Modeling the impact of testing, contact tracing and household quarantine on second waves of COVID-19

Alberto Aleta[1], David Martín-Corral[2,3], Ana Pastore y Piontti [4], Marco Ajelli [5,6], Maria Litvinova[1], Matteo Chinazzi [4], Natalie E. Dean [7], M. Elizabeth Halloran [8,9], Ira M. Longini, Jr.[7], Stefano Merler[5], Alex Pentland[10], Alessandro Vespignani[4,1,*], Esteban Moro[2,10,*] & Yamir Moreno[11,12,1,*]

[1]*Institute for Scientific Interchange Foundation, Turin, Italy*

[2]*Department of Mathematics and GISC, Universidad Carlos III de Madrid, Leganés, Spain.*

[3]*Zensei Technologies S.L., Madrid, Spain.*

[4]*Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston, MA, USA.*

[5]*Bruno Kessler Foundation, Trento Italy*

[6]*Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA*

[7]*Department of Biostatistics, College of Public Health and Health Professions, University of Florida, Gainesville, FL, USA*

[8]*Fred Hutchinson Cancer Research Center, Seattle, WA, USA*

[9]*Department of Biostatistics, University of Washington, Seattle, WA, USA*

[10]*Connection Science, Institute for Data Science and Society, MIT, Cambridge, US*

[11]*Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Spain*

[12]*Department of Theoretical Physics, Faculty of Sciences, University of Zaragoza, Spain*

*Corresponding authors.

While severe social distancing measures have proven effective in slowing the COVID-19 pandemic, second-wave scenarios are likely to emerge as restrictions are lifted. Here, we integrate anonymized, geo-localized mobility data with census and demographic data to build a detailed agent-based model of SARS-CoV-2 transmission in the Boston metropolitan area. We find that a period of strict social distancing followed by a robust level of testing, contact-tracing, and household quarantine could keep the disease within the capacity of the health care system while enabling the reopening of economic activities. Our results show that a response system based on enhanced testing and contact tracing can play a major role in relaxing social distancing interventions in the absence of herd immunity against SARS-CoV-2.

Introduction

The first report of a new infectious disease, later coined COVID-19, appeared on 31 December 2019[1]. As of 15 July 2020, the virus has spread to 188 countries with more than 13.3 millions confirmed cases worldwide, and killing more than 579,500 people[2]. As the number of confirmed COVID-19 cases increased and the expansion of the disease entered into a global exponential growth phase, a large number of affected countries were forced to adopt non-pharmaceutical interventions at an unprecedented scale. Given the absence of specific antiviral prophylaxis, therapeutics or a vaccine, non-pharmaceutical interventions ranging from case isolation and quarantine of contacts, to the lock-down of entire populations have been implemented with the aim of suppressing/mitigating the epidemic before it could overwhelm the health care system. Although these ag-

gressive measures appear to be successful in reducing the number of deaths and hospitalizations[3,4], and in reducing the transmission of the SARS-CoV-2 virus, the absence of herd immunity after the first wave of the epidemic points to a large risk of resurgence when interventions are relaxed and societies go back to a "business as usual" lifestyle[5–7]. It is therefore of paramount importance to analyze different mitigation and containment strategies aimed at minimizing the risk of potential additional waves of the COVID-19 epidemic while providing an acceptable trade-off between economic and public health objectives.

In the present work, through the integration of anonymized and privacy-enhanced data from mobile devices and census data, we build a detailed sample of the synthetic population of the Boston metropolitan area in the United States (see Figure 1a and 1b). This synthetic population (Figure 1a) is used to define a data-driven agent-based model of SARS-CoV-2 transmission and to provide a quantitative analysis of the evolution of the epidemic and the effectiveness of social distancing interventions. The model allows us to explore strategies concerning the lifting of social distancing interventions in conjunction with testing and isolation of cases and tracing and quarantine of exposed contacts. Our results indicate that after the abatement of the epidemic through the "stay at home" orders and halt to all nonessential activities, a proactive policy of testing, contact tracing and contacts' household quarantine allows the gradual reopening of economic activities and workplaces, with a low COVID-19 incidence in the population and a manageable impact on the health care system.

To provide a quantitative estimate of the contact patterns for the population of agents and

to build the synthetic population of the Boston Metropolitan Area (BMA), we used detailed mobility and socio-demographic data and generated a network that encodes the contact patterns of about 85,000 agents in the area during a period of six months (see Supplementary Material, SM). Agents are chosen to be representative of the different census areas in the Boston area following the methodology used in Ref. [8]. This defines a weighted multilayer network consisting of three layers representing the network of social interactions at (1) workplace/community level (W+C), (2) households, and (3) schools, as shown in Figure 1a. Connections between two agents in the W+C layer are estimated from the data by the probability of both being present in a specific place (e.g. restaurant, workplace, shopping) weighted according to the time they have spent in the same place. A second layer represents the households of each anonymous individual. Using the home census block group of each anonymous user we associate each individual to a specific household profile based on socio-demographic data at US census block group level[9]. Families are generated by randomly mixing nodes from the community living in the same census block group, following the statistical features of family types and sizes. Finally, a third layer represents the contacts in the schools (i.e., every node represents one synthetic student and has contacts only with other individuals attending the same school). To study the evolving dynamics of the infection, we implemented a stochastic, discrete-time compartmental model in which individuals transition from one state to the other according to key time-to-event intervals (e.g., incubation period, serial interval, and time from symptom onset to hospital admission) as from available data on SARS-CoV-2 transmission. The natural history of the disease is captured by the epidemiological model represented in Figure 1c, where we also show the transition rates among compartments[8,10–12]. The model considers that

susceptible individuals (S) become infected through contact with any of the infectious categories (infectious symptomatic (I_S), infectious asymptomatic (I_A) and pre-symptomatic (P_S)), transitioning to latent compartments (L_S) and (L_A), where they are infected but not infectious yet. Latent individuals branch out in two paths according to whether the infection will be symptomatic or not. We also consider that symptomatic individuals experience a pre-symptomatic phase and that once they develop symptoms, they can experience diverse degrees of illness severity, from mild symptoms to being hospitalized (H) or in need of an intensive care unit (ICU)[13]. Finally, individuals transition in the removed compartment (identifying recovered or dead individuals). The model assumes a basic reproductive number $R_0 = 2.5$, which together with the rest of the parameters (see SM, table S1) yields a generation time $T_g = 6.6$ days. We consider a 25% fraction of asymptomatic individuals. We report the full set of parameters used in the model and an extensive sensitivity analysis in the SM file. The model is not calibrated to account for the specific evolution of the COVID-19 epidemic in Boston as it is aimed at showing the effect of different non-pharmaceutical interventions rather than providing a forensic analysis of the outbreak in the BMA. Furthermore current data on the detailed reopening policies in the BMA during the summer are still under evaluation pending the future evaluation of the epidemic trajectory. Details on the generation of the synthetic population network and the infection transmission model are provided in the SM.

Results

To provide a baseline of the COVID-19 impact in the Boston metropolitan area, we have first investigated an unmitigated scenario in which no interventions are implemented. Results for the unmitigated scenario are shown in Figure 2, panels a-c. A COVID-19 unmitigated epidemic would have a peak of daily incidence of 25.2 (95% C.I: 23.8-26.4) newly infected individuals per 1,000 people. The epidemic follows a typical trajectory, namely, when the effective reproduction number R_t as a function of time (panel c) becomes smaller than 1, the transmission dynamics slow down and eventually vanish after having infected about 75% of the population (Figure 2b). Figure 3a shows the evolution of the estimated number of new severely affected patients who require hospitalization and admission into ICUs. At the peak of the unmitigated epidemic, the number of ICU beds needed exceeds by far the available capacity (dashed horizontal line in Figure 3a) by more than a factor of 10, thus indicating that the health care system would suffer large service disruptions, resulting in additional deaths due to hospitals overcrowded with patients with COVID-19[14]. It is worth noting that current estimated fatality rates consider the general availability of ICU beds and critical care capacity. If this would not be possible the fatality rate may increase dramatically. Here we do not report fatality estimates as it goes beyond the scope of our analysis and should consider specific data on the BMA, as well as changing medical treatment and therapeutics in future months.

To avoid the harmful effects of an unmitigated COVID-19 epidemic, governments and policy makers across the world have relied on the introduction of aggressive social distancing measures.

In the United States, as of April 15 2020, it was estimated that more than 95% of the population was under a "stay at home" or "shelter in place" order[15,16]. To model the social distancing policies implemented in the whole Boston metropolitan area, we have considered March 17, 2020 as the average starting date of social distancing policies that include school closures, the shut down of all non-essential work activities as well as mobility restrictions (see SM for details). This scenario mimics the social distancing intervention implemented in most of the high income countries, in Europe and across states in the US. Such extreme social distancing policies come with very large economic costs and social disruption effects[17], thus prompting the question of what exit strategy can be devised to restart economic activities and normal societal functions[18]. For this reason, we explore two different scenarios for lifting social distancing interventions:

- Lift scenario (LIFT): the "stay at home" order is lifted after 8 weeks by re-opening all work and community places, except for mass-gathering locations such as restaurants, theaters, and similar locations (see SM). The latter partial re-opening is enforced for another 4 weeks, which is followed by a full lifting of all the restrictions that remained. We consider that schools will remain closed given the impending summer break in July and August, 2020. In fact, some school systems, like the Boston Public schools, have announced they will remain closed through the 2019-2020 school year.

