.50)	9.72 (0.10)	3.11 (0.10)	0.00 (0.00)	0.00 (0.03)	3.60 (0.01)	6.06 (0.24)

Table 1: Topological features of the simulated contact networks (as a function of the clustering probability p for the DC network). Mean values across 100 network realizations (standard deviation in parentheses). Clustering coef = clustering coefficient; Deg-clust corr = Pearson correlation coefficient between nodes' degree and their clustering coefficient; Av path length = Average of the shortest path lengths; Diameter = Maximum of the shortest path lengths.

3.8 For the DC network, Table 1 also shows that how we modified the configuration model efficiently generated increasing levels of clustering as the clustering probability p increased. As expected, when $p = 0$, meaning that we did not force ego's neighbors to close triads, the DC network exhibited a very low level of clustering, essentially comparable to the ER network. By contrast, when $p = 1$, meaning that we forced the maximum number of links among a focal agent's neighbors, the level of clustering increased to 0.57, the maximum level we could reach given the structural constraints imposed by the actual degree distribution and the size of our synthetic population.

3.9 As COMES-F's contact-diaries did not contain information on possible close-range contacts among a respondent's contacts, we were not able to empirically calibrate the clustering coefficient of the DC network. To the best of our knowledge, only Danon et al. (2012, 2013) in the UK designed contact-diaries to collect information from which clustering of close-range social encounters could be estimated. Their data showed an average value around 0.46, with a considerable range of variation from approximately 0.07 to 0.7 depending on age category, meeting place, and distance from home (see, in particular, Danon et al. 2012, Figure 2b and Figure 2c; Danon et al. 2013, Figure 3 and Figure 5). These estimations rely on complex rescaling procedures that may overestimate (up to a factor of 1.8) the true level of clustering (see, on this point, Danon et al. 2013: SI, § 5.3). For this reason, we opted to study our model over the range of possible clustering levels generated by our algorithm. This range includes the 0.46 estimate.

3.10 The DC network was characterized by a negative correlation between the nodal degree and clustering, a correlation that becomes stronger as the overall level of clustering increases. This means that the higher the degree of a node the lower the fraction of ties among its neighbors. Thus, high-contact nodes spanned across the network more than they clustered together. This pattern was found on real-world close-range contact networks

in the UK (see Danon et al. 2012, Figure 2a). As discussed by Barabási (2014, pp. 232–237), this negative correlation between nodal degree and clustering is the statistical signature of the presence of community structure within the network, a topological feature that makes hub-centered interventions especially effective: attacking the hubs means interrupting (or slowing down) communication among the modules (p. 236). By contrast, the correlation between nodal degree and clustering in the ER network was virtually nil.

3.11 Finally, when $p = 0$, meaning zero probability of closing triads among a node's neighbors, in the DC network average path length and diameter were comparable to those of an Erdős–Rényi network with the same size and degree, consistent with that usually observed in pure scale-free models (see Albert & Barabási 2002, p. 74). Highly connected individuals were effective in bringing many parts of the network together. As we increased the clustering, the DC network's average path length and diameter increase, too, but remained low, achieving the combination of high clustering and reachability characteristic of small-world topologies (Watts & Strogatz 1998).

3.12 To sum up, the DC network displayed the actual long-tailed distribution of close-range contacts observed in representative data, while incorporating important topological features among which clustering and short path length that previous studies have shown to be consequential for the spread of disease. Our goal was to assess whether our proposed mitigation strategy of targeting hubs robustly increased epidemic control in these degree-calibrated networks[2].

Agent-based SEIR model

3.13 We modelled disease propagation through the *Degree-Calibrated* (DC) and *Erdős-Rényi* (ER) networks by building a stochastic agent-based implementation of a SEIR model (Martcheva 2015). The SEIR model is a type of compartmental model that has been previously applied to the COVID-19 outbreak (Brethouwer et al. 2020; Kucharski et al. 2020; Li et al. 2020; Prem et al. 2020). In particular, we followed recent empirical parametrization (see Salje et al. 2020a) to determine how agents unidirectionally move from being (S)usceptible, to (E)xposed, (I)nfectious, and eventually (R)ecovered. Each iteration corresponds to one day. The time it takes for an agent to move from one state to the next was calibrated accordingly.

3.14 Upon infection, agents first entered E where they stayed four days; during this period, they were not infectious (for this value, see Salje et al. 2020a, p. 10). They then moved to I where they become infectious and could contaminate other agents over the course of four days (for this value, see Salje et al. 2020a, p. 10). Infected agents move to R with probability following a normal distribution with average 0.993 (and possible range at the agent-level between 0.990 and 0.996) (for the average value, see again Salje et al. 2020b) provided they have spent a number of days in I at least equal to a given recovery time. The recovery time followed a Poisson distribution centered on 2 weeks (with possible range at the agent-level between 1 and 6 weeks) (for these values, see empirical estimates in World Health Organization 2020, p. 14).

3.15 We combined this basic compartment structure with our network topologies so that agents who an infectious agent could infect were determined by the network of close-range contacts (see Barrat et al. 2008, ch. 9). During each day, an infectious agent could only transmit the disease to its direct contacts. The dyadic, meaning agent-to-agent, transmission probability r was assumed to be normally distributed with mean equal to 0.03, 0.05 or 0.07, and standard-deviation equal to 0.02[3].

3.16 To the best of our knowledge, currently there is no data that allows us to estimate the transmission probability at the dyadic level. For this reason, we followed a common procedure that simulates the model under different values of the likelihood of infection in order to assess whether the intervention of interest is robust across epidemics of different sizes (see, for instance, Block et al. 2020). In our model, the values of 0.03, 0.05 and 0.07 were chosen because they were able to trigger, on the DC network, epidemics where approximately 20%, 60% and 80% of agents were ever infected, thus allowing us to assess the effect of hub-targeted versus random interventions under very different scenarios[4].

3.17 All simulations started with five (randomly chosen) initially exposed infected agents. This is the lowest number of seeds that prevents excessive variability across simulation trials in our model population[5].

Interventions

3.18 We followed prior studies by considering targeted interventions that offer a set of agent protection against the virus (Pastor-Satorras & Vespignani 2002; Herrmann & Schwartz 2020b): agents present in the Susceptible, Exposed, or Infectious compartment were moved to the Recovered compartment. The intervention thus

prevented future infection of the targeted individuals, if susceptible, or prevented further spread from these targeted agents to other agents, if already infected. The intervention could represent any combination of measures, such as a vaccine against Covid-19, medical testing and quarantining-if-positive, protections in high-risk professions and targeted informational campaigns (Banerjee et al. 2020). We assumed a government with a fixed daily (medical / technological / financial / ethical) capacity to intervene on b individuals. This was implemented as follows: On day 1 (iteration 1), b agents are selected from among all agents that are either in S, E, or I and moved to R, on day 2 (iteration 2) b additional agents are selected who are currently in S, E, or I, and moved to R, and so on. We studied four budgets for each intervention: $b = 1, 3, 5$, and 10.

3.19 We considered three methods for selecting agents for intervention. The first method, "NO-TARGET", simply randomly samples b agents for intervention each day and is intended as a benchmark against which to contrast the other two methods. This method corresponds to what Pastor-Satorras & Vespignani (2002) refer to as "uniform intervention."

3.20 The second method, "CONTACT-TARGET", follows the strategy described in Cohen & Havlin (2010), whereby each day b random agents are sampled who each select one random contact (without replacement) for intervention. Because of the friendship paradox (Feld 1991), these targets have above-average expected degree (Christakis & Fowler 2010). This is so because high-degree nodes are by definition overrepresented among other nodes' contacts (Feld 1991). The CONTACT-TARGET strategy is implementable in practice as a government could in fact randomly sample from the known population and have sampled individuals suggest their contacts. This was implemented as follows: On day 1, b random agents are sampled from among all agents. For each sampled node a random network neighbor is sampled. The intervention is targeted at these b random neighbors. On day 2, again b random nodes are sampled. For each sampled node a random network neighbor is sampled who had not previously been intervened on. The intervention is targeted at these b random neighbors, and so on[6].

3.21 The third method, "HUB-TARGET", assumes that agents' numbers of contacts were perfectly observed. During each iteration, nodes are targeted in strictly decreasing order of their network degree, starting with the b largest hubs. This was implemented as follows: On day 1, the b nodes with the b highest degrees are selected and immunized; on day 2, the b nodes with degree rank $b + 1$ through $2b$ are targeted, and so on.

Results

4.1 Figure 3 shows the number of concurrently infected individuals over time for the four simulated networks when no interventions are taken. Results for the ER network are represented with dashed black curves and the three DC networks with low, medium and high clustering are shown as solid curves in respectively black, blue and red. Shaded areas represent variability in the inner 90% of simulated runs, that is, between percentiles 5 and 95. Panels A, B, and C show results for various dyadic transmission probabilities r, respectively 0.03, 0.05, and 0.07. Peaks were naturally higher at higher transmission probabilities, with vertical axes rescaled to accommodate these base differences across panels. In panel A peaks are not easily identified due to minimal spread.

4.2 A comparison between the ER and the DC network with virtually no clustering ($C_c = 0.01$) showed that greater variability in degree generates higher peaks (for low dyadic transmission probabilities), and earlier and higher peaks (for middle and high dyadic transmission probabilities). This was consistent with theoretical results from formal models showing that in networks with high degree variance viral spread is faster, whatever the transmission probability is (Barthélemy et al. 2005). These results illustrate the impact of hubs: Highly connected individuals were more likely connected to the seeds and their neighbors. Once infected, they exposed others early on, thus catalyzing viral diffusion. In the ER network, by contrast, there were no hubs to accelerate spread.

4.3 The epidemic size measured as the total number of ever infected agents is not easily seen in Figure 3. Table 2 shows these estimates along with 95% intervals. The DC network without clustering produced a larger epidemic than the ER network at low and medium dyadic transmission probabilities. At a high transmission probability, the ER network rather produced a larger epidemic. This pattern can be understood as follows: At high dyadic probabilities anyone is nearly guaranteed to eventually become infected except those with few ties. In DC networks, there are many more agents with fewer ties than in an ER network. By contrast, at low transmission probabilities most agents are likely to escape the pandemic except hubs. In DC networks there are more hubs than in ER networks.

4.4 A comparison between the three DC networks in each panel of Figure 3 shows that epidemics are monotonically slower in networks with greater clustering. This is consistent with theoretical results from formal models showing that in networks with high degree variance increasing clustering attenuate hubs capacity to accelerate viral spread (Eguíluz & Klemm 2002; Serrano & Marián 2006). The mechanism is straight-forward: In networks

with high levels of clustering, sources of infection are more likely to expose the same target rather than distinct targets, reducing overall exposure (see Molina & Stone 2012, Figure 3).
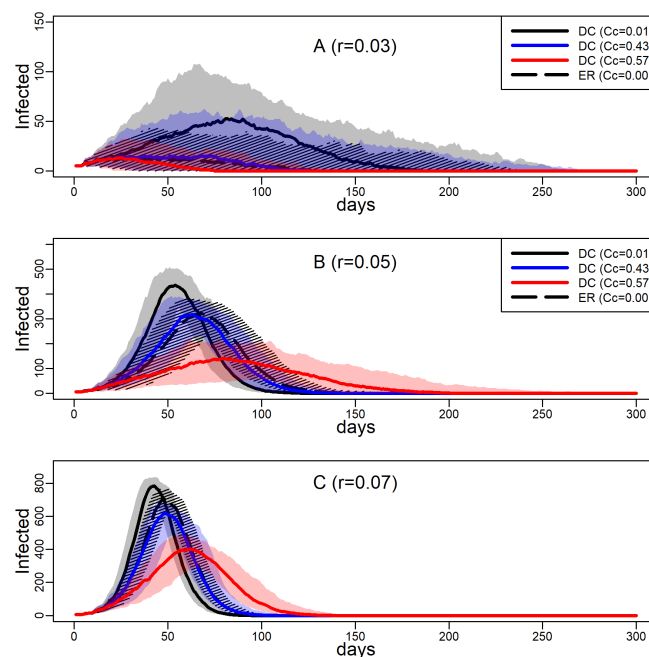

Figure 3: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) as a function of increasing values of the dyadic transmission probability r and clustering (see Legend). Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. $n = 2{,}029$ agents. Solid line: *Degree-calibrated* (DC) networks; dashed line: *Erdős–Rényi* (ER) network with the same average degree.

	$r = 0.03$			$r = 0.05$			$r = 0.07$		
	peak height	time	Epidemic size	peak height	time	Epidemic size	peak height	time	Epidemic size
DC ($C_c = 0.01$)	**52.5** [0; 96]	81	427.5 [13.9; 546.3]	**436.5** [307.75; 501.1]	54	1241.5 [1181; 1312.05]	**783.5** [687.7; 836.15]	43	1587 [1560; 1628.05]
DC ($C_c = 0.43$)	**15.5** [0; 60.10]	68	151 [7.95; 439.1]	**317** [218.1; 373.3]	63	1173.5 [1108.85; 1248]	**619.5** [410.55; 684.45]	48	1547.51 [509.95; 1586]
DC ($C_c = 0.57$)	**13** [2; 26.15]	19	46.5 [7.95; 158.65]	**140** [46.45; 210]	79	996.5 [822.6; 1117.55]	**401.5** [286; 474.15]	63	1510 [1414.75 1566]
ER	**13** [1.95; 30.05]	24	102 [8.95; 313.6]	**322** [208.65; 376.35]	68	1210 [1136.9; 1288.15]	**704.5** [584.05; 766.05]	48	1661.5 [1619.95 1713.05]

Table 2: Peak height (maximum # concurrently infected agents), time (in days), and epidemic size (# ever infected agents) on the *Degree-calibrated* (DC) networks with increasing clustering (C_c), and *Erdős–Rényi* (ER) network with the same average degree (rows) under low, middle and high dyadic transmission probability r (column). Shown are median, 5% and 95% percentiles across 100 iterations.

4.5 Figure 4 shows the impact of the three intervention methods on viral diffusion in the DC network with low clustering, assuming an intermediate dyadic transmission probability. Peak reductions and timing are reported in Table 3. Spread under intervention regimes is displayed as dashed curves in Figure 4. Solid curves represent the no-intervention scenario, in contrast. Panel A shows the results for the NO-TARGET procedure, whereby each day b randomly selected susceptible, exposed or infectious agents were intervened on. The NO-TARGET

procedure's maximally achievable impact, using the most generous budget considered, $b = 10$, corresponding to about 10% of the population being treated during the first 20 days, leaves the peak at 40% (Table 3, 174 / 436.5) of what it would have been without any intervention. Also, the peak occurs on about the same day.

4.6 Panels B show results for the CONTACT-TARGET procedure, which assumes that no global information on connectivity is available. Lacking this information, it attempts to find high-degree nodes by drawing a random sample of agents with unknown degree and selecting a random neighbor of each sampled agent for intervention. The figure shows this procedure is more effective than the NO-TARGET intervention. At $b = 5$ CONTACT-TARGET interventions produces an impact comparable to a NO-TARGET intervention regime with $b = 10$ daily interventions. At $b = 10$, the CONTACT-TARGET intervention achieves a reduction down to only 14% of the peak in the no-intervention scenario (Figure 4, 62 / 436.5). The peak is reached nine days earlier, after 45 days (CONTACT-TARGET) instead of 54 days (NO-TARGET).