- Lift and enhanced tracing (LET) scenario: The "stay at home" order is lifted as in the previous scenario. Once partial reopening is implemented, we assume that 50% of symptomatic COVID-19 cases can be identified for SARS-CoV-2 infection, on average, within 2 days

after the onset of symptoms and that they are isolated at home and their household members are quarantined successfully for 2 weeks (a sensitivity analysis for lower rate of isolation and quarantine is presented in the SM). Although COVID-19 tests are highly specific, 50% detection accounts also for imperfect testing. We also assume that a fraction of the non-household contacts (we show results for 20% and 40%) of the symptomatic infections can be traced and quarantined along with their household as well − note that we consider that the contacts are identified with a rate proportional to the time duration of the interaction with the symptomatic individual.

The above scenarios are mechanistically simulated on the multilayer network of Figure 1a, by allowing different interactions (between effective contacts) according to the simulated strategy. As a result, the average number of interactions in the W+C layer goes from 10.86 (95% C.I.:1.51-42.39) under the unmitigated scenario, to 4.10 (95% C.I.:0-23.79) for the partial lock-down and only 0.89 (95% C.I.: 0-8.39) contacts for the stay at home policy, see Methods and SM for more details. This result is in agreement with previously published work[19] and recent reports in the New York City area[20]. It is worth remarking that the fluctuations in the number of contacts in the stay at home order is due to a large extent to contacts that take place in grocery stores and other public venues.

The numerical results show that the LIFT scenario, while able to temporally abate the epidemic incidence, does not prevent the resurgence of the epidemic and a second COVID-19 wave when the social distancing measures are relaxed. In Figure 2d, we show that following the lifting

of social distancing the infection incidence starts to increase again, and the effective reproductive number, that dropped by circa 75% and reached values below 1 with the intervention, increases to values up to 2.05 (95%CI: 1.73-2.47) (see Figure 2f). Indeed, at the time of lifting the social distancing intervention the population has not achieved the level of herd immunity that would protect it from the resurgence of the epidemic. It is important to stress that here we do not consider additional mitigation measures such as behavioral changes in the population, mask wearing, etc (see the SM). We also estimate that a second wave of the epidemic still has the potential to infect a large fraction of the population (Figure 2e) and to overwhelm the health care systems, as shown in Figure 3b. The number of ICU beds needed, although half the unmitigated scenario, is still exceeding by far the estimated availability, as pointed out in similar scenario analysis [5–7,21]. This suggests that lifting social distancing without the support of additional containment strategies is not a viable option.

In the case of the LET scenario, the lifting of the social distancing intervention goes along with a significant amount of contact tracing and precautionary quarantine of potentially exposed individuals. The quarantine is not limited to the contacts of the identified symptomatic COVID-19 case, but extended to their households. This strategy amounts to a simplified tracing of contacts of contacts, that would not require extensive investigations. In other words, this strategy does not require the tracking of a large number of single contacts but leverage on the contacts' households as the basic unit [22]. Households could be monitored though, with daily calls or messages to ascertain the onset of symptomatic infections, and provide medical support as needed.

Figure 2g shows results obtained for different levels of tracing (no tracing, 20% and 40%) of the contacts of the symptomatic isolated COVID-19 cases. By comparing Figure 2d with Figure 2g (for no tracing), we find that quarantining households of symptomatic subjects alone is not enough to significantly change the course of the epidemic and the conclusions reached for the first of these scenarios. When 40% or more of the contacts of the detected symptomatic infections are traced and they and their households quarantined, the ensuing reduction in transmission leads to a noticeable flattening of the epidemic curve and appears to effectively limit the possible resurgence of a second epidemic wave. It is also worth noticing that we assume the absence of other additional and minimally disruptive social distancing policies such as crowd control, smart working, wearing of masks, etc., that could lead to a further reduction of the transmissibility of the virus with respect to our estimates. In the SM file we investigate the positive synergistic effect of the combination of the LET strategy in the presence of different levels of transmissibility reduction due to additional mitigation policies. It is important to stress that the contact tracing proposed here works at the level of household unit, simplifying also the monitoring and follow up process, by contacting only one member of the household to monitor the onset of symptoms among all members (we further explore other isolation/quarantine strategies in the SM). Figure 3c and Table1 show the burden in hospitalization and ICU demand in the unmitigated situation and the two mitigation scenarios. The LET scenario allows relaxation of the social distancing interventions while maintaining the hospital and ICU demand at levels close to the health-care availability and surge capacity. For the sake of completeness in the SM file we report the analysis for a LIFT scenario including the school and university reopening in the fall. The results show that in absence of additional containment

policies the tracing effort would need to be raised of about 50% to cope with the increased number of infections.

Discussion

The efforts in the suppression and mitigation of COVID-19 are pursuing the objectives of preserving the health care system from disruptive failures due to overwhelming stress imposed by the large number of severe cases, and of minimizing the morbidity and mortality related to the epidemic. The aggressive social distancing interventions implemented by many countries in response to the COVID-19 pandemic appear to have achieved the interruption of transmission and the abatement of the epidemic, although at the price of huge societal disruption and economic costs. In such a context, the identification of "exit strategies" that allow restarting economic and social activities while still protecting the healthcare systems and minimizing the burden of the epidemic is of primary importance. Several modeling studies have already pointed out that resuming economic activities and social life is likely to lead to a resurgence of the COVID-19 epidemic, and combined social distancing interventions of different degrees and intensity have been proposed to substantially delay and mitigate the epidemic[17,21]. These interventions still generate economic loss and widespread disruption to social life. Here we show how testing, contact tracing strategies at scale, based on home isolation of symptomatic COVID-19 cases and the quarantine of a fraction of their households' contacts, has the potential to provide a viable course of action to manage and mitigate the epidemic when social distancing interventions are progressively lifted[23,24]. These strategies present us with logistic challenges that include large-scale and rapid diagnostic capacity, and a

large surge in the number of contact tracers. We have investigated what fraction of the population would be isolated/quarantined under the proposed contact tracing and isolation strategy. Figure 4a shows the fraction of households that needs to be quarantined. Assuming the identification of 50% of the symptomatic infections, and tracing of 40% of their contacts and households, only about 9% of the population would be quarantined at any time. While this is certainly a relevant fraction of the population, it is a much better option if compared with massive social distancing policies affecting the entire population that last for months.

In Table 1, we report the number of symptomatic infections for which the contact tracing investigation should be performed in the basic scenarios. This number provides an estimate of the contact tracers per 1,000 individuals. It is important to note that the more effective is the contact tracing starting from each individual, the smaller is the number of generally traced households because the epidemic has lower incidence rates. In addition, as illustrated in Figure 4b, the health status of the vast majority of quarantined individuals is unknown as contact tracing does not imply testing. The curves in Figure 4a constitute the upper bounds for each simulated case. If we assume that the capacity to do massive testing would likely ramp up in the near future, then it is expected that the actual number of people in quarantine could be significantly lowered by testing the quarantined household. This would also alleviate the burden on household members that could not go to work and increase compliance of quarantine for the positive cases. It is also worth remarking that many of the logistic challenges faced with massive contact tracing might possibly be eased by digital technologies that are currently being investigated across the world following the examples of COVID-19 response in Asian countries[24]. Also it may be difficult to quarantine

the entire household of individuals who are potentially exposed, since this is a hardship suffered with great uncertainty about their risk of infection. Offering other logistic quarantine solutions (quarantine centers, hotel rooms) might significantly raise the rate of compliance.

These results were obtained under several assumptions. There are very large uncertainties around the transmission of SARS-CoV-2, in particular, the fraction of sub-clinical and asymptomatic cases and their transmission. Estimates of age-specific severity are informed from the analysis on individual-level data from China and other countries, and subject to change as more US data become available. We also do not include specific co-morbidities or pre-existing conditions of the specific BMA population. For this reason, in the SM we perform an extensive sensitivity analysis showing that the modeling results discussed here are robust to the plausible range of parameter values for the key time-to-event intervals of COVID-19 (e.g., incubation period, serial interval, and time from symptom onset to hospital admission, etc.) as well as the fraction of presymtpomatic and asymptomatic transmission. We are also not considering here potential changes to the virus transmissibility due to environmental factors, in particular, seasonal drivers such as temperature and humidity. The modeling does not consider possible reintroduction of SARS-CoV-2 in the population from infected travelers. Strategies based on testing, isolation and contact tracing might be hampered by the importation through travel of a large number of infections, thus travel restrictions and screening may need to be introduced to/from places that show sustained local transmission. Finally, we report in the SM the effect of the widespread use in the population of masks or other personal protective equipment that lead to a reduction of the transmissibility of SARS-Cov-2. These active protection measures improve the effectiveness of the exit strategies modeled here.

The modeling of the impact of testing, contact tracing and isolation on second-wave scenarios of the COVID-19 epidemic could be instrumental to national and international agencies for public health response planning. While we show that contact tracing and household quarantine at scale may be effective even assuming a complete lifting of the social distancing measures, future decisions on when and for how long to relax policies will need to be informed by ongoing surveillance. For instance, smart working from home for people who can adhere to it without serious disruptions should be encouraged. This, as well as other minimally disturbing policies together with efficient and large-scale testing, contact tracing and monitoring of the epidemic should be considered in the definition of exit strategies from large scale social-distancing policies.

Methods

Weighted synthetic population. Our synthetic population is constituted by circa 85,000 nodes, of which 64,000 are adults and 21,000 are children (defined as individuals aged 17 or less), see Figure 1b. The total number of interactions among these individuals before social distancing is given by more than 5M edges, see SM for a finer description.

1) Community weighted contact network. The community network is approximated using 6 months of data observation in the Boston area from anonymized users who have opted-in to provide access to their location data anonymously, through a GDPR-compliant framework provided by Cuebiq. Individuals performing the analysis were legally required to never single out identifiable individuals and not make attempts to link these data to third party data about an indi-

vidual. In this layer each agent in our synthetic population represents an anonymous individual of the real population. The data allow us to understand how infection can propagate in each layer by estimating co-location of two individuals in the same setting. We use a large database of 83,000 places from Foursquare API in the BMA. Specifically, the weight, $\omega_{C_{ij}}$, of a link between individuals i and j within the workplace plus community layer is computed according to the expression:

$$\omega_{C_{ij}} = \sum_{p}^{n} \frac{T_{ip}}{T_i} \frac{T_{jp}}{T_j}, \qquad \forall i, j$$

where T_{ip} is the total time that individual i was observed at place p and T_i is the total time that individual i has been observed at any place set within the workplace plus community layer. Note that while the mobility data set we use is large, co-location events between individuals are still quite sparse. Because of this sparsity, and to protect individual privacy in our analysis, we have adopted this probabilistic approach to measure co-presence in all locations mapped in the dataset. Since agents are representative of the different census areas and groups of the Boston area, our probabilistic approach is a good proxy for the real probability of co-presence between those groups/areas when networks are scaled up to the total population of the Boston area, that is approximately 4,628,910 inhabitants. Finally, for robustness and computational reasons, we have included only links for which $\omega_{C_{ij}} > 0.001$.

2) Household weighted contact network. We first localize individuals' approximate home place according to the US census block group. Then we assign a type of household based on Table B11016: Household Type by Household Size from US Census 2018[9], and mix individuals that live in the same block according to statistics of household type and size. Finally, children are

assigned to households (see the SM for a more detailed description). We also assign individuals an age group based on Table B01001: Sex and age from the US Census 2018. To assign weights, we assume that the probability of interaction within a household is proportional to the number of people living in the same household (well-mixing). Therefore, the weight, $\omega_{H_{ij}}$, of a link between individuals i and j within the same household is given by:

$$\omega_{H_{ij}} = \frac{1}{(n_h - 1)}$$

where n_h is the number of household members. This fraction is assumed to be the same for all individuals in the population.

3) School weighted contact network. To calculate the weights of the links at the school layer, we mix together all children that live in the same school catchment area. Interactions are considered well-mixed, hence, the probability of interaction at a school is proportional to the number of children at the same school. Therefore, the weight, $\omega_{S_{ij}}$, of a link between children i and j within the same school is given by:

$$\omega_{S_{ij}} = \frac{1}{(n_s - 1)}$$

where n_s is the number of school members.

Calibration of intra-layer links. Within each connected component of the network in each layer (e.g., a household, a school), the links between nodes are weighted to account for the effective daily number of contacts. For example, if we consider a school, while a student can potentially contact all her/his schoolmates, she/he only meet a relatively small fraction of them on a daily

basis as estimated in empirical studies on mixing patterns [25,26]. To account for this, we calibrate the weight of the links in each layer of the synthetic network [27] so that the mean number of daily contacts matches the estimation provided in Mistry et al. [28] (see SM for details). Based on the analysis of contact survey data from 9 countries[25,26,29,30], this study estimated the mean number of daily contacts at 10.86, 4.11 and 11.41 in the community+workplace, household and school layers, respectively.

Stochastic simulations of the COVID-19 dynamics. We describe the SARS-CoV-2 transmission process using a discrete time Susceptible-Latent-Infected-Removed (SLIR) stochastic model, with some extra compartments to incorporate the special characteristics of SARS-CoV-2 infection, Figure 1c. In particular, at each time-step t (1 day), the infectious asymptomatic (I_A), infectious symptomatic (I_S) and pre-symptomatic (P_S) individuals can transmit the disease to susceptible (S) subjects with probability $r\beta$, β and β_S, respectively. If the transmission is successful, the susceptible node will move to the latent asymptomatic state (L_A) with probability p or to the latent symptomatic state (L_S) with probability $(1-p)$. A latent asymptomatic individual becomes infectious asymptomatic after a period $(\epsilon^{'})^{-1}$, whereas latent symptomatic subjects transition, after a period ϵ^{-1}, to the pre-symptomatic (P_S) compartment. The average period to develop the disease and move to the infectious symptomatic state is γ^{-1}. Infectious asymptomatic nodes will be removed (R) after an average of μ steps. Conversely, infectious symptomatic nodes can either recover after that period with probability $(1-\alpha)$ or, with probability α, these nodes will need hospitalization. It is considered that due to their symptoms they will self-isolate at home after an

average period of μ^{-1}. Then, depending on the severity of the symptoms, after a period δ^{-1} the individual will end in hospitalization with probability $(1-\chi)$ or require hospitalization and ICU care with probability χ. Finally, individuals that are either hospitalized or at ICU become removed with probability μ_{H} or μ_{ICU}, respectively. We initialize the model in the city of Boston by selecting an attack rate on the 17th of March of 1.5% (a sensitivity analysis of this quantity is provided in the SM).

Social distancing strategies. To simulate social distancing measures, we modify the synthetic population such that:

- School closures are simulated by removing all the schools from the system simultaneously.
- Partial "stay at home" assumes that all places are open except restaurants, nightlife and cultural places. Closures of these places are simulated by removing the interactions that occur in any place that falls into that category according to Foursquare's taxonomy of places. This is the situation after the first reopening.
- Full lock-down and confinement assume that schools and all non-essential workplaces are closed. Here we close all workplaces except essential ones and remove interactions that occur at them. Essential workplaces are: Hospitals, Salons, Barbershops, Grocery Stores, Dispensaries, Supermarkets, Pet Stores, Pharmacies, Urgent Care Centers, Dry Cleaners, Drugstores, Maternity Clinics, Medical Supplies and Gas Stations.

The connectivity distributions for each of the scenarios simulated as well as other statistics related

to the effects of the lock-down are shown in the SM.

Data availability

The data that support the findings of this study is available from Cuebiq through their Data for Good program, but restrictions apply to the availability of these data, which were used under licenses for the current study, and so are not publicly available. Aggregated data used in the models are however available from the authors upon reasonable request and permission of Cuebiq. Other data used comes from the American Community Survey (5-year) from the Census, which is publicly available.

Code availability

Custom code that supports the findings of this study is available from the corresponding author upon request.

References

1. World Health Organization, "Novel Coronavirus – China" (2020). URL `https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/`.

2. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases* **20**, 533–534 (2020).

3. Kraemer, M. U. *et al.* The effect of human mobility and control measures on the COVID-19

epidemic in China. *Science* **368**, 493–497 (2020).

4. Flaxman, S. *et al.* Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in European countries: technical description update. Preprint at arXiv: https://arxiv.org/abs/2004.11342 (2020).

5. Walker, P. G. T. *et al.* The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science* eabc0035 (2020).

6. Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of sars-cov-2 through the postpandemic period. *Science* **368**, 860–868 (2020).

7. Laura Di Domenico , Giulia Pullano, Chiara E. Sabbatini , Pierre-Yves Boëlle , Vittoria Colizza. Expected impact of lockdown in Île-de-France and possible exit strategies. Preprint at medRxiv: https://www.medrxiv.org/content/10.1101/2020.04.13.20063933v1 (2020).

8. Liu, Q.-H. *et al.* Measurability of the epidemic reproduction number in data-driven contact networks. *Proceedings of the National Academy of Sciences* **115**, 12680–12685 (2018).

9. U.S. Census Bureau. 2018 American Community Survey 5-Year Data (2019). URL `https://www.census.gov/programs-surveys/acs`.

10. Poletto, C., Meloni, S., Colizza, V., Moreno, Y. & Vespignani, A. Host mobility drives pathogen competition in spatially structured populations. *PLoS computational biology* **9** (2013).