4.7 The CONTACT-TARGET method would be more effective if randomly chosen agents would be able to select a random neighbor for intervention among relatively high degrees at a higher chance than network structure per se allows. Survey data suggest that targeting of certain professions may help effectively to identify high-degree agents (see Appendix A). To evaluate the maximally achievable impact of any degree-based intervention, panels C of Figure 4 show the impact of the HUB-TARGET policy, whereby each day the b previously untargeted agents with highest degree were targeted. A budget of 3 agents per day ($b = 3$) reduced the peak to 30% (Figure 2, 131.5 / 436.5). The peak occurred at the same time as without the intervention. This reduction in peak daily infections achieved with $b = 3$ exceeded what NO-TARGET intervention achieved with ten agents per day ($b = 10$). With ten agents the pandemic was effectively prevented.

4.8 In sum, these results suggest that insights from formal models on abstract networks concerning the effectiveness of degree-targeting extend to networks with degree distributions that concord with contact survey data. And, by recalculating the Figure 4 and Table 3 for the *Erdős–Rényi* (ER) with the same average degree as the empirical degree distribution but lower degree variance (see Table 1 above), it can be proven that it is precisely the skewness in the empirical distribution of close-range contact that makes hub targeting more effective than random targeting (see respectively Appendix B, Figure 10 and Figure 5).

Degree-calibrated (DC) network		no intervention: Peak height = 436.5 [307.75; 501.1]; Time = 54					
	b = 1		**b = 3**		**b = 5**		**b = 10**
	peak height	time	peak height	time	peak height	time	peak height	time
NO-TARGET	394 [306.5; 462.1]	55	341.5 [144.55; 407.35]	55	295 [196.45; 362]	52	174 [45.8; 226.3]	52-53
CONTACT-TARGET	382.5 [222.85; 447.65]	54	262.5 [131.95; 338.00]	54-57	184 [84.5; 255.15]	54	62 [11.85; 112.2]	45
HUB-TARGET	298.5 [152. 364.2; 182]	56	131.5 [15.6; 192.95]	53	52.5 [3.95; 112.45]	44	16 [3.95; 35.05]	23-32

Table 3: Peak height (maximum # concurrently infected agents) and time (in days) under three interventions (rows) and four budgets (column) on the *Degree-calibrated* (DC) network. Dyadic transmission probability $r = 0.05$ & Local clustering ($C_c = 0.01$). Shown are median, 5% and 95% percentiles across 100 iterations.

4.9 Figures 5 and 6 explore the robustness of this result under increasing levels of clustering in the DC networks. Panels A, B, and C again represent results for the three intervention methods separately, again assuming an intermediate dyadic transmission probability ($r = 0.05$). We consistently found a substantial improvement in virus-spread control of the contact- and hub-targeting methods over the random targeting method. In the Appendix B, we further explored the robustness of our results under different assumptions on transmission probabilities and clustering (see in particular, Figure 11-16). Results were qualitatively unchanged.
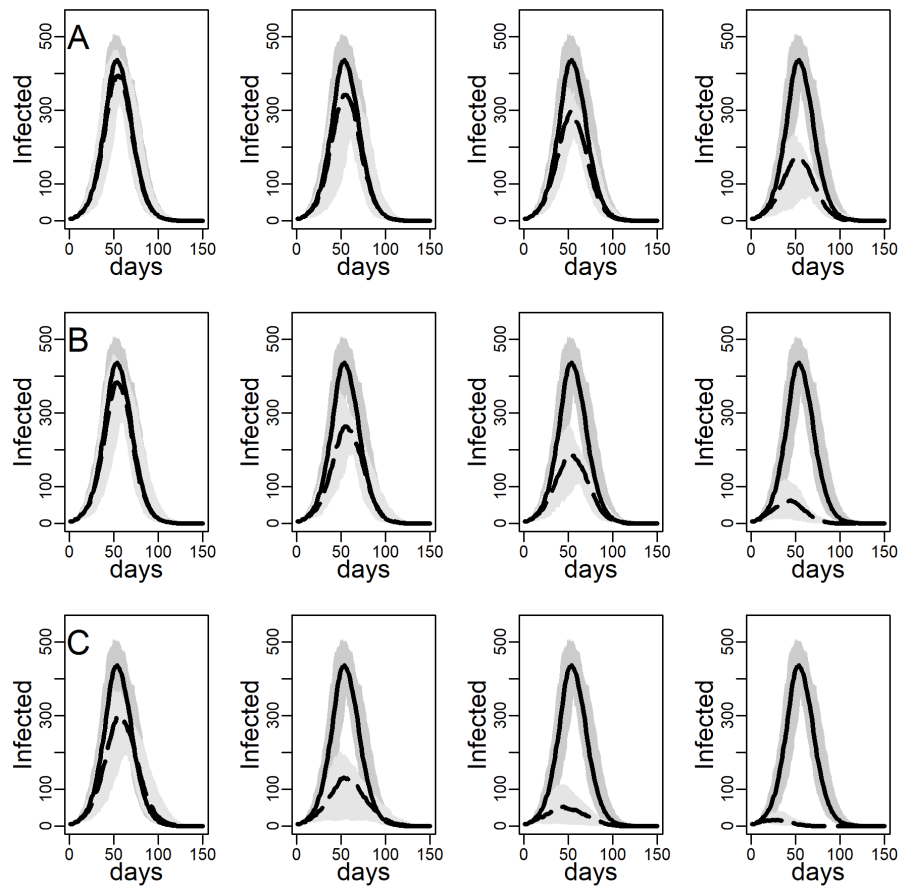
Figure 4: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Empirical-degree* (ED) network; dashed line: interventions. Dyadic transmission probability $r = 0.05$ & Local clustering ($C_c = 0.01$). $n = 2,029$ agents.

Discussion

5.1 In the absence of a vaccine, countries worldwide seek to contain viral spread through a combination of social distancing, protective measures, informational campaigns, testing, and contact tracing (Sustained Suppression, *Nature Biomedical Engineering* 2020). Yet there are clear limits on medical, technological, and financial resources and on the ability to durably restrict individual mobility, raising the question of how to prioritize. Our results suggest that all these interventions will generally be more efficient when targeted at individuals suspected or known to have close-range contact with many others. Once an effective vaccine has been developed, it may remain available in small quantities only for some time and/or face skepticism by large fractions of the population (Peretti-Watel & the COCONEL Group 2002). Based on our simulation results, we can expect vaccination to reduce the spread to a greater degree when high-contact individuals are given the first vaccines.

5.2 How could public policy effectively try and identify high-contact individuals? We propose two possible methods. First, the approach we systematically studied in our simulations is agnostic of who the high-degree individuals are and targets random acquaintances of random individuals, who statistically have high expected degree. This method was found to be effective in detecting past flu outbreaks (Christakis & Fowler 2010), and robust against missing network data (Rosenblatt et al. 2020). In our simulations, this method was reasonably effective yet at the same time was conservative in assuming no knowledge of degree or use thereof. If individuals did not nominate random contacts but rather those they knew to have had many other contacts, the difference that targeted intervention could make would be greater.

5.3 The second method we suggest exploits the covariation that seems to exist between individuals' occupation and the volume of their daily close-range contacts. Previous large-scale surveys on face-to-face encounters
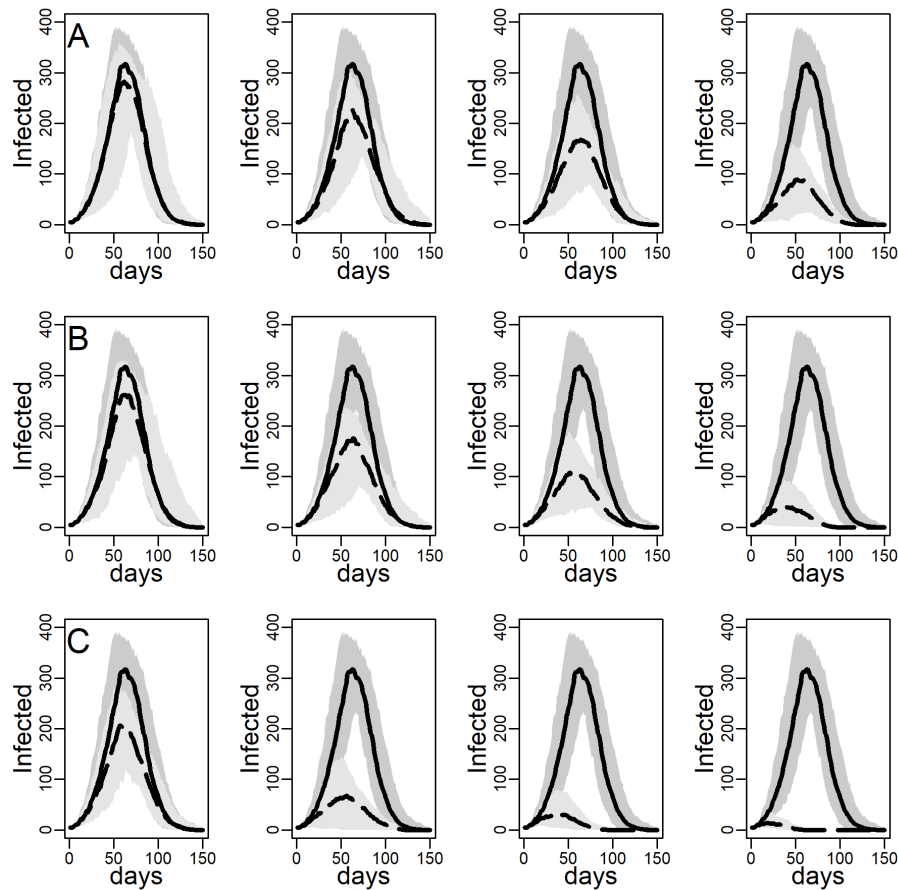
Figure 5: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability $r = 0.05$ & Local clustering ($C_c = 0.43$). $n = 2,029$ agents.

have documented that some professions (like teachers, service workers or health care workers) are especially exposed to close-range contacts (see Danon et al. 2013, Figure 4). We replicate this result in a survey on six different countries: Some professions involve ten times as many close-range contacts than others, with elementary school teachers, cashiers, order clerks, retail salespersons, and administrators topping the list (see Appendix A, Table 4).

5.4 The correlation between profession and contact frequency could be exploited in two different ways. One may want to directly target workers in professions characterized by high frequencies of contact. This approach would have the advantage of making it possible to set interventions on the basis of category-based institutional lines, thus avoiding potential privacy and discrimination issues associated with targeting individuals. For instance, the occupational categories used in the international comparative survey we exploited follow a common international standard used by the US Bureau of Labor Statistics. Preferential protective legislation could be set on its basis. On the other hand, however, this approach ignores the variability in workers' exposure to social contacts that also occurs within a given occupational category (see Appendix A, Figure 8). Undifferentiated category-based interventions, by protecting individuals within the category who are below the average exposure, waste resources. Another way to exploit the covariation between individuals' occupation and the volume of their daily close-range contacts would thus be to inject this information within the method of targeting random acquaintances of random individuals. According to this hybrid approach, randomly sampled individuals may be asked to preferentially report random social contacts within a given list of highly socially exposed professions. These contacts would then have a higher expected degree, rendering the method more effective.

5.5 One may expect this approach to especially benefit low-income workers. Quasi-experimental evidence suggests that the substantial income gradient in the impact of the pandemic on mortality is strongly mediated by
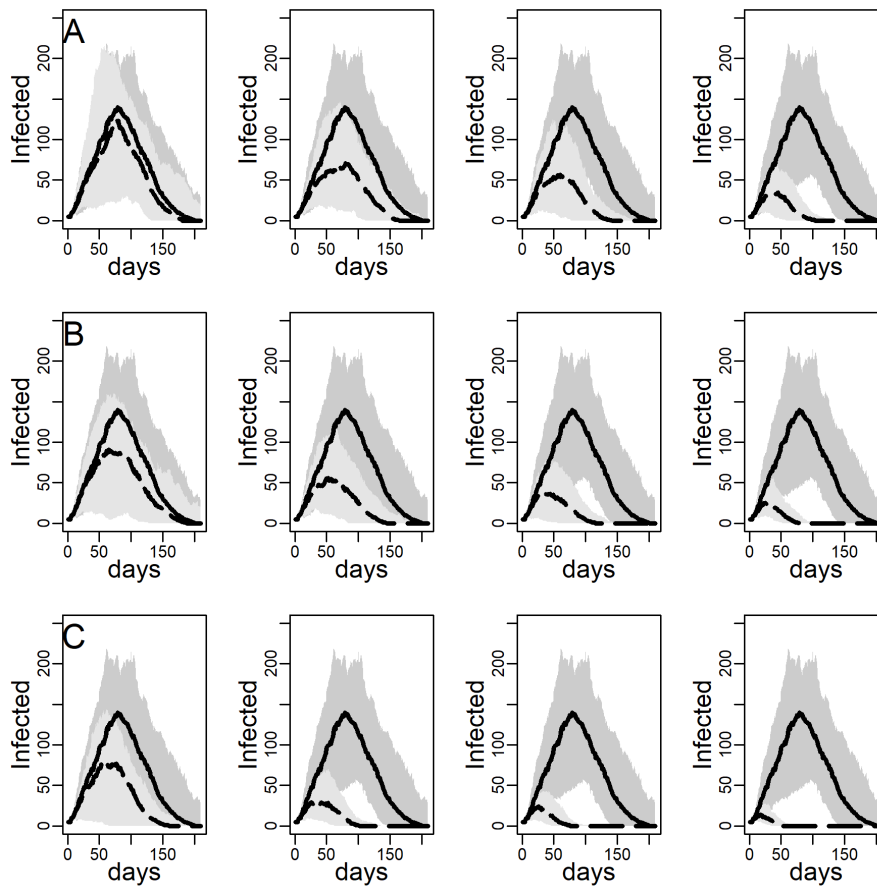
Figure 6: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability $r = 0.05$ & Local clustering ($C_c = 0.57$). $n = 2,029$ agents.

low-income workers being trapped in low-paid jobs with high exposure to social contacts (Brandily et al. 2020). A hybrid strategy that combines occupational exposure with random acquaintances of random individuals to identify high-contact individuals to be protected/tested preferentially may thus be highly effective in reducing the overall death toll associated with SARS-CoV-2.