11. Zhang, J. *et al.* Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. *The Lancet Infectious Diseases* **20**, 793–802 (2020).

12. Li, Q. *et al.* Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *New England Journal of Medicine* **382**, 1199–1207 (2020).

13. Verity, R. *et al.* Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases* **20**, 669–677 (2020).

14. American Hospital Directory and are based on public records obtained from the US Centers for Medicare and Medicaid Services. (2020). URL {https://www.ahd.com/data_sources.html}.

15. The White House. 15 days to slow the spread. (2020). URL {https://www.whitehouse.gov/wp-content/uploads/2020/03/0316.20_coronavirus-guidance_8.5x11_315PM.pdf}.

16. CNN. These states have implemented stay-at-home orders. Here's what that means for you (2020). URL {https://www.cnn.com/2020/03/23/us/coronavirus-which-states-stay-at-home-order-trnd/index.html}.

17. IMF Blog. Global Uncertainty Related to Coronavirus at Record High (2020). URL {https://blogs.imf.org/2020/04/04/global-uncertainty-related-to-coronavirus-at-record-high/}.

18. Gottlieb, S., Rivers, C. & McClellan, M. B. National coronavirus response: A road map to reopening (2020). URL https://www.aei.org/research-products/report/national-coronavirus-response-a-road-map-to-reopening/.

19. Zhang, J. *et al.* Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* eaba8001 (2020).

20. Bakker, Michiel and Berke, Alex and Groh, Matt and Pentland, Alex and Moro, Esteban. Effect of social distancing measures in the New York City metropolitan area (2020). URL http://curveflattening.media.mit.edu/posts/social-distancing-new-york-city/.

21. Leung, K., Wu, J. T., Liu, D. & Leung, G. M. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *The Lancet* **395**, 1382–1393 (2020).

22. Kucharski, A. J. *et al.* Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *Lancet Infect. Dis.* (2020).

23. Hellewell, J. *et al.* Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health* **8**, e488–e496 (2020).

24. Ferretti, L. *et al.* Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, eabb6936 (2020).

25. Mossong, J. *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine* **5** (2008).

26. Ajelli, M. & Litvinova, M. Estimating contact patterns relevant to the spread of infectious diseases in russia. *Journal of Theoretical Biology* **419**, 1–7 (2017).

27. Litvinova, M., Liu, Q.-H., Kulikov, E. S. & Ajelli, M. Reactive school closure weakens the network of social interactions and reduces the spread of influenza. *Proc Natl Acad Sci USA* **116**, 13174–13181 (2019).

28. Mistry, D. *et al.* Inferring high-resolution human mixing patterns for disease modeling. Preprint at arXiv: https://arxiv.org/abs/2003.01214 (2020).

29. Béraud, G. *et al.* The french connection: the first large population-based contact survey in france relevant for the spread of infectious diseases. *PloS one* **10** (2015).

30. Zhang, J. *et al.* patterns of human social contact and contact with animals in shanghai, china. *Scientific reports* **9**, 1–11 (2019).

Acknowledgements NED, IML, MEH, APP and AV acknowledge the support of NIH/NIAID AI139761. MC and AV acknowledge support from Google Cloud Healthcare and Life Sciences Solutions via the GCP research credits program. MEH acknowledges support from the NIH/NIGMS U54 GM111274. MC, APP, AV, acknowledge the support of the McGovern Foundation. EM acknowledges partial support by MINECO (FIS2016-78904-C3-3-P). YM acknowledges partial support from the Government of Aragon and FEDER funds, Spain through grant E36-17R (FENOL), and by MINECO and FEDER funds (FIS2017-87519-P).

AA and YM acknowledge support from Intesa Sanpaolo Innovation Center. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions AA, DMC, MA, AV, EM and YM designed research; AA performed research with contributions from DMC; AA, DMC, MA, AV, EM, and YM analyzed the results. AV and YM wrote the first draft of the manuscript, and all other authors discussed results and edited the manuscript. All authors approved the final version.

Competing Interests MEH reports grants from the National Institute of General Medical Sciences during the conduct of the study; AV reports grants and personal fees from Metabiota, Inc., outside of the submitted work; MC and APyP report grants from Metabiota, Inc., outside of the submitted work. The authors declare no other relationships or activities that could appear to have influenced the submitted work.

Correspondence Correspondence and requests for materials should be addressed to AV (alexves@gmail.com), EM (esteban.moroegido@gmail.com) and YM (yamir.moreno@gmail.com).

Figure 1 **Model components.** Panel **a** is a schematic illustration of the weighted multilayer synthetic population built from mobility data in the metropolitan area of Boston. The agent-based system is made up by around 64000 adults and 21000 children, whose geographical distributions are shown in panel **b**. Nodes are connected by more than 5 million weighted edges. Community layers (that include workplaces), are further classified into categories according to Foursquare's taxonomy of places. Panel **c** displays the compartmental model used to describe the natural history of the disease as well as the transition rates between the different states. Specifically, we consider Susceptible (S), Latent asymptomatic (L_A), Latent symptomatic (L_S), Pre-symptomatic (P_S), Infectious asymptomatic (I_A), Infectious symptomatic (I_S), Hospitalised (H), Hospitalized in intensive care (ICU) and Recovered (R) individuals. More details of the model and the transitions between compartments are provided in Methods and the SM.

Figure 2 **Impact of COVID-19 under different scenarios.** Evolution of the number of new cases (**a, d, g**), the outbreak size (**b, e, h**) and the effective reproductive number (**c, f, i**) as a function of time in each situation studied. Results of the SARS-CoV-2 transmission dynamics are shown for the unmitigated scenario (top panels **a-c**), and the two social distancing interventions considered, LIFT (**d-f**) and LET scenarios (**g-h**). In both cases, we considered the closure of schools and non-essential places for 8 weeks. This is the strictest lock-down period, which is followed by a partial lifting of the stay-at-home policy whose duration is set to 4 weeks. During the partial lifting, all places in the community

Scenario			Hospitalization	ICU	Individuals traced
Unmitigated			4.57 (4.10-5.03)	2.56 (2.21-2.91)	-
LIFT			3.22 (2.80-3.67)	1.87 (1.55-2.20)	-
LET	Detection 30%	No Tracing	2.70 (2.29-3.12)	1.58 (1.27-1.88)	-
		Tracing 20%	0.86 (0.65-1.10)	0.55 (0.39-0.72)	0.52 (0.36-0.69)
		Tracing 60%	0.35 (0.21-0.50)	0.22 (0.12-0.34)	0.17 (0.08-0.27)
LET	Detection 50%	No Tracing	2.35 (1.97-2.75)	1.39 (1.11-1.68)	-
		Tracing 20%	0.44 (0.28-0.62)	0.28 (0.16-0.42)	0.39 (0.23-0.55)
		Tracing 40%	0.29 (0.18-0.43)	0.15 (0.08-0.26)	0.14 (0.05-0.23)

Table 1: Mean and 95% C.I. of the number of normal hospitalizations, ICU hospitalizations and symptomatic individuals identified/traced (when applicable) at the peak of the epidemic per 1,000 people. The estimated availability of ICU beds is 0.21 beds per 1,000 people.

layer are open except mass-gathering locations (restaurants, theaters, etc, see SM). Finally, a full reopening takes place after the period of partial lifting ends (relevant events are marked with the vertical lines). Panels **d-f** consider that no other measures are adopted concurrently to the lifting of the restrictions, whereas the results in panels **g-i** have been obtained when the reopening is accompanied by an active policy consisting of testing the symptomatic individuals, home isolating them, and quarantining their household and the households of a fraction of their contacts, as indicated in the legend of the bottom panels. Note that the vertical scales of panels **a**, **d**, and **g** are not the same and that both the number of new cases and total cases are per 1,000 inhabitants. In all panels the solid line represents the average over 10,000 simulations and the shaded region the 95% C.I.

Figure 3 **Impact on the Boston health care system.** Estimated number of individuals per 1,000 inhabitants that would need hospitalization (H), and intensive care (ICU) for each of the three scenarios considered in Figure 2. Panel **a** corresponds to the unmitigated situation, whereas results for the LIFT and LET strategies are shown in panels **b** and **c**, respectively. The horizontal dotted-dashed lines represent the ICU basal capacity of the Boston health care system. The dotted line in panel **c** indicates 30% of the ICU basal capacity.

Figure 4 **Affordability of the best way-out scenario.** LET strategy with 50% detection and 40% tracing. (**a**) Fraction of the population that needs to be put under quarantine as a function of time and percentage of contact tracing. (**b**) Health status of the individuals that

are quarantined for a contact tracing level of 40%. Note that only symptomatic individuals are tested, which implies that a large fraction of the quarantined population is of unknown status. This fraction of individuals quarantined with unknown health condition could be reduced if the capacity to do more tests increases. As it is shown, the pandemic might span over several months depending on the level of contact tracing. (**c**) Number of individuals whose contacts are traced each day per 1,000 persons. Relevant intervention actions are signaled by the vertical dashed lines in all panels.