5.6 Finally, let us emphasize that targeting hubs could also help other mitigation strategies at a lower cost. It has been proposed that the limitation of mobility networks can diminish the reachability of contact networks, through reducing the contact network's diameter, which in turn could slow down virus propagation (see, for instance, Brethouwer et al. 2020; Chinazzi et al. 2020). Our study suggests that this same result can be achieved by targeting hubs rather than halting movements of large fractions of the population, if hub connections make up a large share of bridge ties (Barabási 2014, p. 64). This phenomenon can be found back in the degree-calibrated networks that we generated. Increasing the standard deviation of the degree distribution (from approximately 7 to approximately 19) through the inclusion of high-degree individuals leads to a reduction of the network average path length from 3.47, 4.38 or 5.52 to 2.83, 3.29 or 3.60, depending on the level of built-in clustering (compare Tables 1 and 6). And our results show that, with all else equal, networks with more and larger hubs, thus lower average path length, systematically lead to faster and larger epidemics (compare for instance, curves in Figure 3 to curves in Figure 17). Thus, given the deep link between the notion of hubs and the small world nature of networks (Albert et al. 2000), our proposed viral mitigation strategy effectively operates through a world-enlargement mechanism.

Conclusion

6.1 In this paper our goal was to assess the effectiveness of preferentially targeting hubs versus undifferentiated interventions for controlling SARS-CoV-2 spread. With this aim in mind, we moved away from the standard compartmental models that rely on random mixing assumptions toward a network-based modeling framework that can accommodate person-to-person differences in both infection risk and ability to infect others stemming from differential connectedness. Unlike past studies, we simulated virtual epidemics on networks with empirically calibrated frequencies of close-range contact. This framework allowed us to model rather than average out the high variability of close-contact frequencies across individuals observed in contact survey data. Results of simulations calibrated with empirical close-range contact distributions exhibiting right skew show large improvements in epidemic containment when shifting from general to targeted interventions. The relative effectiveness of preferentially targeting hubs proved highly robust across changes in degree skewness, clustering, and infection probability, as well as across epidemics of various sizes.

6.2 Our study has several limitations. First, the recommendation of prioritization of hubs in interventions is based on an assessment of effects on overall containment. There may be reasons to prioritize differently, e.g. protecting those in the medical profession dealing with SARS-CoV-2 as to maintain maximum capacity to treat. Alternatively, the protection of highly vulnerable subpopulations may reduce the overall death toll. The present paper does not speak to these alternative considerations as medical capacity and death rates are not modelled.

6.3 Second, our model is silent on the specific content of the actions to be performed on each finally selected individual. We only provide a method to maximize the efficiency of that selection. In the absence of a vaccine, moving from the model to the real-world world, an intervention could involve a combination of: (a) testing and quarantining-if-positive, (b) additional provision and mandatory protective equipment such as face masks and transparent physical barriers, (c) closer monitoring and tracking with mobile devices, and (d) targeted and contextualized informational messages stressing the importance of certain acts of social distancing and use of protective measures. Targeted messages can be relatively inexpensive as they are performed at a distance (Marcus 2020) and evidence suggests they have strong health-behavioral effects (Noar et al. 2007), including recent field-experimental evidence to this effect for SARS-CoV-2 (Banerjee et al. 2020).

6.4 Third, our simulation models varied only heterogeneity in the number of close-range contacts across individuals and levels of clustering. By doing so, we may have overlooked more global features of close-range contact networks that drive viral propagation, e.g., fragmentation and community structure. Such topological factors could moderate the impact of hubs. For example, when hubs span different communities they may accelerate disease spread more than when contributing many ties to an already tightly knit community. The ego-centric survey data we drew upon, while providing representative measures of degree heterogeneity, could not provide such global measures, which require population level data. A promising source of data is smartphones, as used in previous small-scale studies (Mones et al. 2018; Sapiezynski et al. 2019; Cencetti et al. 2020). Future studies may be able to draw on large-scale data from smartphone apps recently rolled out to monitor and prevent the spread of SARS-CoV-2.

6.5 Finally, a factor that could limit the super'spreader status of hubs and the effectiveness of hub targeting is contact time, namely if contact time were inversely proportional to the number of contacts. In this case hubs' shorter average per-tie duration of contact may be associated with lower risks of contracting and spreading the coronavirus (provided the probability of transmission is negatively correlated with the contact duration). Our contact diary data revealed that individuals with many close-range contacts on average spend a similar amount of time per contact as those with few close-range contacts. The evidence suggests that the augmented risk associated with greater contact numbers are not offset by shorter durations. While these findings reinforce the critical role hubs may play in disease propagation, lack of data on contact duration for individuals with (much) higher contact volume than those we could observe prevent us from identifying the point above which a negative correlation between contact volume and average contact duration appears. While theoretical studies that make the probability of transmission inversely contingent on a node's degree do not univocally find that this negative correlation attenuates the importance of hubs (Olinky & Stone 2004), in order to settle this question, we need better data on the relationship between transmission probability, contact duration and contact frequency.

Model Documentation

All statistical analyses of individual-level data were performed with R language (release 3.6.3) or StataSE 14. Network statistics are partly computed with the R's igraph packages (version 1.2.5) and partly with NetLogo's "stat"

and "nw" extensions. The agent-based model is written and simulated in NetLogo (release 6.0.3). COMES-F survey can be downloaded at: `https://figshare.com/articles/Data_file_for_Comes_F/1466917`; detailed documentation on the dataset, the full text of the questionnaire and the paper contact diary can be found in Bèraud et al. (2015) supplementary information. Belot et al. (2020)'s survey that we analyze in the Appendix can be downloaded at `https://osf.io/gku48/`. All .nlogo and .nls files containing the simulation model's code are accessible at `https://www.comses.net/codebases/25d1ac60-7a5b-4331-a499-6163607241d2/releases/1.1.0/`. The model folder also contains the specific portion of French survey data needed to calibrate the network as a separate .txt file. All simulations can exactly be replicated throughout the list of random-generator-seeds displayed in the code. Please read the "Read.me" file for more details.

Notes

[1]The power law is fitted using the R implementation (in package igraph) of the maximum likelihood method developed by Clauset et al. (2009).

[2]All networks we study are fully connected, meaning that there is at least one path connecting each pair of agents. In less than 1% of the network realizations of our algorithm are any nodes disconnected, often only 1 or 2. When this is the case, we exclude these nodes, following common practice in the study of viral diffusion on networks (e.g., Montes et al. 2020). Note that disconnected regions are empirically implausible in face-to-face encounter networks of open, unquarantined societies.

[3]Rather than limiting transmission over the short period of time where infectiousness seems to be highest, an alternative specification would consist in using a reversed-U shape (discrete or continuous) probability function over a larger infectious period. Herrmann & Schwartz (2020a, Appendix), for instance, on the basis of epidemiological data concerning 94 Chinese patients, span dyadic transmission probabilities over a period of 14 days on a probability interval starting at 0.01, peaking at 0.3 (for days 5, 6, and 7), and progressively going back to 0.01. On a purely simulated network, and assuming perfect knowledge of the degree distribution, the authors target nodes with the highest degree, and show the effectiveness of this strategy to mitigate the virus spread. Thus, a modeling choice that requires making assumptions on a much larger number of values than ours leads to results that, as to the role of hubs, are in line with our own results.

[4]In terms of basic reproductive R_0, if one computes this quantity for a network with heterogeneous degree (see, in particular, Olinky & Stone 2004, Equation 1), the chosen values of the dyadic transmission probability correspond, for the DC network (with no clustering), to virus spread characterized by R_0 respectively equal to approximately 1.52, 2.53, and 3.55, which allow to cover a wide range of R_0 (or R_t) values observed in different contexts, and/or at different time, in the current COVID-19 crisis (for France, see for instance, Salje et al. 2020b; Roux et al. 2020; in a comparative perspective, see Flaxman et al. 2020.

[5]In simulations on larger networks, which take much longer to run, in which we implemented the same degree distribution and used the same number of seeds (smaller fraction), we find that peaks naturally occur later, while the interventions we present next show qualitatively the same relative effects (results are available upon request).

[6]The procedure is implemented in such a way that: (a) if a randomly selected agent has no neighbor who had not been intervened on before, a new randomly selected agent is sampled as long as the condition is met; (b) the required number of agents to be intervened on is constantly adjusted as a function of the available population.

Appendix A: Variations of close-range contact heterogeneity by gender, age, and profession

We documented high variability in the number of close-range contacts across COMES-F's respondents (see Figure 1). Here we showed that this variability persisted within major demographic categories. Figure 7's upper panel shows the distribution of per day self-reported close-range contacts by respondent's gender. Past analyses of COMES-F data found that women (mainly adult women) tend to have a higher average number of contacts than men (see Bèraud et al. 2015, 6 and Table 1). Figure 7 shows that behind this average difference, there exists a large degree of variation within genders. Women and men have a nearly identical distribution of contact frequencies.

Figure 7's bottom panel shows the distribution of (diary-based) close-range contacts per day by respondent's age. Age is the most recurrent variable used in epidemiological models for representing socially structured social interactions. Age assortativity (and dissortativity at home) is found to be one of the most robust empirical regularities in epidemiological social contact surveys, as also found in multivariate analyses of the COMES-F data (see Bèraud et al. 2015, pp. 7–8). This motivates the use of average contacts per (more or less disaggregate) age-groups in age-structured compartment models (see, for some recent examples, (Di Domenico et al., Roux et al. 2020, Salje et al. 2020a, pp. 3–4; in a comparative perspective, see Walker et al. 2020). Net of the main effects of age on the likelihood of having more social contacts (which is indeed found in these data, see Bèraud et al. 2015, Table 1), Figure 7's bottom panel again shows high variability within age-groups. Median numbers of contacts vary little by age and the interquartile ranges of the various age groups mostly overlap.

Focusing analysis on adult respondents in employment, we found a similar pattern for broad occupational groups (see Figure 8). Among both diary-based contacts (top panel) and extra job-related contacts declared by (bottom panel), occupational groups vary little in the median number of close-range contacts and again the interquartile ranges overlap. To the extent that there is any systematic variability, farmers seem to have somewhat smaller contact networks, while high-contact individuals seem especially concentrated among high (e.g., elementary school teachers, teaching assistants) and low routine non-manual workers (e.g., bank tellers, tellers in public administration) and service class (e.g., university professors, politicians, journalists or doctors). Most variation across individuals in the number of self-reported close-range contacts clearly occurs within rather than between occupational groups.
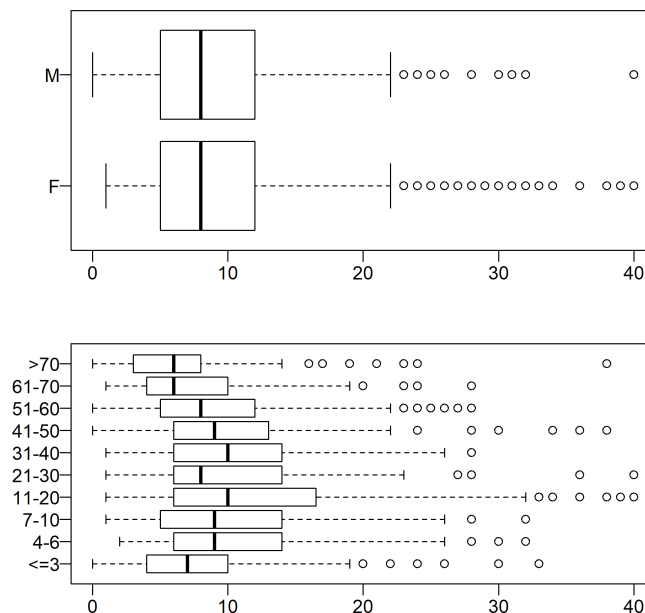

Figure 7: Distribution of self-reported per day close-range contacts (x-axis) by gender (top panel; F [1136], M [897]) and age groups (bottom panel; ≤ 3 [240], 4-6 [169], 7-10 [196], 11-20 [276], 21-30 [155], 31-40 [109], 41-50 [135], 51-60 [195], 61-70 [357], > 70 [201]) (y-axis).

While these data suggest that social contacts within occupations are much more dispersed than one could expect under a distribution symmetrically centered around the mean, COMES-F does not provide a detailed list of jobs. The professional categories in the COMES-F data were too coarse to evaluate the effectiveness of a method for targeting hubs on the basis of employment status. We therefore exploited a recent survey that has a somewhat less thorough measurement of close-range contact (Belot et al. 2020) but fine-grained professional categories.

Belot et al. (2020)'s survey was conducted in the third week of April, 2020 in the midst of the Covid-19 epidemic in six countries: China, South Korea, Japan, Italy, the UK and four states in the US: California, Florida, New York, and Texas. The sample consisted of roughly 1,000 individuals from each country for a total of 6,082 respondents. Data was collected using market research companies Lucid and dataSpring, using gender and income quota. With regard to close-range contact, instead of being asked to keep a two-day-long diary, respondents were asked: "On a typical working day (before the outbreak of Covid-19), with how many people would you have
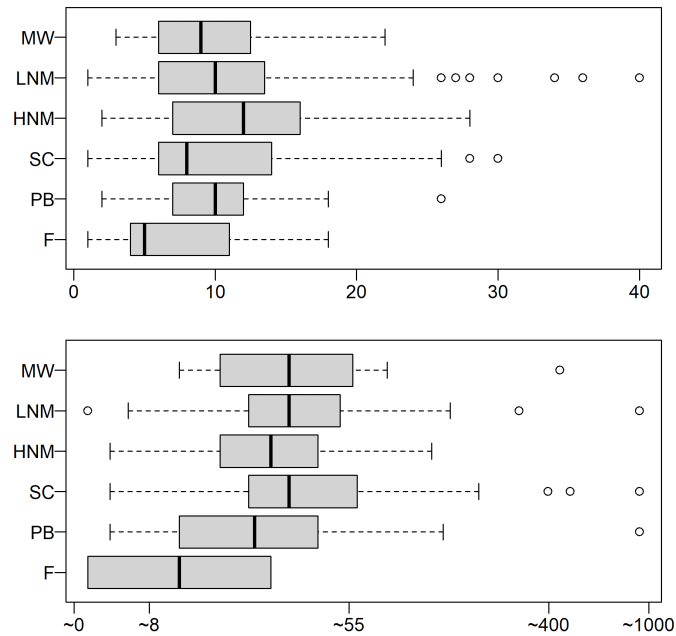
Figure 8: Top: Distributions of daily close-range contacts among employed respondents ($n = 436$) by occupational category (y-axis); *Bottom* (logarithmic scale with labels corresponding to values on the original linear scale): Distribution of daily job-related contacts among respondents regarding their occupation as especially exposed to close-range contact ($n = 257$) by occupation category (y-axis). Occupational categories (n) in parenthesis, top panel first): F = farmers ($n_{TOP} = 14$), ($n_{BOTTOM} = 2$), PB = petty bourgeoisie (craftsmen and shopkeepers) and entrepreneurs ($n_{TOP} = 26$), ($n_{BOTTOM} = 18$), SC= Service class (managers, high-skilled administrators, intellectual, scientific and liberal professions) ($n_{TOP} = 123$, $n_{BOTTOM} = 91$), HNM=High routine non-manual worker ($n_{TOP} = 47$), ($n_{BOTTOM} = 41$), LNM=Low routine non-manual workers ($n_{TOP} = 183$), ($n_{BOTTOM} = 93$), MW=manual workers ($n_{TOP} = 43$), ($n_{BOTTOM} = 12$).

close social contact (at less than one meter distance) and how long would you interact with them? (indicate approximate numbers - leave blank if the answer is zero)". Respondents' professions were classified in terms of the O-Net classification used by the US Bureau of Labor Statistics.