Supplementary Material: Modeling the impact of testing, contact tracing and household quarantine on second waves of COVID-19

Alberto Aleta, David Martín-Corral, Ana Pastore y Piontti, Marco Ajelli, Maria Litvinova, Matteo Chinazzi, Natalie E. Dean, M. Elizabeth Halloran, Ira M. Longini, Jr., Stefano Merler, Alex Pentland, Alessandro Vespignani, Esteban Moro & Yamir Moreno

July 16, 2020

Contents

1 Mobility data

The mobility data was obtained from Cuebiq, a location intelligence and measurement company. The dataset consists of anonymized records of GPS locations from users that opted-in to share the data anonymously in the Boston metropolitan area over a period of 6 months, from October 2016 to March 2017. Data was shared in 2017 under a strict contract with Cuebiq through their Data for Good program where they provide access to de-identified and privacy-enhanced mobility data for academic research and humanitarian initiatives only. All researchers were contractually obligated to not share data further or to attempt to de-identify data. Mobility data is derived from users who opted in to share their data anonymously through a General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA) compliant framework.

From the data we extracted the "stays", as the places where anonymous users stayed (stopped) for at least 5 minutes. Some of the stays happen within places (Points of Interest). We use a dataset of 86k Points of Interest in the Boston metropolitan area collected using the Foursquare API. Stays are then aggregated at place level. Finally we estimate the home Census Block Group of the anonymous users as that in which they are more likely located during nighttime. This results in a dataset of the places people stayed including the points of interest that anonymous users visited and the most likely census block group of where the device owner lives.

2 Network structure

2.1 Agents

Our population consists of two different sub-populations, adults and children. Adults are sampled from anonymous individuals in the mobility data collected by Cuebiq, each adult is associated with a home location assigned to a US Census block group which is provided by our location data provider. With this data we designed a population building pipeline that consists of three steps.

- First step, we build synthetically the number of households, their size and the presence of children based on our adult population and the US Census [1] tables B11016 (Household Type by Household Size) and B11003 (Family Type by Presence and Age of Own Children)
- Second step, we assign adults to households and in case of presence of children we generate them up to reach the size of the household assigned in the first step.
- And final step, we assign ages to nodes using table B01001 (Ages by Sex) of age distribution within the Census Block Group.

This process generates our synthetic population consisting of 85k agents (2% of the population in the Boston Metropolitan Area), 64k (75%) of them are adults and 21k (25%) are children. Age groups are distributed as follows: 6,027 (7%) agents for the age group between zero and five years old, 16,250 (18.9%) agents for the age group between six and eighteen years old, 36,207 (42,2%) agents for the age group between nineteen and fifty years old, 13,176 (15,4%) agents for the age group between fifty one and sixty five years old, and final group, 13,945 (16,2%) agents for the age group between sixty six and older. All of agents together form 43,167 households distributed as follows: 23,293 (53.9%) households with only one agent, 7,886 (18.2%) with two agents, 4,959 (11.4%) with three agents, 4,486 (10.4%) with four agents, 1,784 (4.1%) with five agents, 514 (1.2%) with six agents, and finally, 245 (0.5%) with seven agents. In Figure 1 we can see the comparison of our synthetic population against census data.

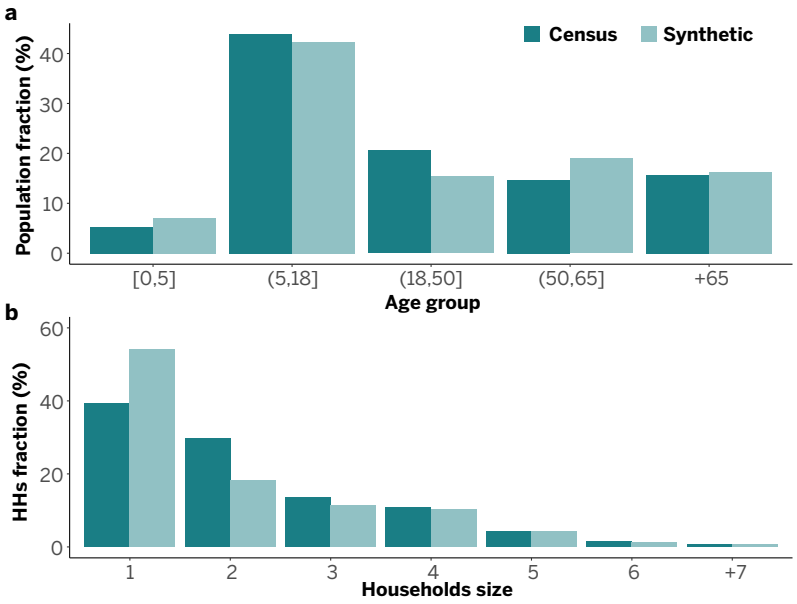

Figure 1: Age groups and households demographics compared against US Census data. (a) Age groups distribution. (b) Households size distribution.

Our sample from devices is very representative of the population in the Boston area. As we can see in Figure 2, population and number of anonymous devices detected in the real data by census area are highly correlated: $\rho = 0.8$ (Pearson correlation) with a CI between 0.77 and 0.82 for county subdivisions.

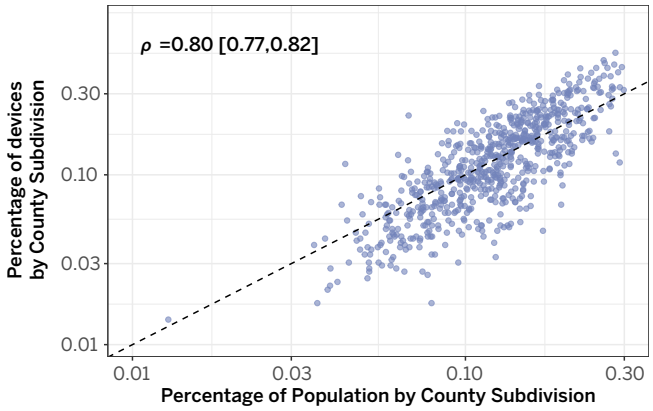

Figure 2: The correlation between the population for each county subdivision and the number of devices in our dataset.

2.2 Contacts

In the unmitigated scenario, our network has a total number of 5,029,888 unique daily contacts, 3,924,694 (78%) of them in the community layer obtained using the mobility data, 160,748 (3%) and 944,446 (19%) are synthetically built for the household and school layers, respectively. The community layer is based on estimation of co-presence of two devices in Points of Interest visited by the anonymous users (see Methods in the main paper). Points of interest (POIs) are categorized using the Foursquare taxonomy of places which has ten main categories: Art & Entertainment (4.4%), Colleges & Universities (4.8%), Food (16.7%), Nightlife Spots (3.9%), Outdoors & Recreation (10.6%), Professional & Other Places (23.7%), Shops & Services (29.1%) and Travel & Transport (6.4%), and 638 subcategories. See [2] for a complete list of them.

2.3 Social distancing policies

We simulated two different scenarios for social distancing policies. This produces three contact networks: i) *baseline*, ii) *medium closure*, and iii) *non-essential closure*, as we see it in more detail in Table 1. Schools are closed in the medium and non-essential closure, but both policies differ in the number of places kept open in the community layer. In Table 2 we can see the distribution of POIs, by main Foursquare category, that remain open during each social distancing policy. In the baseline scenario, we keep all the categories and thus the average number of contacts in the community layer is 63 (median 47, [15-150] 90% confidence interval), with few anonymous individuals having a large number of contacts (that could eventually lead to super-spreading events). In the medium closure scenario, POIs in the Art & Entertainment, Restaurants and Nightlife categories are closed; this drastically reduces the average number of contacts to 27 (median 15, [0-92] 90%CI). Lastly, when all non-essential places are closed, we only keep open the following subcategories: Hospital, Salon / Barbershop, Grocery Store, Dispensary, Supermarket, Pet Store, Pharmacy, Urgent Care Center, Dry Cleaner, Drugstore, Maternity Clinic, Medical Supply, and Gas Station. In this situation, the average number of contacts is reduced to 6 (median 0, [0-29] 90%CI). The distribution for the number of contacts in the community layer in these three scenarios is shown in Figure 3.

	Baseline		Medium closure			Non-essential closure		
Layers	Contacts	%.	Contacts	%	% Diff.	Contacts	%	% Diff.
Community	3,924,694	78	1,378,054	27.4	-72.6	357,144	7.1	-92.9
Households	160,748	3.2	160,748	3.2	0	160,748	3.2	0
Schools	944,446	18.8	0	0	-100	0	0	-100
Total	5,029,888	100	1,538,802	30.6	-69.4	517,892	10.3	-89.7

Table 1: Number of daily contacts by layer and social distancing policy.