The distribution of close contact frequency in the Belot et al. (2020) data is displayed in Figure 9. The distribution is severely right-skewed, as we also observed for the French survey data. This provides confidence that the existence of hubs is not a measurement artifact but a robust feature of contact networks: Using different methods for measuring contact the same distributional characteristic is obtained.
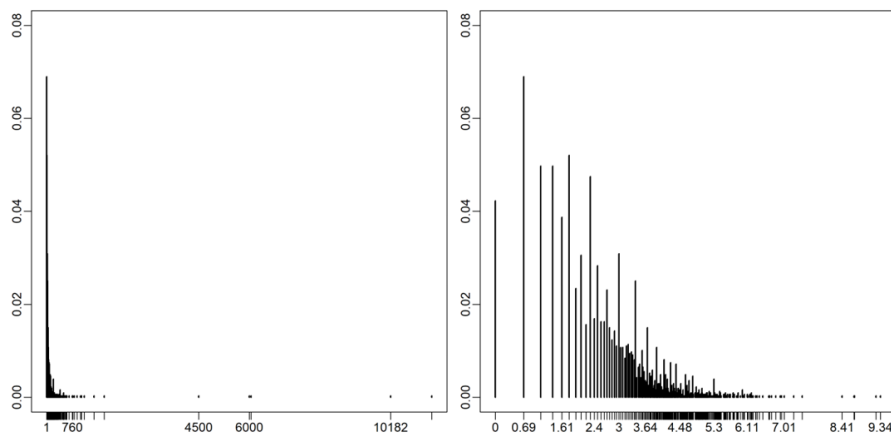

Figure 9: Fraction of cases (y-axis) reporting a given number of close-range contacts (x-axis) in the Belot et al. (2020) data ($n = 4,103$). Left: Linear scale. Right: Logarithmic scale.

Table 4 shows the mean and median number of close-range contacts by profession, in descending order of mean contact frequency, combining short- and long-duration contacts, excluding zero answers and professions with 15 or fewer cases. Table 4 has face validity, topped by professions that clearly involve close contact with many individuals — elementary school teachers, cashiers — and at the very bottom individuals who mostly work from home — computer programmers. Some professions have an order of magnitude greater mean close-range contact than others. The spread is substantial, especially when considering the ambiguity in the possible interpretation of the phrase "social contact" used in the questionnaire, translated into different languages, and the difficult task of estimating such numbers without use of a contact diary, which may produce noise that suppresses measured occupational differences.

Thus, given the systematic covariation that seems to exist between the fraction of high-contact individuals and specific occupation, when searching for hubs, targeting selected professions may be a reasonably effective strategy for finding hubs in contact networks.

Profession	Mean # contacts	Median # contacts	N
Elementary School Teacher	120	50	17
Cashier	76	40	20
Order Clerks	70	34	34
Teacher Assistants	67	40	18
Retail Salespersons	62	17	29
Administrative Services Managers	59	16	47
Childcare Workers	49	30	19
Bill and Account Collectors	45	19	21
Sales Managers	45	17	20
Computer and Information Systems Managers	34	14	33
Financial Analysts	34	19	25
Customer Service Reps	31	18	49
Audio and Video Equipment Technicians	31	19	27
Construction Managers	30	10	33
Construction Laborers	28	12	24
Architectural Drafters	25	11	16
File Clerks	24	12	38
Civil Engineers	23	16	53
Data Entry Keyers	22	16	17
Credit Checkers	22	14	16
Ophthalmic Laboratory Technicians	21	14	18
Computer Network Support Specialists	19	12	28
Financial Managers, Branch or Department	17	7	25
Financial Examiners	17	6	20
Computer Programmers	15	9	16

Table 4: Close contacts by profession in Belot et al. (2020) data.

Appendix B: Results for the *Erdős–Rényi* (ER) network and for the *Degree-calibrated* (DC) networks under various transmission probabilities

Figure 4 and Table 3 show large differences in the effectiveness of interventions that do and do not target high-contact individuals for intervention. Here, we explored how instrumental the skewness in the empirical distribution of close-range contact was for the effectiveness of hub targeting. We did so by recalculating the Figure 4 and Table 3 results for the *Erdős–Rényi* (ER) network with the same average degree as the empirical degree distribution (see Table 1), respectively Figure 10 and Table 5.

A comparison of panels A between Figures 10 and 4 shows that NO-TARGET interventions were less effective in DC networks with high degree skew than in ER networks with low degree variance. This suggests that models that did not account for the empirically observed inhomogeneity of network degree may overestimate the expected impact of interventions. Comparing panels B and C across figures we found that HUB-TARGET and
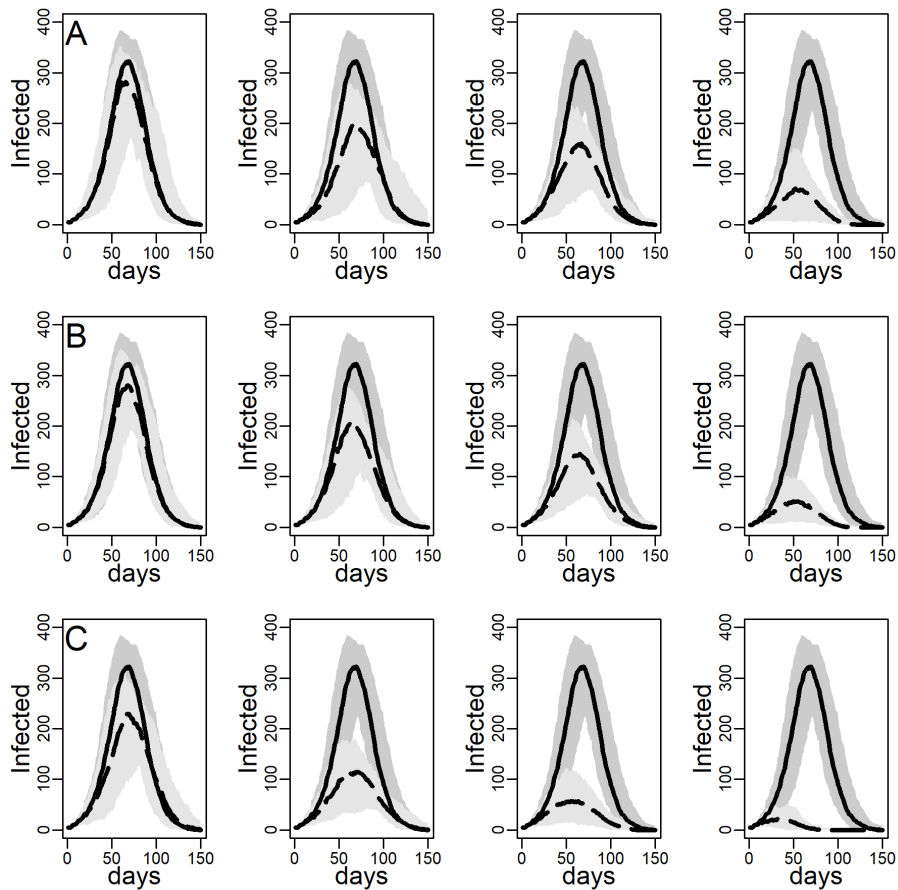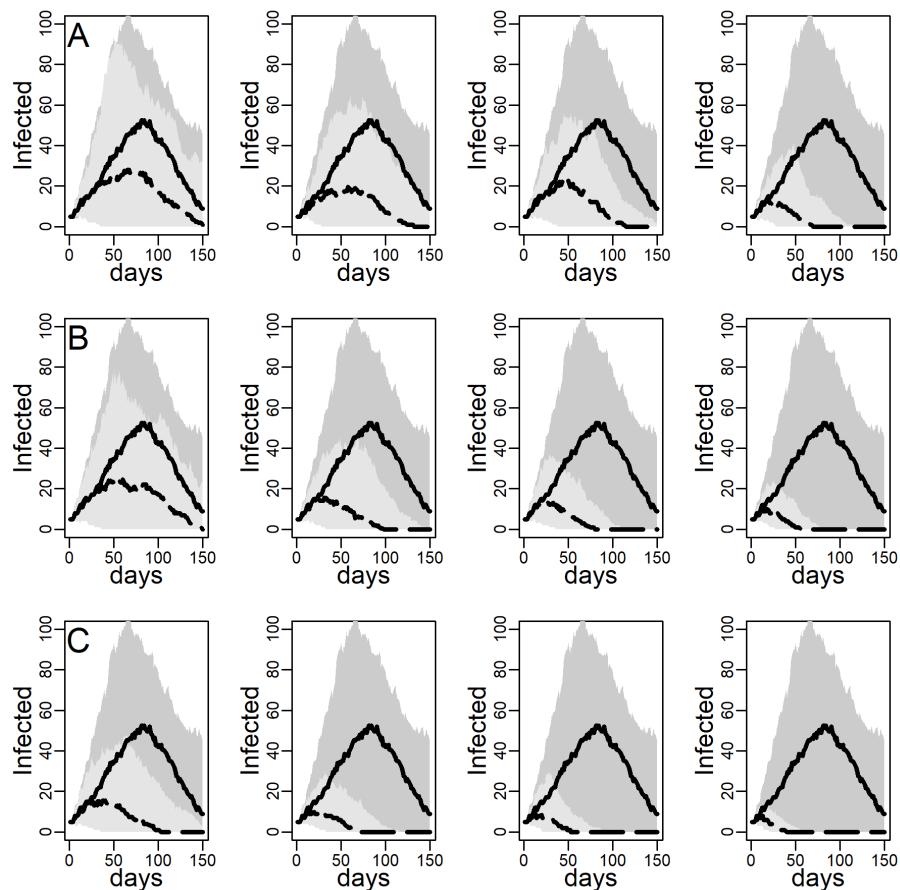
Figure 10: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Erdős–Rényi* (ER) network; dashed line: interventions. Dyadic transmission probability $r = 0.05$ & Local clustering ($C_c = 0.01$). $n = 2,029$ agents.

Erdős–Rényi (ER) network					no intervention: Peak height = 322 [208.65; 376.35] Time = 68			
	b = 1		**b = 3**		**b = 5**		**b = 10**	
	peak height	time	peak height	time	peak height	time	peak height	time
NO-TARGET	281.5 [141.6; 340.25]	66	200 [58.75; 264.40]	69	160.5 [60; 214.05]	67	71.5 [3.95; 150.15]	52
CONTACT-TARGET	281 [178.55; 338.15]	68	208 [73.9; 271.05]	65	145 [51.95; 204]	65	51.5 [10.85; 98.25]	50
HUB-TARGET	229 [95.85 ; 302.20]	66	114.5 [33.9; 175.35]	70	56.5 [14.95; 118.009]	56	22 [3; 43.35]	38

Table 5: Peak height (maximum # concurrently infected agents) and time (in days) under three interventions (rows) and four budgets (column) on the *Erdős–Rényi* (ER) network. Dyadic transmission probability $r = 0.05$ & Local clustering ($C_c = 0.01$). Shown are median, 5% and 95% percentiles across 100 iterations.

CONTACT-TARGET interventions were much more effective in the DC network than in ER networks, where the to-be-immunized agents had lower network degree.

We next checked the stability of the results shown in Figure 4 and Table 3 about the effectiveness of interventions that did and did not target high-contact individuals for intervention on the DC network when we decreased/increased the infectiousness of the disease. To this aim, below, we recalculated Figures 4-6 under low and high dyadic transmission probabilities r, in both cases over the entire range of clustering levels we built in our DC contact networks. Results were qualitatively unchanged. The relative differences across targeting methods were attenuated for small-size epidemics triggered by low dyadic transmission probabilities (see Figures 11-113) whereas they are enhanced for larger epidemics associated with high dyadic transmission probabilities (Figures 14-16). However, the effectiveness of the contact- and hub-targeting methods in mitigating the epidemic relatively to the random targeting method was still observed across all combinations of transmission probability and clustering levels.

Figure 11: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated (DC)* network; dashed line: interventions. Dyadic transmission probability $r = 0.03$ & Local clustering ($C_c = 0.01$). $n = 2,029$ agents.
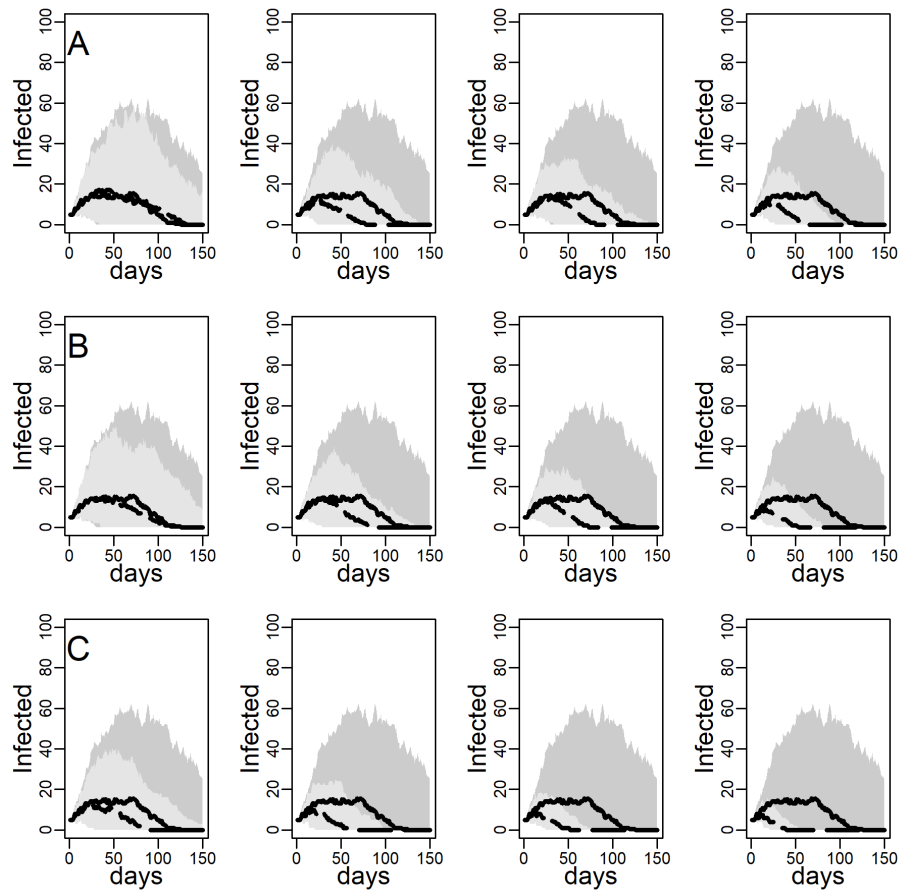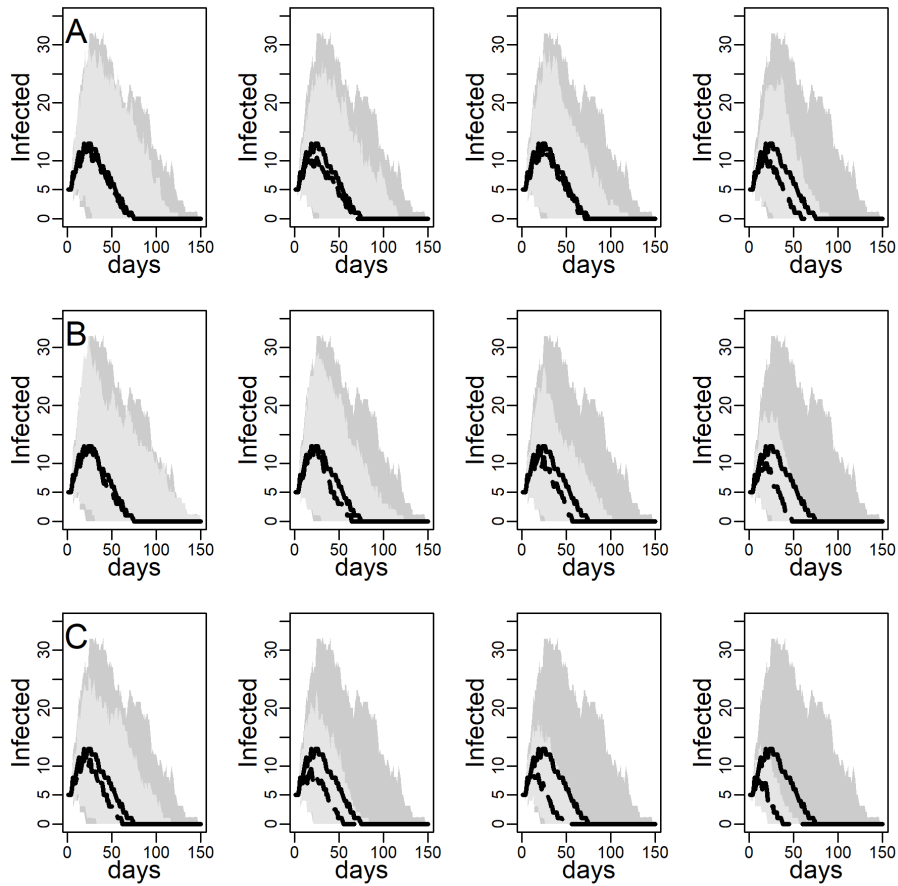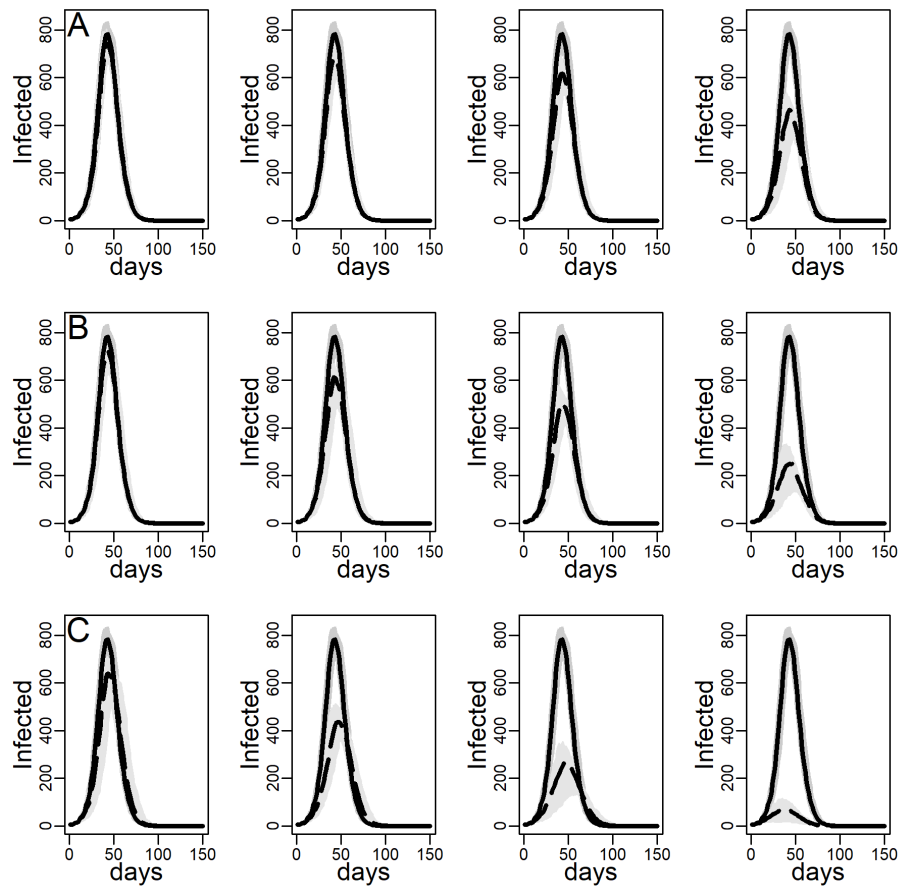
Figure 12: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability $r = 0.03$ & Local clustering ($C_c = 0.43$). $n = 2,029$ agents.
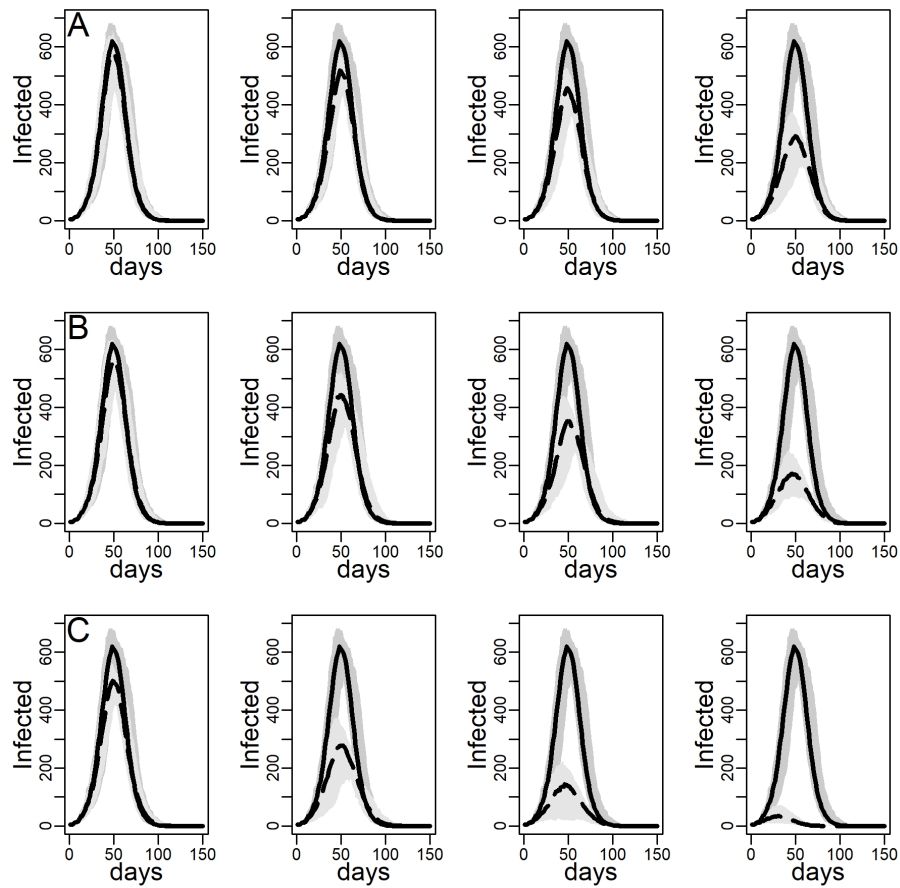
Figure 13: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability $r = 0.03$ & Local clustering ($C_c = 0.57$). $n = 2,029$ agents.

Figure 14: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability $r = 0.07$ & Local clustering ($C_c = 0.01$). $n = 2,029$ agents.
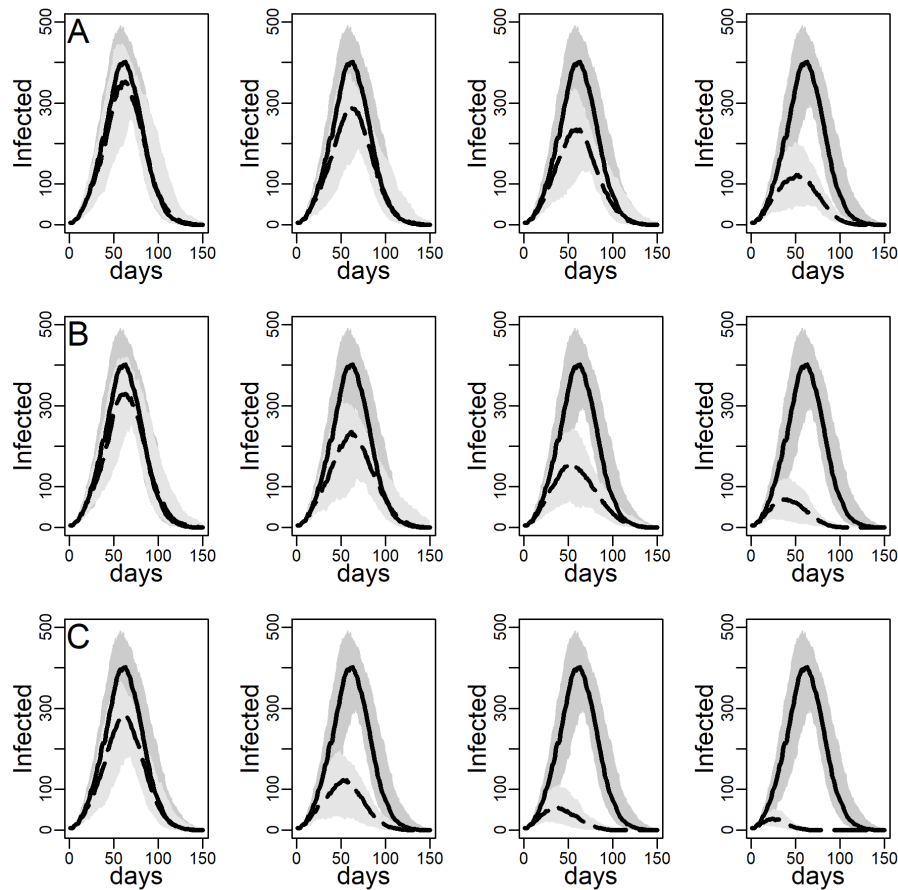
Figure 15: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability $r = 0.07$ & Local clustering ($C_c = 0.43$). $n = 2,029$ agents.

Figure 16: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability $r = 0.07$ & Local clustering ($C_c = 0.57$). $n = 2,029$ agents.

Appendix C: Results for the *Degree-Calibrated* (DC) network using additional professional contact data

Here, we reported on the results obtained by calibrating the synthetic network through a measure of daily close contact for each respondent that combines diary-based contacts (Figure 1's left plot) and job-related extra contacts (Figure 1's right plot). In particular, for respondents in employment who self-reported job-related extra contacts ($n = 259$), these contacts were summed up to the number of contacts recorded through the diary. However, we limited the portion of job-related extra contacts to be added in such a way that the total number of contacts is never higher than 134. Existing contact survey data suggests that the monotonic increase in the relationship between the number of close-range contacts and the total amount of time of these contacts that we described in Figure 2 (left plot) started to decrease above 100 contacts (see Danon et al. 2012, Figure S3c; Danon et al. 2013, Figure 2. This suggests that the relationship between the total number of contacts and the average contact length may start to become negative above this threshold. We choose the specific value of 134 to be consistent with previous studies of COMES-F data, where supplementary professional contacts were censored at 134 (see Bèraud et al. 2015, pp. 5, 9). As a by-product, this choice prevented a few nodes (like the two with 500 or the three with 999 job-related contacts) from having contacts with a substantial fraction of our simulated population of $\sim$2k agents, thus reducing the risk of artificially overestimating the impact of hubs.

Table 6 shows network statistics computed over 100 realizations of the DC and ER networks. Compared to the networks including only diary-based contacts (see Table 1), apart from a higher average degree, it is noteworthy the larger standard-deviation (approximately 19 versus approximately 7) of the DC network, which reflected the

larger portion of high-contact respondents that are now in the network. If one considers the right-tail of the degree distribution of this network (which again is well approximated by a power law with a scale parameter 2.54 for respondents with close-range contact above 17), this tail now contained 445 nodes (compared to 175 nodes for the power-law-like right-tail of the network including only diary-based contacts). Compared to the diary-based contact network, such a ticker right-tail translated into a lower average path length (even when built-in clustering increases), which clearly shows that the larger, and more numerous, the hubs the stronger their capacity to create bridges across otherwise distant parts of the network (on the connection between hubs and small-world behavior, see Albert et al. 2000).

DEL	Average degree	Median degree	Stdev degree	Clustering coef	Deg-clust corr	Av path length	Diameter
Empirical-degree (ED) networks							
$p = 0$	14.87 (0.00)	9.00 (0.00)	19.58 (0.00)	0.04 (0.00)	-0.15 (0.01)	2.83 (0.00)	4.8 (0.40)
$p = 0.5$	14.72 (0.04)	9.00 (0.00)	19.19 (0.04)	0.42 (0.01)	-0.50 (0.00)	3.29 (0.02)	5.01 (0.01)
$p = 1$	14.77 (0.04)	9.00 (0.00)	19.38 (0.12)	0.50 (0.01)	-0.44 (0.01)	3.60 (0.08)	6 (0.62)
Erdős-Rényi (ER) network							
ER	14.86 (0.14)	14.89 (0.31)	3.86 (0.14)	0.01 (0.01)	-0.00 (0.03)	3.09 (0.01)	5 (0.00)

Table 6: Topological features of the simulated contact networks (as a function of the local tie probability p, for the *Degree-Calibrated* (DC) network). Mean values across 100 network realizations (standard deviation in parentheses). Clustering coef = clustering coefficient; Deg-clust corr = Pearson correlation coefficient between nodes' degree and their clustering coefficient; Av path length = Average of the shortest path lengths; Diameter = Maximum of the shortest path lengths.