	Baseline		Medium closure.			Non-essential closure.		
POIs categories	Open	%.	Open	%	% Diff.	Open	%	% Diff.
Arts and Entertainment	3,692	4.44	0	0	-100	0	0	-100
Colleges and Universities	4,016	4.83	4,016	4.83	0	0	0	-100
Restaurants	13,860	16.7	0	0	-100	0	0	-100
Nightlife Spots	3,288	3.95	0	0	-100	0	0	-100
Outdoors and Recreation	8,840	10.64	8,840	10.64	0	229	0.27	-97.4
Professional and Other Places	19,692	23.71	19,692	23.71	0	415	0.5	-97.8
Shops and Services	24,310	29.27	24,310	29.27	0	5,139	6.19	-78.8
Travel and Transport	5,370	6.46	5,370	6.46	0	0	0	-100
Total	83,608	100	62,228	74.91	-25.6	5,783	6.96	-93.1

Table 2: Number of POIs open in the Community layer by the different social distancing measures and full non-essential closure. Percentages are calculated with respect to the total number of POIs in the baseline.

2.4 Calibration of intra-layer links

As described in the main text, we define ω_{ij} as the weight associated to the link between node i and j.

In the community+workplace layer, we estimate the mean number of daily effective contacts (η_C) by

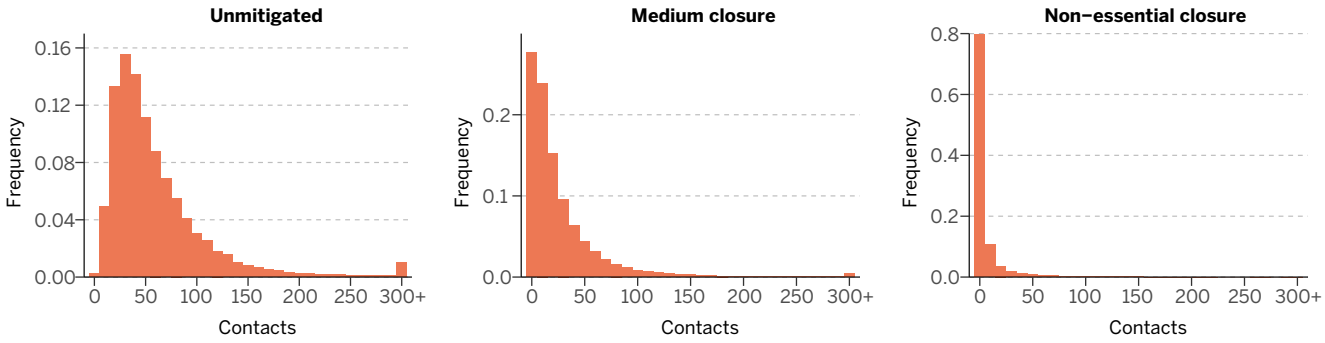

Figure 3: Degree distribution in the community layer under normal conditions, soft social distancing measures and full non-essential closure.

using $\omega_{C_{ij}}$, which is based on the co-presence probability estimation. η_C can thus be calculated as

$$\begin{aligned} \eta_C &= 1/n \sum_{i\in\{1,\ldots,n\}} \sum_{j|j\neq i\wedge j\in\{1,\ldots,n\}} \omega_{C_{ij}} \\ &= 2.5 \end{aligned}$$

By construction, the weights for the household layer were assigned as $\omega_{H_{ij}} = 1/(h-1)$, where h is the number of household members, so that $\eta_H = 1$. Analogously, by construction, for schools we have $\eta_S = 1$.

It is important to note that $\eta_{C,H,S}$ refer to the mean number of daily effective contacts in the synthetic (non-calibrated) network. Based on the analysis of contact survey data from 9 countries [3, 4, 5, 6, 7], the estimated number of daily effective contacts by social setting is 10.86 in community+workplace, 4.11 in household, and 11.41 in school.

To calibrate the weights of intra-layer links ($\hat{\omega}_l$), we associate to each layer a single rescaling factor w_l such that the mean number of daily effective contacts in that layer matches mean number of daily effective contacts in the corresponding social setting. Therefore, the calibrated mean number of daily effective contacts in the community+workplace layer $\hat{\eta}_C$ is

$$\begin{aligned} \hat{\eta}_C &= 1/n \sum_{i\in\{1,\ldots,n\}} \sum_{j|j\neq i\wedge j\in\{1,\ldots,n\}} \omega_{C_{ij}}\ w_{C+W} \\ &= 1/n \sum_{i\in\{1,\ldots,n\}} \sum_{j|j\neq i\wedge j\in\{1,\ldots,n\}} \hat{\omega}_{C_{ij}} \\ &= 2.50 \times 10.86/2.50 \\ &= 10.86 \end{aligned}$$

where $w_{C+W} = 10.86/2.5$. Analogously for household and school layers we obtain $w_H = 4.11$ and $w_S = 11.41$.

3 SARS-CoV-2 transmission model

The values of all the disease parameters used for simulating the transmission dynamics are given in table 3. Figure 4 shows the numerical distributions of these parameters as resulting from simulations of the model.

Parameters	Description	Age group	Value	Ref.
r	relative infectiousness of asymptomatic individuals	-	50%	†
ϵ^{-1}	latent period	-	3 days	[8]
ϵ'^{-1}	latent period	-	5 days	[8]
p	proportion of asymptomatic	-	25%	[9]
γ^{-1}	pre-symptomatic period	-	2 days	[8]
μ^{-1}	time to removed/home stay	-	2.5 days	$*$
α	symptomatic case hospitalization ratio (%)	0-4 5-17 18-49 50-64 65+	0.0 0.025 2.672 9.334 15.465	[10]
χ	ICU % among hospitalized	0-4 5-17 18-49 50-64 65+	5.0 5.0 5.38 17.10 44.71	[11]
δ^{-1}	days from home stay to hospital admission	-	2	[12]
μ_H^{-1}	days in hospital	-	8	[10]
μ_{ICU}^{-1}	days in ICU	-	13	[10]
k	proportion of presymptomatic transmission	-	15%	[13]
R_0	basic reproduction number	-	2.5	†
β	transmission for symptomatic and asymptomatic individuals	-	$\frac{R_0\mu}{pr+(1-p)/(1-k)}$	
β_S	transmission for pre-symptomatic individuals	-	$\frac{\beta\gamma k}{\mu(1-k)}$	

Table 3: Baseline set of parameters. †: assumed (see Section 6 for sensitivity analysis);$*$: calibrated to the generation time T_g.