We re-ran all the analyses on the extended DC networks and the ER network with the same average degree. Results are reported below. Figure 17 reproduces Figure 3 in the main text; Figures 18-20 (dyadic transmission probability $r = 0.05$), 21-23 (dyadic transmission probability $r = 0.03$), and 24-26 (dyadic transmission probability $r = 0.07$) respectively reproduces Figures 4-6 in the main text and Figures 11-13 and 14-16 Appendix B. We do not comment in detail on these figures because results are in line with our main analysis. As to the effects of degree skewness and clustering on epidemic's size and paste where no intervention is in place, (Figure 17), we found the same patterns as in the network including only diary-based contact. Furthermore, as to the effectiveness of the contact- and hub-targeting methods in mitigating the epidemic relatively to the random targeting method, the relative gradient between these strategies is still observed across all combinations of transmission probability (r) and clustering levels. The superiority of targeting hubs only appeared more clearly because of the larger size and fraction, of high-contact nodes.
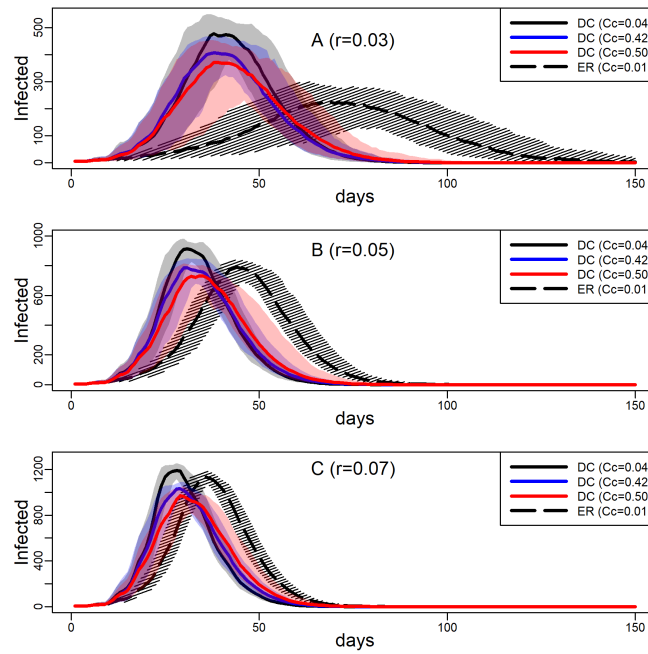
Figure 17: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) as a function of increasing values of the dyadic transmission probability r and clustering (see Legend). Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. $n = 2,029$ agents. Solid line: *Degree-Calibrated* (DC) networks; dashed line: *Erdős–Rényi* (ER) network with the same average degree.
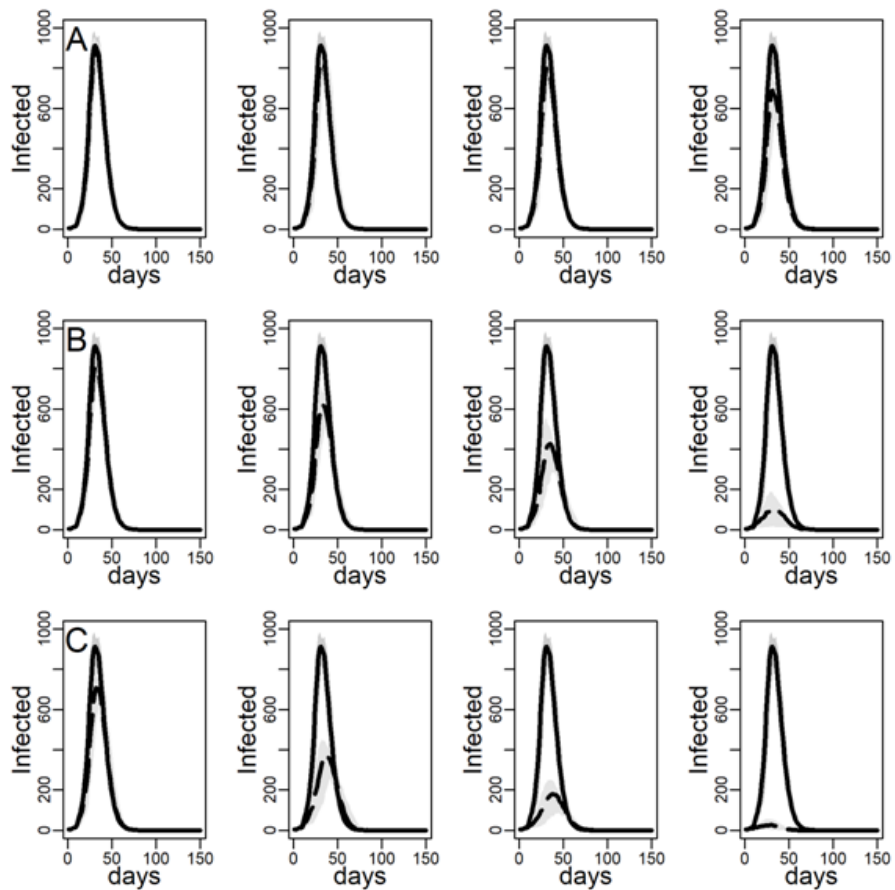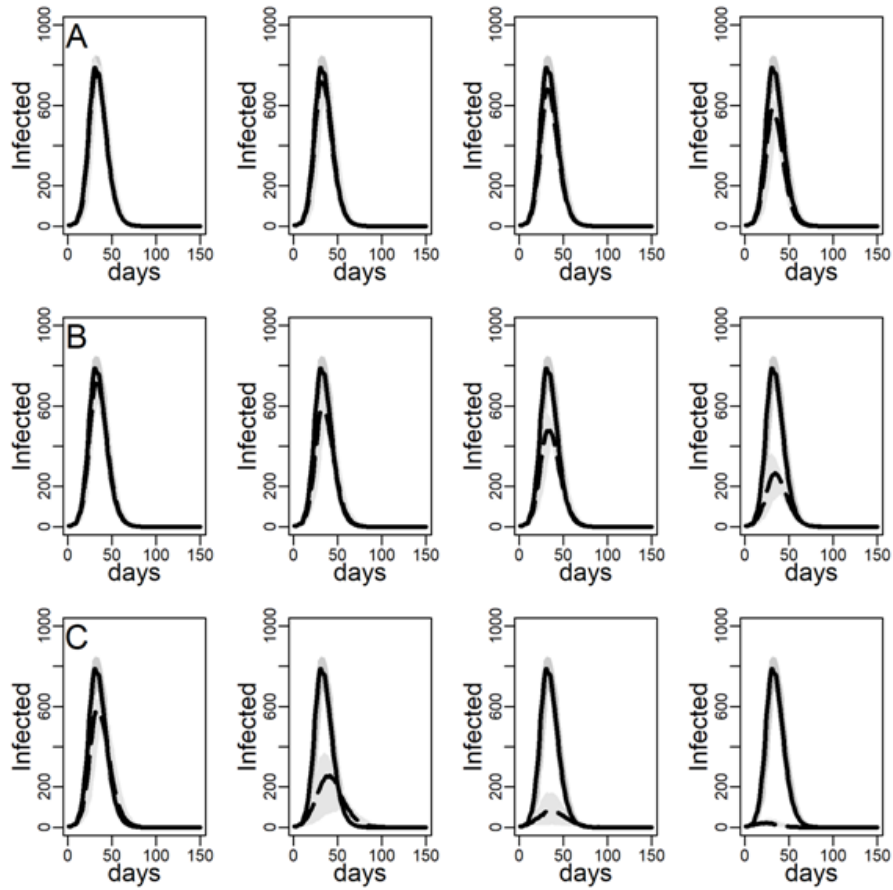
Figure 18: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: Degree-Calibrated (DC) network; dashed line: interventions. Dyadic transmission probability ($r = 0.05$) & Local clustering ($C_c = 0.04$). $n = 2,029$ agents.
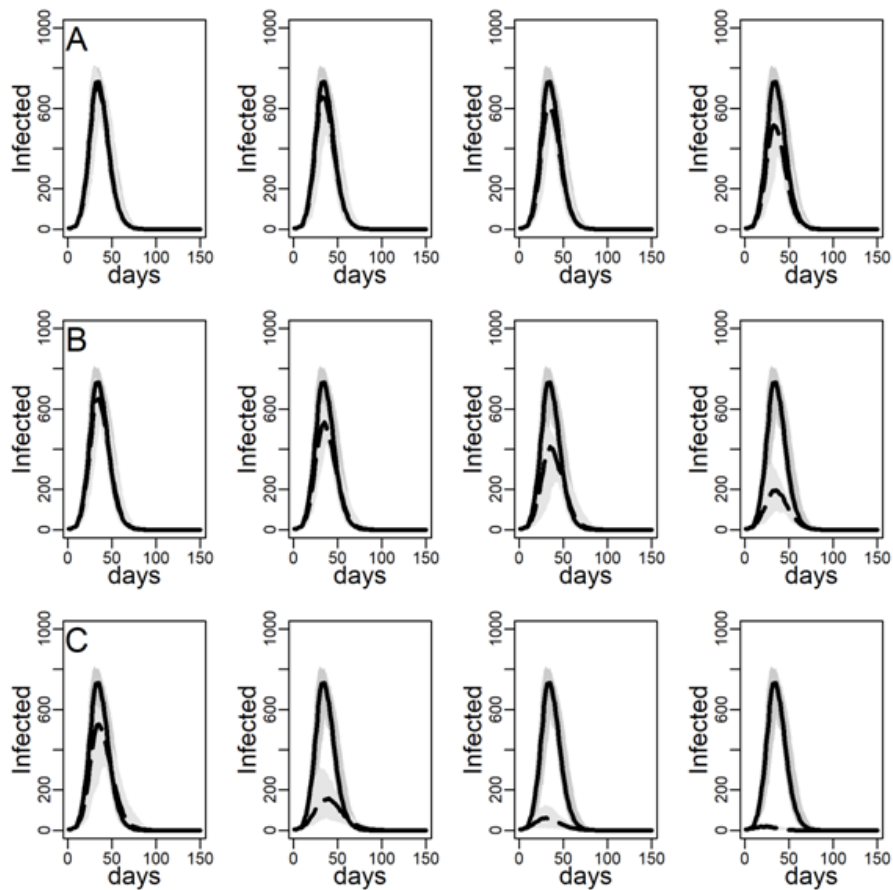
Figure 19: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability ($r = 0.05$) & Local clustering ($C_c = 0.42$). $n = 2{,}029$ agents.

Figure 20: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability ($r = 0.05$) & Local clustering ($C_c = 0.50$). $n = 2,029$ agents.
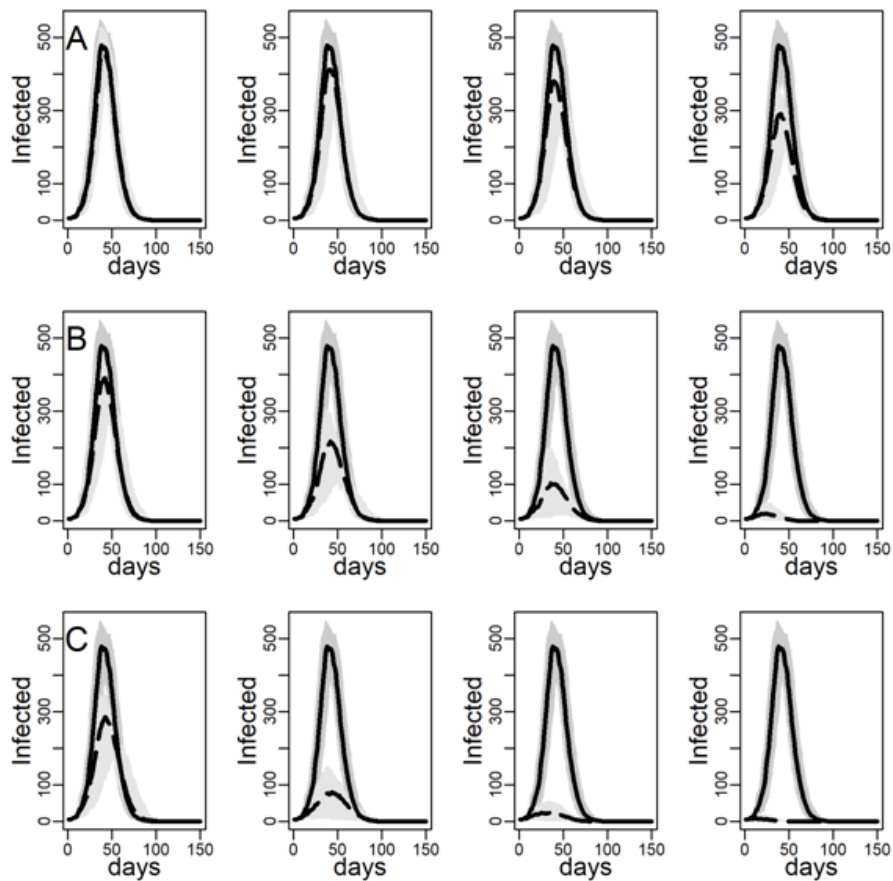
Figure 21: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability ($r = 0.03$) & Local clustering ($C_c = 0.04$). $n = 2,029$ agents.
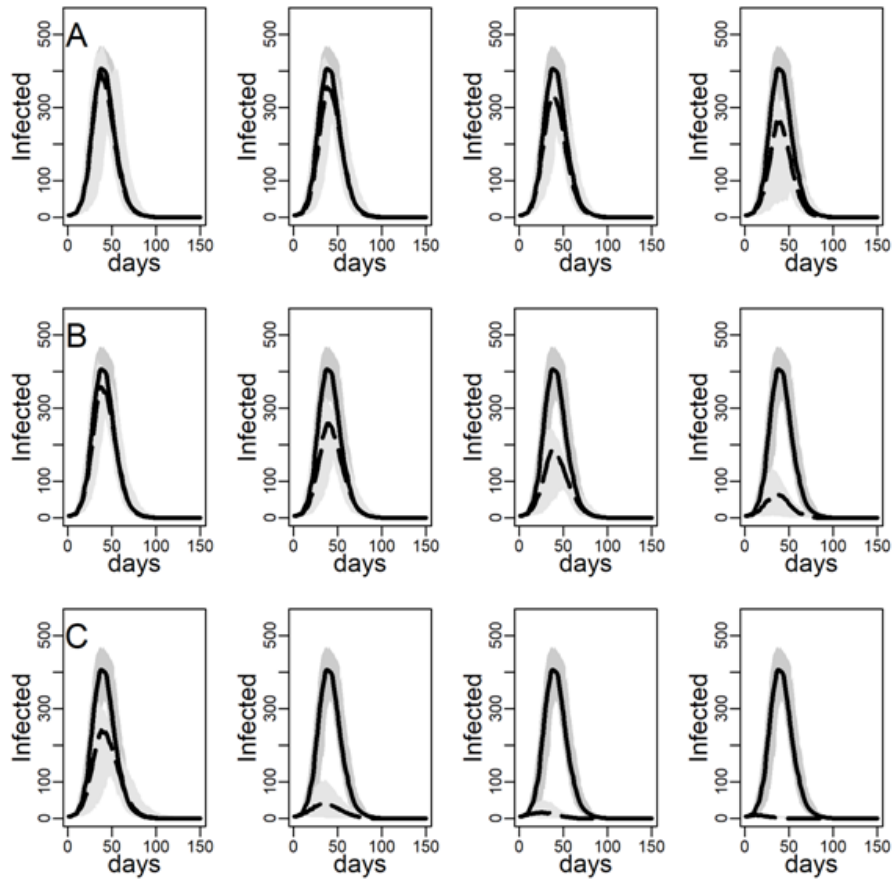
Figure 22: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability ($r = 0.03$) & Local clustering ($C_c = 0.42$). $n = 2,029$ agents.
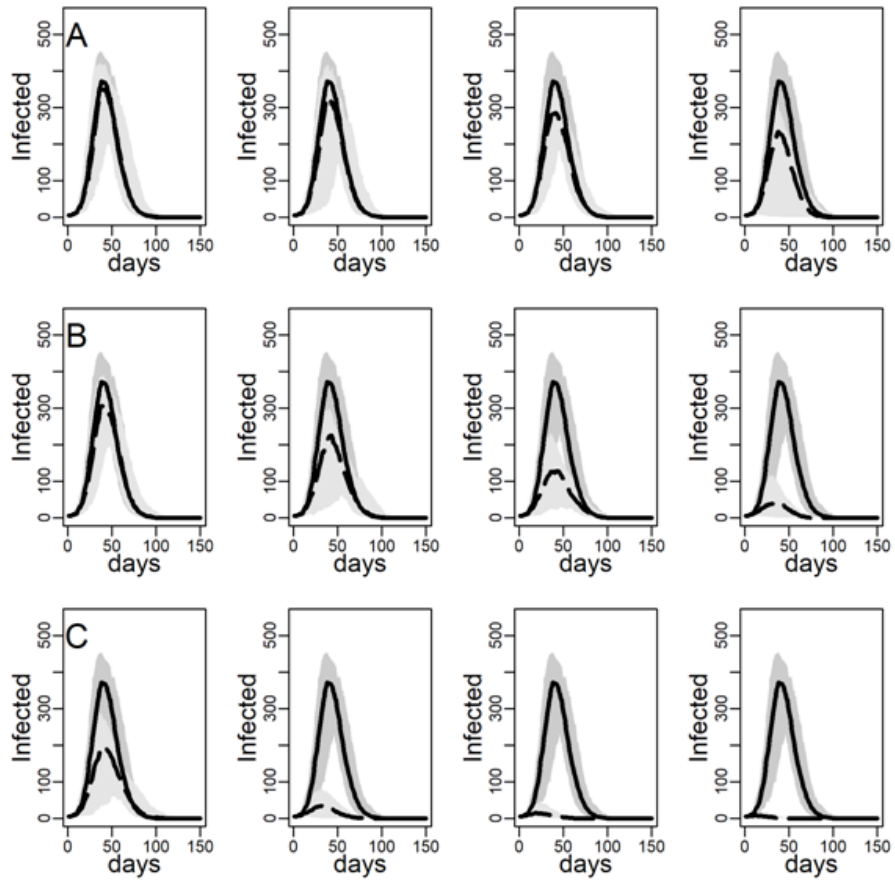
Figure 23: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability ($r = 0.03$) & Local clustering ($C_c = 0.50$). $n = 2{,}029$ agents.
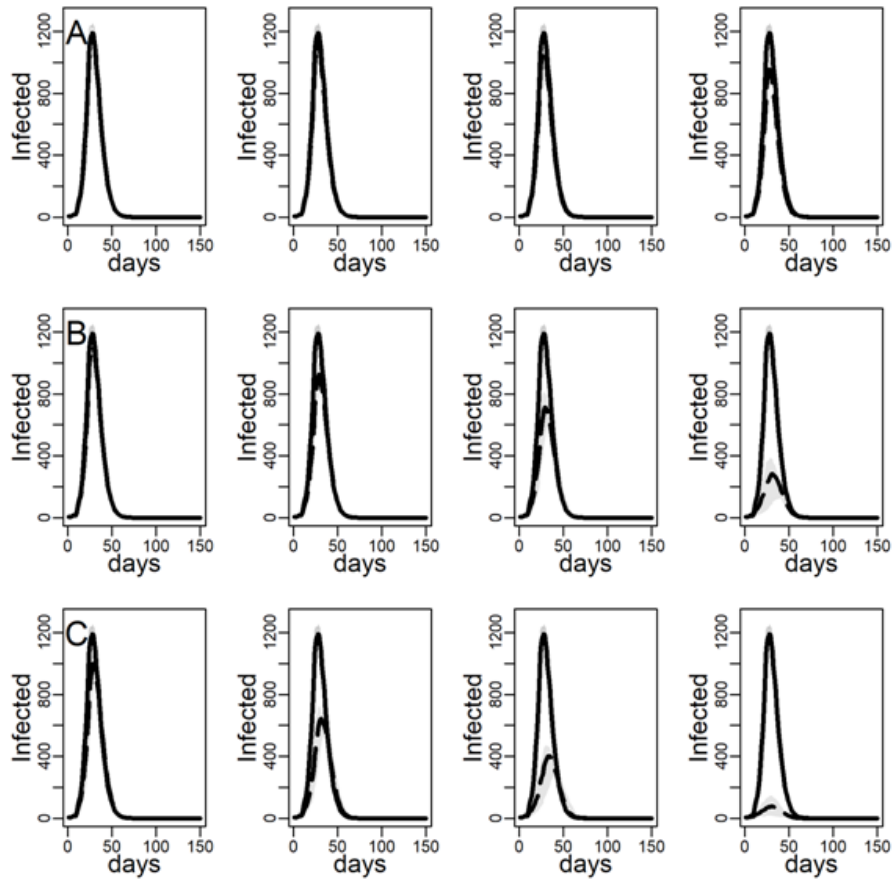
Figure 24: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability ($r = 0.07$) & Local clustering ($C_c = 0.04$). $n = 2,029$ agents.

Figure 25: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability ($r = 0.07$) & Local clustering ($C_c = 0.42$). $n = 2,029$ agents.
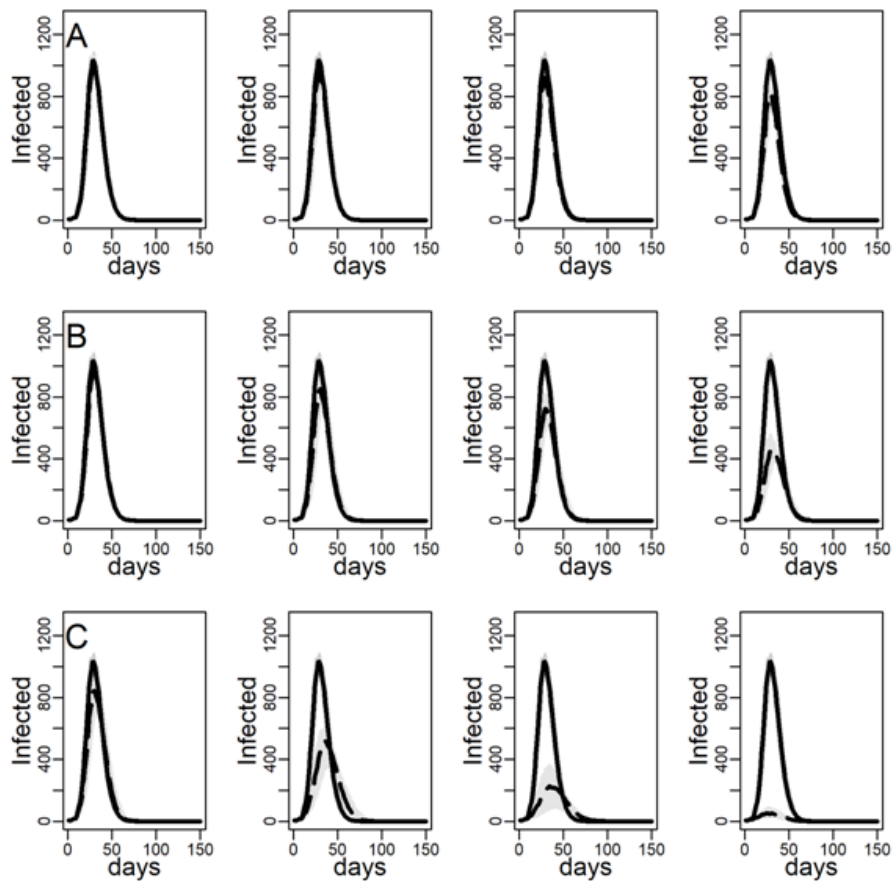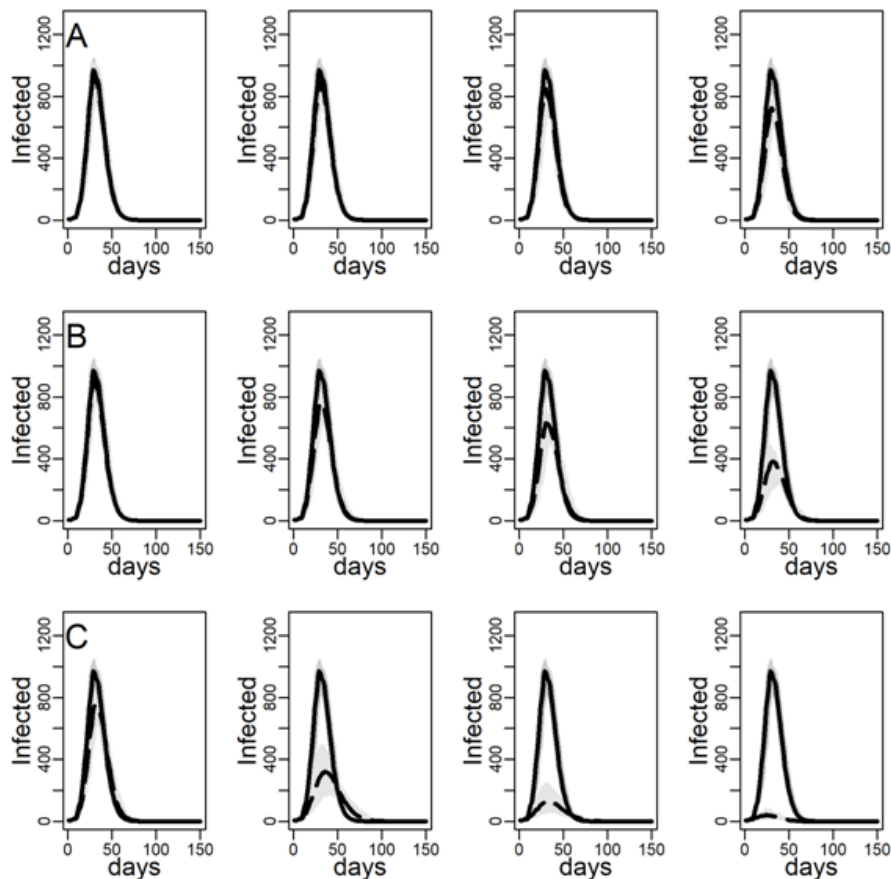
Figure 26: Number of infected agents (y-axis) by days (x-axis) (median of 100 replications) under three different interventions (rows) targeting 1, 3, 5, or 10 agents per day (columns). A – NO-TARGET intervention; B – CONTACT-TARGET intervention. C – HUB-TARGET intervention. Lower and upper bounds of the shaded areas correspond to the 5th percentiles and 95th percentiles of the 100 replications. Solid line: *Degree-Calibrated* (DC) network; dashed line: interventions. Dyadic transmission probability ($r = 0.07$) & Local clustering ($C_c = 0.50$). $n = 2,029$ agents.

References

Adam, D., Wu, P., Wong, J., Lau, E., Tsang, T., Cauchemez, S., Leung, G. & Cownling, B. (2020). Clustering and superspreading potential of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in Hong Kong. Preprint (Version 1). Available at:https://www.researchsquare.com/article/rs-29548/latest.pdf

Adamic, L. A. & Huberman, B. A. (2002). Zipf's law and the internet. *Glottometrics*, *3*(1), 143–150

Albert, R. & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Review of Modern Physics*, *74*(1), 47–97

Albert, R., Jeong, H. & Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, *406*(6794), 378–382

Banerjee, A., Breza, E., Chandrasekhar, A. G., Chowdhury, A. G., Duflo, E., Alsan, M., Goldsmith-Pinkham, P. & Olken, B. A. (2020). Messages on COVID-19 prevention in India increased symptoms reporting and adherence to preventive behaviors among 25 million recipients with similar effects on non-recipient members of their communities. National Bureau of Economic Research (No. w27496)

Barabási, A.-L. (2014). *Linked*. New York, NY: Basic Books

Barabási, A.-L. & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512

Barrat, A., Barthélemy, M. & Vespignani, A. (2008). *Dynamical Processes on Complex Networks*. Cambridge: Cambridge University Press

Barthélemy, M., Barrat, A., Pastor-Satorras, R. & Vespignani, A. (2005). Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *Journal of Theoretical Biology*, *235*(2), 275–288

Belot, M., Choi, S., Jamison, J., Papageorge, N. W., Tripodi, E. & van den Broek-Altenburg, E. (2020). Six-country survey on Covid-19. IZA Discussion Paper No. 13230. Available at SSRN: `https://ssrn.com/abstract=3596697`

Bèraud, G., Kazmerczak, S., Beutels, P., Levy-Bruhl, D., Lenne, X., Mielcarek, N., Yazdanpanah, Y., Boëlle, P. Y., Hens, N. & Dervaux, B. (2015). The French connection: The first large population-based contact survey in France relevant for the spread of infectious diseases. *PLoS ONE*, *10*(7), 1–22

Bi, Q., Yongsheng, W., Shujiang, M., Ye, C., Zou, X., Zhang, Z., Liu, X., Wei, L., Truelove, S. A., Zhang, T., Gao, W., Cheng, C., Tang, X., Wu, X., Sun, B., Huang, S., Sun, Y., Zhang, J., Ma, T., Lessler, J. & Feng, T. (2020). Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: A retrospective cohort study. *The Lancet Infectious Diseases*, *20*(8), 911–919

Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R. & Mills, M. C. (2020). Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nature Human Behavior*, *4*, 588–596

Brandily, P., Brebion, C., Briole, S. & Khoury, L. (2020). A poorly understood disease? The unequal distribution of excess mortality due to COVID-19 across French municipalities. MedRxiv 2020.07.09.20149955

Brethouwer, J. T., van de Rijt, A., Lindelauf, R. & Fokkink, R. (2020). "Stay nearby or get checked": A Covid-19 lockdown exit strategy. ArXiv preprint, arXiv:2004.06891

Broido, A. D. & Clauset, A. (2019). Scale-free networks are rare. *Nature Communications*, *10*(1)

Cencetti, G., Santin, G., Longa, A., Pigani, E., Barrat, A., Cattuto, C., Lehmann, S., Salathe, M. & Lepri, B. (2020). Digital proximity tracing in the COVID-19 pandemic on empirical contact networks. MedRxiv 2020.05.29.20115915

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Pastore y Piontti, A., Mu, K., Rossi, L., Sun, K., Viboud, C., Xiong, X., Halloran, M. E., Longini, I. M. & Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, *368*(6489), 395–400

Cho, S. Y., Kang, J. M., Ha, Y. E., Park, G. E., Lee, J. Y., Ko, J. H., Lee, J. Y., Kim, J. M., Kang, C. I., Jo, I. J., Ryu, J. G., Choi, J. R., Kim, S., Huh, H. S., Ki, C. S., Kang, E. S., Peck, K. R., Dhong, H. J., Song, J. H., Chung, D. R. & Kim, Y. J. (2016). MERS-CoV outbreak following a single patient exposure in an emergency room in South Korea: An epidemiological outbreak study. *The Lancet*, *388*(10048), 994–1001