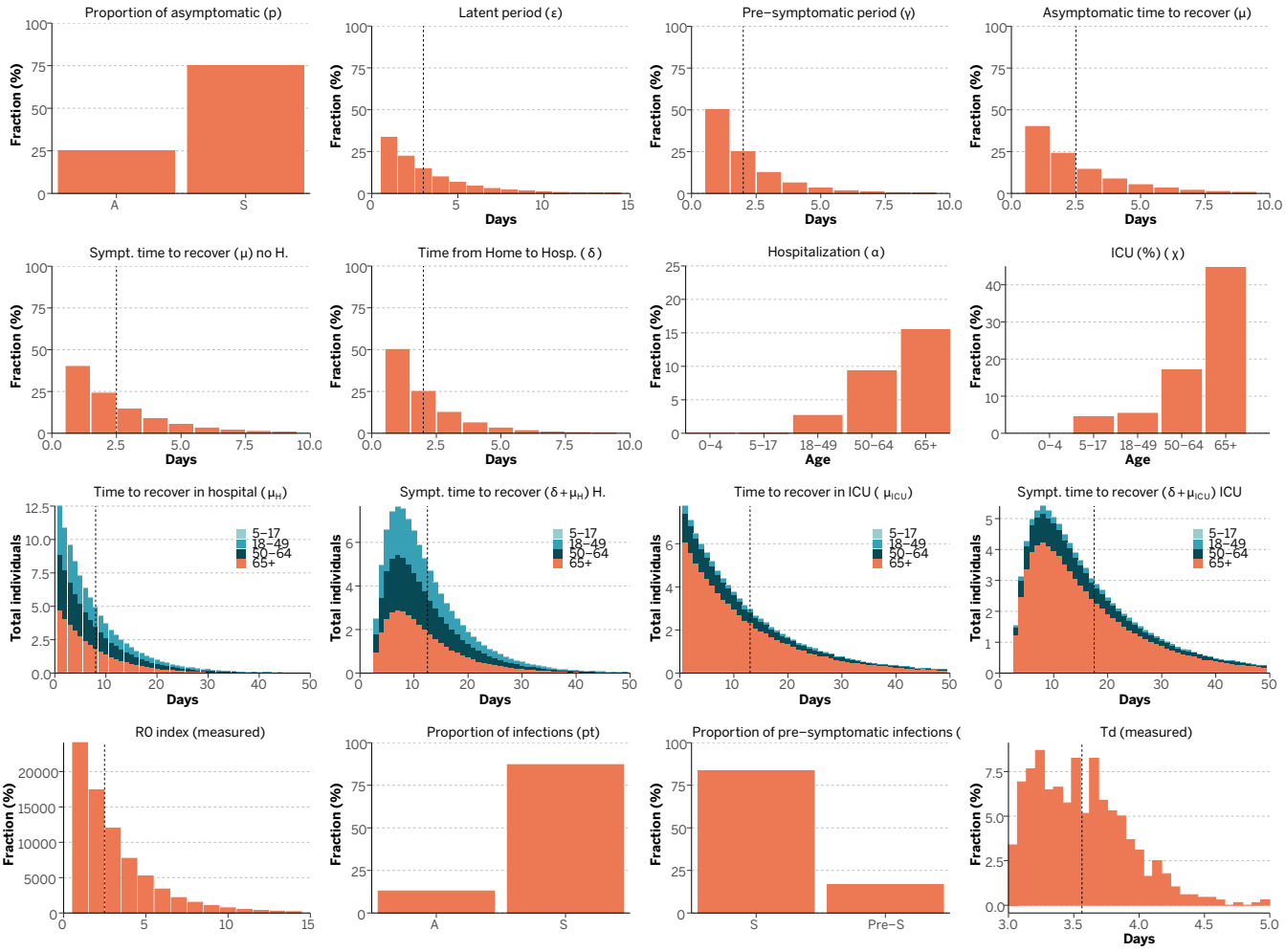

Figure 4: Model's parameters and distributions summary.

3.1 Estimation of the effective reproduction number

We assume that the daily number of new infectious individuals $C(t)$ at time t can be approximated as

$$C(t) \approx Pois\left(R(t)\sum_{s=1}^{t}\phi(s)C(t-s)\right), \tag{1}$$

where ϕ is the generation time distribution and $R(t)$ is the effective reproduction number at time t. The generation time is directly measured in the simulations and fitted to a log-normal distribution with mean 6.6 days.

In each single run, the likelihood L of the observed time series of cases from day 1 to T is given by

$$L = \prod_{t=1}^{T} P\left(C(t), R(t)\sum_{s=1}^{t}\phi(s)C(t-s)\right), \tag{2}$$

where $P(k,\lambda)$ is the probability mass function of a Poisson distribution. The posterior distribution of $R(t)$ is then explored using MCMC sampling [14, 15].

After the implementation of social distancing policies, a certain amount of nodes will be completely disconnected from the system (individuals living alone who lose all their links in the community layer).

Denoting this quantity as $Z(t)$, the actual number of infectious individuals that can produce new cases is $C(t) - Z(t)$. We modify expression (2) accordingly so that

$$L = \prod_{t=1}^{T} P\left(C(t), R(t) \sum_{s=1}^{t} \phi(s)\left[C(t-s) - Z(t-s)\right]\right) . \tag{3}$$

Note that $Z(t-s)$ will always be 0 in the unmitigated scenario and larger than zero only in a few steps after the total closure.

In Figure 5 we compare the effective reproduction number estimated using equations (2) and (3). Right after the closure the estimation obtained using (2) shows a small valley followed by an increment of $R(t)$. This is produced by nodes that lose all their connections and thus cannot spread the disease at all. Once they recover, since the disease will be still spreading through a small fraction of the system, $R(t)$ seems to increase. Using equation (3) corrects this problem and $R(t)$ shows a continuously decreasing trend until the partial reopening.

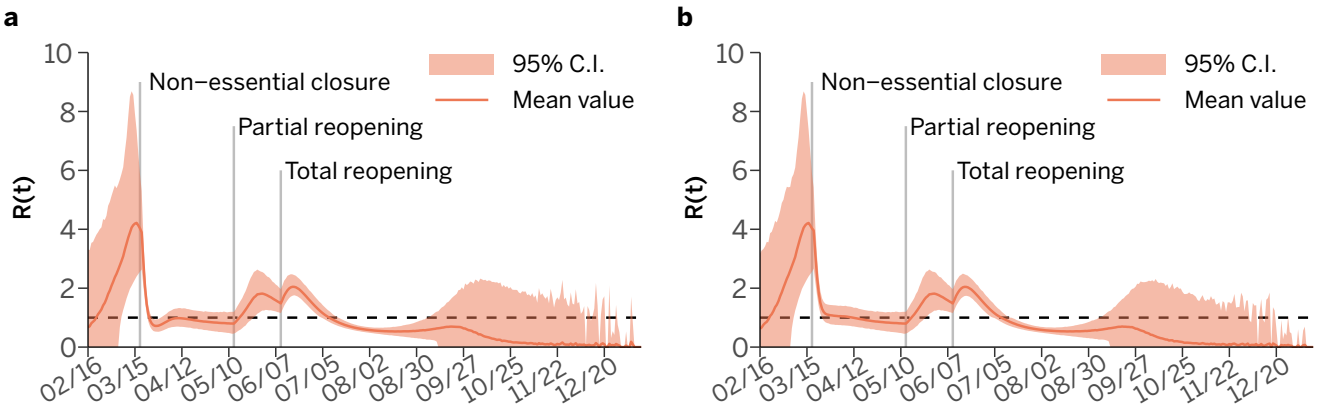

Figure 5: Estimation of the effective reproduction number using equation (2) (panel a) or (3) (panel b). Solid lines represent the mean and the shaded region the 95% C.I.

3.2 Strategy implementation

On March 17th all links in the school layer are removed, as well as all links in the community+workplace layer associated to non-essential places. After 8 weeks, all these links in the community+workplace layer are restored, except the ones belonging to restaurants, nightlife, museums, etc (see Section 2.3 above). Lastly, after 4 weeks the rest of the links of the community+workplace layer are added to the system (while schools remain closed). We align the computational step with the real dates so that on March 17th the attack rate in the population is 1.5% in the Figures shown in the main text. In section 6.4 we explore the effect of choosing different attack rates for initializing our simulations.

Then, in the LET scenario, quarantines are applied starting on the same day of the partial reopening (i.e., 8 weeks after the initial closure). We have explored 3 types of quarantines:

i) A fraction qH of all the symptomatic individuals is identified after, on average, qI days of the onset of symptoms. After identification, they are completely isolated in special locations (i.e., outside of their households). A fraction qT of their contacts are traced and quarantined without testing in special locations where they are also completely isolated.

ii) A fraction qH of all the symptomatic individuals is identified after, on average, qI days of the onset of symptoms. After identification, they are isolated at their homes. A fraction qT of their contacts are traced and quarantined without testing at their homes.

iii) A fraction qH of all the symptomatic individuals is identified after, on average, qI days of the onset of symptoms. After identification, they are isolated in their households and the rest of the household is set in quarantine (i.e., they cannot contact anyone outside their household). A fraction qT of their contacts are traced and their households are quarantined without testing.

After 14 days, all individuals isolated or under quarantine are allowed to interact again, regardless of their health status.

In Figure 6, we explore the ICU usage at the peak as a function of the amount of detection (qH) and the amount of tracing (qT) when symptomatic individuals are identified, on average, in 1, 2 or 3 days (qI). The dashed line represents the boundary of ICU saturation. For any pair of qT, qH values on the left of the dashed line, the burden of ICU usage will surpass the current capacity. Note that the maximum capacity is the same in all cases, but different values of detection and tracing produce a different maximum usage. These results indicate that investing in contact tracing and symptomatic detection is a better strategy than simply multiplying the number of available ICU beds. Indeed, doubling the ICU capacity would only reduce the amount of symptomatic detection necessary by 10%.

We further explore the value of this threshold in Figure 7. We observe that increasing the time of symptomatic detection and isolation increases both the number of symptomatic individuals that have to be detected and the amount of tracing performed. We also note that even though the quarantine of whole households seems to be the best option, isolating and quarantining individuals outside of their households is also a viable strategy. However, as we show next, even though this measure has a lower burden on the families and their individuals economy, it will increase the costs for the overall society.

In Figure 8, we show the maximum number of individuals whose contacts are traced in a single day. Furthermore, note that identifying those individuals as symptomatic also requires testing. Hence, this measure will be proportional to the amount of tests that have to be performed daily (the total number will depend on the positive test rate). Note that even though isolating and quarantining individuals in special locations has a similar effect as doing so in their households from the ICU usage point of view, the number of tracers needed in the former strategy is larger than in the latter. This implies that the amount of testing will also be higher. Added to the cost of creating and managing such special facilities, these results indicate that quarantining whole households is the best strategy overall. However, whenever possible, the possibility of isolating and quarantining single individuals in special locations should be offered to the families who would suffer the most due to the complete quarantining of their household members.

4 Distribution of infections

In Figure 9 we show how the infections that are produced for each strategy are distributed across layers. In the unmitigated scenario the disease spreads initially through the community+workplace layer. By the time the incidence peaks (around 04/12) infections are transmitted roughly equally through each layer.

In the LIFT scenario the fraction of infections that take place in households is larger than in the unmitigated one during the first closure. Once places in the community+workplace layer are reopened, the disease starts to spread through that layer again.