Christakis, N. A. & Fowler, J. H. (2010). Social network sensors for early detection of contagious outbreaks. *PLoS ONE*, *5*(9), e12948

Clauset, A., Shalizi, C. R. & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661–703

Cohen, R. & Havlin, S. (2010). *Complex Networks: Structure, Robustness and Function*. Cambridge: Cambridge University Press

Cohen, R., Havlin, S. & Ben-Avraham, D. (2003). Efficient immunization strategies for computer networks and populations. *Physical Review Letters*, *91*(24), 247901

Danon, L., Read, J. M., House, T. A., Vernon, M. C. & Keeling, M. J. (2012). Social encounter networks: Collective properties and disease transmission. *Journal of the Royal Society Interface*, *9*(76), 2826–2833

Danon, L., Read, J. M., House, T. A., Vernon, M. C. & Keeling, M. J. (2013). Social encounter networks: Characterizing Great Britain. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1765), 20131037

Dezső, Z. & Barabási, A.-L. (2002). Halting viruses in scale-free networks. *Physical Review E*, *65*(5), 055103

Di Domenico, L., Pullano, G., Sabbatini, C. E., Boëlle, P.-Y. & Colizza, V. (????). Expected impact of reopening schools after lockdown on COVID-19 epidemic in Île-de-France

Dunbar, R. I. M. (2016). Do online social media cut through the constraints that limit the size of offline social networks? *Royal Society Open Science*, *3*(1), 3150292

Eguíluz, V. M. & Klemm, K. (2002). Epidemic threshold in structured scale-free networks. *Physical Review Letters*, *89*(10), 108701

Endo, A., Abbott, S., Kucharski, A. J., Funk, S. & Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group (2020). Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Research*, *5*, 67

Feld, S. L. (1991). Why your friends have more friends than you do. *American Journal of Sociology*, *96*(6), 1464–1477

Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., Monod, M., Whittaker, C., Zhu, H., Berah, T., Eaton, J. W., Ghani, A. C., Donnelly, C. A., Riley, S., Vollmer, M. A. C., Ferguson, N. M., Okell, L. C., Bhatt, S. & Imperial College COVID-19 Response Team (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, *584*(7820), 257–261

Galvani, A. P. & May, R. M. (2005). Dimensions of superspreading. *Nature*, *438*(7066), 293–295

Goolsbee, A. & Syverson, C. (2020). Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline. NBER Working Paper, No. 27432 (June 2020)

Hamner, L., Dubbel, P., Capron, I., Ross, A., Jordan, A., Lee, J., Lynn, J., Ball, A., Narwal, S., Russell, S., Patrick, D. & Leibrand, H. (2020). High SARS-CoV-2 attack rate following exposure at a choir practice — Skagit County, Washington, March 2020. *Morbidity and Mortality Weekly Report*, *69*(19), 606–610

Herrmann, H. A. & Schwartz, J.-M. (2020a). Using network science to propose strategies for effectively dealing with pandemics: The COVID-19 example. MedRxiv

Herrmann, H. A. & Schwartz, J. M. (2020b). Using network science to propose strategies for effectively dealing with pandemics: The COVID-19 example. MedRxiv

Hoang, T., Coletti, P., Melegaro, A., Wallinga, J., Grijalva, C. G., Edmunds, J. W. & Hens, N. (2019). A systematic review of social contact surveys to inform transmission models of close-contact infections. *Epidemiology*, *30*(5), 723–736

Hodcroft, E. B. (2020). Preliminary case report on the SARS-CoV-2 cluster in the UK, France, and Spain. *Swiss Medical Weekly*, *150*

Jackson, M. O. (2008). *Social and Economic Networks*. Princeton, NJ: Princeton University Press

James, A., Pitchford, J. W. & Plank, M. J. (2007). An event-based model of superspreading in epidemics. *Proceedings of Biological Sciences*, *274*(1610), 741–747

Jones, J. H. & Handcock, M. S. (2003). Sexual contacts and epidemic thresholds. *Nature*, *423*(6940), 605–606

Kay, J. (2020). COVID-19 superspreader events in 28 countries: Critical patterns and lessons. Quilette, Available at: `https://quillette.com/2020/04/23/`. Retrieved April 23, 2020

Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E. & Makse, H. A. (2010). Identification of influential spreaders in complex networks. *Nature Physics*, *6*(11), 888–893

Kucharski, A. J., Russell, T. W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R. M. & Centre for Mathematical Modelling of Infectious Diseases COVID-19 working group (2020). Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *The Lancet Infectious Diseases*, *20*(5), 553–558

Li, Q., Xuhua, G., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K. S. M., Lau, E. H. Y., Wong, J. Y., Xing, X., Zhao, J., Liu, M., Tu, W., Chen, C., Li, D., Liu, T., Zhao, J., Liu, M., Tu, W., Chen, C., Jin, L., Yang, R., Waang, Q., Zhou, S., Wang, R., Liu, H., Luo, Y., Liu, Y., Shao, G., Li, H., Tao, Z., Yang, Y., Deng, Z., Liu, B., Ma, Z., Zhang, Y., Shi, G., Lam, T. T. Y., Wu, J. T., Gao, G. F., Cowling, B. J., Yang, B., Leung, G. M. & Feng, Z. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *The New England Journal of Medicine*, *382*(13), 1199–1207

Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E. & Aberg, Y. (2001). The web of human sexual contacts. *Nature*, *411*(6840), 907–908

Little, S. J., Pond, S. L. K., Anderson, C. M., Young, J. A., Wertheim, J. O., Mehta, S. R., May, S. & Smith, D. M. (2014). Using HIV networks to inform real time prevention interventions. *PLoS ONE*, *9*(6), e98443

Lloyd-Smith, J., Schreiber, S., Kopp, P. & Getz, W. M. (2005). Superspreading and the effect of individual variation on disease emergence. *Nature*, *438*(7066), 355–359

Manzo, G. (2020). Complex social networks are missing in the dominant epidemic COVID-19 epidemic models. *Sociologica*, *14*(1), 31–49

Manzo, G., Gabbriellini, S., Roux, V. & M'mbogori, F. N. (2018). Complex contagions and the diffusion of innovations: Evidence from a small-N study. *Journal of Archaeological Method and Theory*, *25*(4), 1109–1154

Marcus, J. (2020). Quarantine fatigue is real. The Atlantic, 11 May 2020. Available at: https://www.theatlantic.com/ideas/archive/2020/05/quarantine-fatigue-real-and-shaming-people-wont-help/611482/

Martcheva, M. (2015). *An Introduction to Mathematical Epidemiology (Vol. 61)*. New York, NY: Springer

Meidan, D., Schulmann, N., Cohen, R., Haber, S., Yaniv, E., Sarid, R. & Barzel, B. (2020). Alternating quarantine for sustainable mitigation of COVID-19. Preprint, arXiv:2004.01453v2

Miller, D., Martin, M. A., Harel, N., Kustin, T., Tirosh, O., Meir, M., Sorek, N., Gefen-Halevi, S., Amit, S., Vorontsov, O., Wolf, D., Peretz, A., Shemer-Avni, Y., Roif-Kaminsky, D., Kopelman, N., Huppert, A., Koelle, K. & Stern, A. (2020). Full genome viral sequences inform patterns of SARS-CoV-2 spread into and within Israel. Preprint medRxiv 2020.05.21.20104521

Mittal, R., Ni, R. & Seo, J.-E. (2020). The flow physics of COVID-19. *Journal of Fluid Mechanics*, *894*, F2

Molina, C. & Stone, L. (2012). Modelling the spread of diseases in clustered networks. *Journal of Theoretical Biology*, *315*, 110–118

Mones, E., Stopczynski, A., Pentland, A. S., Hupert, N. & Lehmann, S. (2018). Optimizing targeted vaccination across cyber–physical networks: An empirically based mathematical simulation study. *Journal of The Royal Society Interface*, *15*(138), 20170783

Montes, F., Jaramillo, A. M., Meisel, J. D., Diaz-Guilera, A., Valdivia, J. A., Sarmiento, O. L. & Zarama, R. (2020). Benchmarking seeding strategies for spreading processes in social networks: An interplay between influencers, topologies and sizes. *Scientific Reports*, *10*(1), 1–12

Morawska, L. & Milton, D. K. (2020). It is time to address airborne transmission of coronavirus disease 2019 (COVID-19). *Clinical Infectious Diseases*, (p. ciaa939)

Mossong, J., Hens, N., Jit, M., Auranen, K., Beutels, P., Mikolajczyk, R., Massari, M., Salmaso, S., Scalia Tomba, G., Wallinga, J., Heijne, J., Sadkowska-Todys, M., Rosinska, M. & Edmunds, J. W. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine*, *5*(3), e74

Newman, M. E. (2002). Spread of epidemic disease on networks. *Physical Review E*, *66*(1), 016128

Newman, M. E. (2003). The structure and function of complex networks. *SIAM Review*, *45*(2), 167–256

Noar, S. M., Benac, C. N. & Harris, M. S. (2007). Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychological Bulletin*, *133*(4), 673–693

Olinky, R. & Stone, L. (2004). Unexpected epidemic thresholds in heterogeneous networks: The role of disease transmission. *Physical Review E*, *70*(3), 030902

Pastor-Satorras, R. & Vespignani, A. (2002). Immunization of complex networks. *Physical Review E*, *65*(3), 036104

Peretti-Watel, P. & the COCONEL Group (2002). A future vaccination campaign against COVID-19 at risk of vaccine hesitancy and politicization. *The Lancet Infectious Diseases*, *20*(7), 769–770

Pichler, A., Pangallo, M., del Rio-Chanona, M., Lafond, F. & Farmer, J. D. (2020). Production networks and epidemic spreading: How to restart the UK economy? Preprint, Available at: https://arxiv.org/pdf/2005.10585.pdf

Prem, K., Cook, A. R. & Jit, M. (2017). Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Computational Biology*, *13*(9), e1005697

Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M., Davies, N., Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Jit, M. & Klepac, P. (2020). The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study. *The Lancet Public Health*, *5*(5), 261–270

Rosenblatt, S. F., Smith, J. A., Gauthier, G. R. & Hébert-Dufresne, L. (2020). Immunization strategies in networks with missing data. Preprint, arXiv, 2005.07632

Roux, J., Massonnaud, C. & Crépey, P. (2020). COVID-19: One-month impact of the French lockdown on the epidemic burden. Preprint, Available at: `http://medrxiv.org/lookup/doi/10.1101/2020.04.22.20075705`

Salje, H., Kiem, C. T., Lefrancq, N., Courtejoie, N., Bosetti, P., Paireau, J., Andronico, A., Hozé, N., Richet, J., Dubost, C. L., Le Strat, Y., Lessler, J., Levy-Bruhl, D., Fontanet, A., Opatowski, L., Boelle, P. Y. & Cauchemez, S. (2020a). Estimating the burden of SARS-CoV-2 in France. Preprint, Available at: `https://www.medrxiv.org/content/10.1101/2020.04.20.20072413v2`

Salje, H., Kiem, C. T., Lefrancq, N., Courtejoie, N., Bosetti, P., Paireau, J., Andronico, A., Hozé, N., Richet, J., Dubost, C. L., Le Strat, Y., Lessler, J., Levy-Bruhl, D., Fontanet, A., Opatowski, L., Boelle, P. Y. & Cauchemez, S. (2020b). Estimating the burden of SARS-CoV-2 in France. *Science*, *369*(6500), 208–211

Sapiezynski, P., Stopczynski, A., Lassen, D. D. & Lehmann, S. (2019). Interaction data from the Copenhagen networks study. *Scientific Data*, *6*(1), 1–10

Serrano, M. A. & Marián, B. (2006). Percolation and epidemic thresholds in clustered networks. *Physical Review Letters*, *97*(8), 088701

Smith, J. A. & Burrow, J. (2018). Using ego network data to inform agent-based models of diffusion. *Sociological Methods & Research*, (p. 0049124118769100)

Stein, R. A. (2011). Super-spreaders in infectious diseases. *International Journal of Infectious Diseases*, *15*(8), 510–513

Stumpf, M. P. & Porter, M. A. (2012). Critical truths about power laws. *Science*, *335*(6069), 665–666

Sun, L., Axhausen, K. W., Lee, D. H. & Cebrian, M. (2014). Efficient detection of contagious outbreaks in massive metropolitan encounter networks. *Scientific Reports*, *4*(5099)

Tolles, J. & Luong, T. (2020). Modeling epidemics with compartmental models. *JAMA*, *323*(24), 2515–2516

Trewick, C., Ranchod, P. & Konidaris, G. (2013). Preferential targeting of HIV infected hubs in a scale-free sexual network. Presented at the Annual Conference of the Computational Social Science Society of the Americas, August 2013

Vázquez, A., Pastor-Satorras, R. & Vespignani, A. (2002). Large-scale topological and dynamical properties of the Internet. *Physical Review E*, *65*(6), 066130

Walker, P. G. T., Whittaker, C., Watson, O., Baguelin, M., Winskill, P., Hamlet, A., Djafaara, B. A., Cucunbá, Z., Bhatia, S., Bhatt, S., Boonyasiri, A., Boyd, O., Brazeau, N. F., Cattarino, L., Cuomo-Dannenburg, G., Dighe, A., Donnelly, C. A., Dorigatti, I., van Elsland, S. L., Fitzjohn, R., Fu, H., Gaythorpe, K. A. M., Geidelberg, L., Grassly, N., Haw, D., Hayes, S., Hinsley, W., Imai, N., Jorgensen, D., Knock, E., Laydon, D., Mishra, S., Nedjati-Gilani, G., Okell, L. C., Unwin, H. J., Verity, R., Vollmer, M., Walters, C. E., Wang, H., Wang, Y., Xi, X., Lalloo, D. G., Ferguson, N. M. & Ghani, A. C. (2020). The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science*, *369*(6502), 413–422

Watts, D. J. & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), 440–442

Wong, G., Liu, W., Zhou, B., Bi, Y. & Gao, G. F. (2015). MERS, SARS, and ebola: The role of super-spreaders in infectious disease. *Cell Host & Microbe*, *18*(4), 398–401

Woolhouse, M. E., Dye, C., Etard, J. F., Smith, T., Charlwood, J. D., Garnett, G. P., Watts, C. H., Hagan, P., Hii, J. L. K., Ndhlovu, P. D., Quinnell, R. J., Chandiwana, S. K. & Anderson, R. M. (1997). Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(1), 338–342

World Health Organization (2020). Report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19). Available at: `https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf`

Zhang, J., Klepac, P., Read, J. M., Rosello, A., Wang, X., Lai, S. & Yang, J. (2019). Patterns of human social contact and contact with animals in Shanghai, China. *Scientific Reports*, *9*(1), 1–11

Zhang, J., Litvinova, M., Liang, Y., Wang, Y., Wang, W. & Ajelli, M. (2020). Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science*, *368*(6498), 1481–1486