Lastly, in the LET strategy, in comparison to the LIFT scenario, we observe a larger fraction of infections within households in the period between partial and total reopening, which is to be expected since symptomatic individuals are force to stay at home. This suggests that providing special facilities to isolate those individuals will be beneficial for their families.

5 Household attack rate

In Figure 10, we explore the secondary attack rate in households. The attack rate in a household, with at least one infected individual, is defined as the fraction of household members (excluding the first infected member) that have suffered the disease. While in the unmitigated scenario this attack rate quickly rises to 80%, in the two other scenarios the situation is slightly different. The initial confinement of individuals at their household initially increases the attack rate, but since the overall burden of the disease is lower by the end of the epidemic, the final attack rate is also lower than in the unmitigated scenario. This also explains why the household attack rate in the LET scenario is small (compared to the other two situations) even though people are forced to stay at home.

6 Sensitivity analysis

6.1 Model parameters

We conducted a sensitive analysis to study the effects of different percentages of asymptomatic cases, higher transmissibility and higher pre-symptomatic transmission. We consider a scenario where all parameters are kept the same as the unmitigated scenario discussed in the main manuscript with the exception that we consider 50% of cases are asymptomatic (henceforth, set A). We also considered the effect of increasing the pre-symptomatic transmission of the virus, setting $k = 0.50$ (henceforth, set B) in line with recent observations [16]. Lastly, we test a scenario of higher transmission, where $R_0 = 3.0$ and the time to hospitalization is increased δ^{-1} to 4 days (henceforth, set C).

Figure 11 shows the results for the sensitivity analysis when we consider the unmitigated scenario (a-c) and the LIFT scenario (d-e) for the first two sets of parameters. As we can see, changing these distributions does not affect the total number of infections. Instead, the main change is related to the velocity of propagation. Increasing the pre-symptomatic transmission slightly accelerates the disease, even though the final outcome is fairly similar.

Figure 12 shows the results for the LET scenario where we consider 50% detection of symptomatic cases and different levels of contact tracing: no tracing, 20%, and 40%. The increase of asymptomatic individuals (set A) reduces the effect of contact tracing. Increasing the amount of pre-symptomatic transmission also increases the incidence, something that is expected since individuals will be able to spread more the disease before showing symptoms and being isolated. In any case, in figure 13, we can see that the LET strategy is still feasible with these parameters, signaling that it is quite robust.

We report a similar analysis in figures 14, 15 and 16 for the set of parameters with larger R_0 (set C). As expected, this change yields a higher number of infections. Nonetheless, the overall behavior holds and, even though tracing might need to be slightly increased, the LET strategy is still feasible.

Summarizing, the sensitivity analysis shows that the modeling results presented in the main text are robust to plausible ranges of parameter values for the key time-to-event intervals of COVID-19.

6.2 Symptomatic detection rate

In the main text we mainly explored the situation in which 50% of all symptomatic individuals can be detected and their contacts traced, on average, in two days. In Figures 17 and 18 we show the evolution of the system with a detection rate of 30% and 70%, respectively. Once again, our results and conclusions hold.

6.3 Contact tracing delay

In the main text we considered that symptomatic individuals could be identified, on average, on two days after the onset of symptoms. However, it was assumed that the contacts of the symptomatic individual could be traced and isolated within that same period of time. In this subsection we explore the effect of needing 1 or 2 days to perform the tracing and isolation of the contacts (figures 19 and 20 respectively). The effect on this delay is fairly small, which is to be expected due to the long incubation period of this disease.

6.4 Effect of different attack rates

In Figure 21 we explore the effect of choosing a different initial attack rate for the implementation of the social distancing policies. We observe that the sooner these policies are implemented (i.e., the disease has not spread through the population and thus the attack rate is low) the smaller the first peak of the disease will be. However, once the restrictions are removed, if there is a second wave the peak will be larger. Conversely, if the policies are implemented once a large fraction of the population has already suffered the disease, the second wave of the epidemic will be much milder.

7 School reopening

In the main scenarios studied in the paper, once schools are closed they remain so indefinitely. In this section, we explore the consequences of reopening schools on the 15th of September. In figure 22 we show the evolution of the epidemic for the LIFT and LET scenarios under such condition. As we can see, since in the LIFT scenario the peak of the epidemic occurs in July, the effect of reopening schools is negligible. A similar observation can be made in the LET scenario without tracing. However, for the LET scenarios with tracing, since the peak of the epidemic has not faded out by that date, reopening schools produces a small increment on the incidence. Nonetheless, in figure 23 we observe that such increment does not saturate the health care system. This signals that the increase in the number of infections is due to children getting infected, since this group is less likely to require hospitalization.

In figure 24 we explore the affordability of the LET scenario when schools are reopened. In the cases of no tracing and 20% tracing, since the peak of quarantined households is set before September, the reopening does not produce a large increase on the households that have to be set under quarantine. However, in the 40% tracing scenario since the peak is close to the date when schools would reopen, the population under quarantine could pass the 10%. It is important to stress that the role of children in the transmission of SARS-Cov-2 is not yet clear. Several papers indicates a differential susceptibility to the infection as well as a smaller probability of developing symptoms. In principle this might lead smaller role of school reopenings that should be assessed in the presence of firm data on differential forward transmission across different age brackets.

Lastly, in figure 25 we show the household attack rate and the fraction of infections that would take place in each layer. We observe that the household attack rate would remain stable and that the number of infections that would take place in schools will increase above 0.

8 Soft distancing measures and transmissibility reduction

In this section we explore the effect that soft social distancing measures, such as wearing masks, increasing the space between individuals in closed places, cleaning more often public spaces, etc. To do so, we reduce

the overall transmissibility of the disease by a certain coefficient since these measures are usually applied to all individuals and not only to the ones affected by the disease.

In figure 26, we show the effect that population-wide measures would have in the LIFT scenario if they are applied on the same date as the total reopening. We observe that even with a 60% reduction the ICU capacity will be surpassed. As such, these measures, on their own, will not be enough to contain the epidemic.

In figure 27, we apply this reduction to the LET scenario with 50% symptomatic detection but without any tracing. We can see that to keep the ICU usage below the threshold a large reduction of at least 60% is needed. Thus, even though these measures can help, they are not enough if only symptomatic individuals are isolated. On the other hand, in figure 28, we analyze the situation in which 20% of the contacts are traced and isolated (note that in the main text the main success scenario requires 40% of contact tracing). In this case the reduction of transmissibility has a dramatic effect, even with just a 20% reduction. Indeed, not only in all the cases analyzed the ICU needs are within the availability limit, but also the fraction of individuals and households that have to be quarantined is much smaller than in the no reduction case. As such, complementing the test and tracing strategy with soft social distancing measures can clearly benefit the containment of the epidemic, specially because less test and tracers will be necessary.

References

[1] U.S. Census Bureau. 2018 American Community Survey 5-Year Data (2019). URL `https://www.census.gov/programs-surveys/acs`.

[2] Foursquare venue categories. `https://developer.foursquare.com/docs/build-with-foursquare/categories/`. Accessed: 2020-04-25.

[3] Mistry, D. *et al.* Inferring high-resolution human mixing patterns for disease modeling. Preprint at arXiv: https://arxiv.org/abs/2003.01214 (2020).

[4] Mossong, J. *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine* **5** (2008).

[5] Béraud, G. *et al.* The French connection: the first large population-based contact survey in France relevant for the spread of infectious diseases. *PloS one* **10** (2015).

[6] Ajelli, M. & Litvinova, M. Estimating contact patterns relevant to the spread of infectious diseases in Russia. *Journal of Theoretical Biology* **419**, 1–7 (2017).

[7] Zhang, J. *et al.* Patterns of human social contact and contact with animals in Shanghai, China. *Scientific reports* **9**, 1–11 (2019).

[8] Backer, J. A., Klinkenberg, D. & Wallinga, J. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance* **25**, 2000062 (2020).

[9] Nishiura, H. *et al.* Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *International Journal of Infectious Diseases* **94**, 154–155 (2020).

[10] Verity, R. *et al.* Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases* **20**, 669–677 (2020).

[11] Ferguson, N. *et al.* Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand (2020). URL `https://doi.org/10.25561/77482`.

[12] Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860–868 (2020).

[13] Du, Z. *et al.* Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. *Emerging Infectious Diseases journal* **26** (2020).

[14] Ajelli, M. *et al.* Spatiotemporal dynamics of the Ebola epidemic in Guinea and implications for vaccination and disease elimination: a computational modeling analysis. *BMC Medicine* **14**, 1–10 (2016).

[15] Liu, Q.-H. *et al.* Measurability of the epidemic reproduction number in data-driven contact networks. *Proceedings of the National Academy of Sciences* **115**, 12680–12685 (2018).

[16] Casey, M. *et al.* Pre-symptomatic transmission of sars-cov-2 infection: a secondary analysis using published data. Preprint at medRxiv: https://www.medrxiv.org/content/10.1101/2020.05.08.20094870v2 (2020).

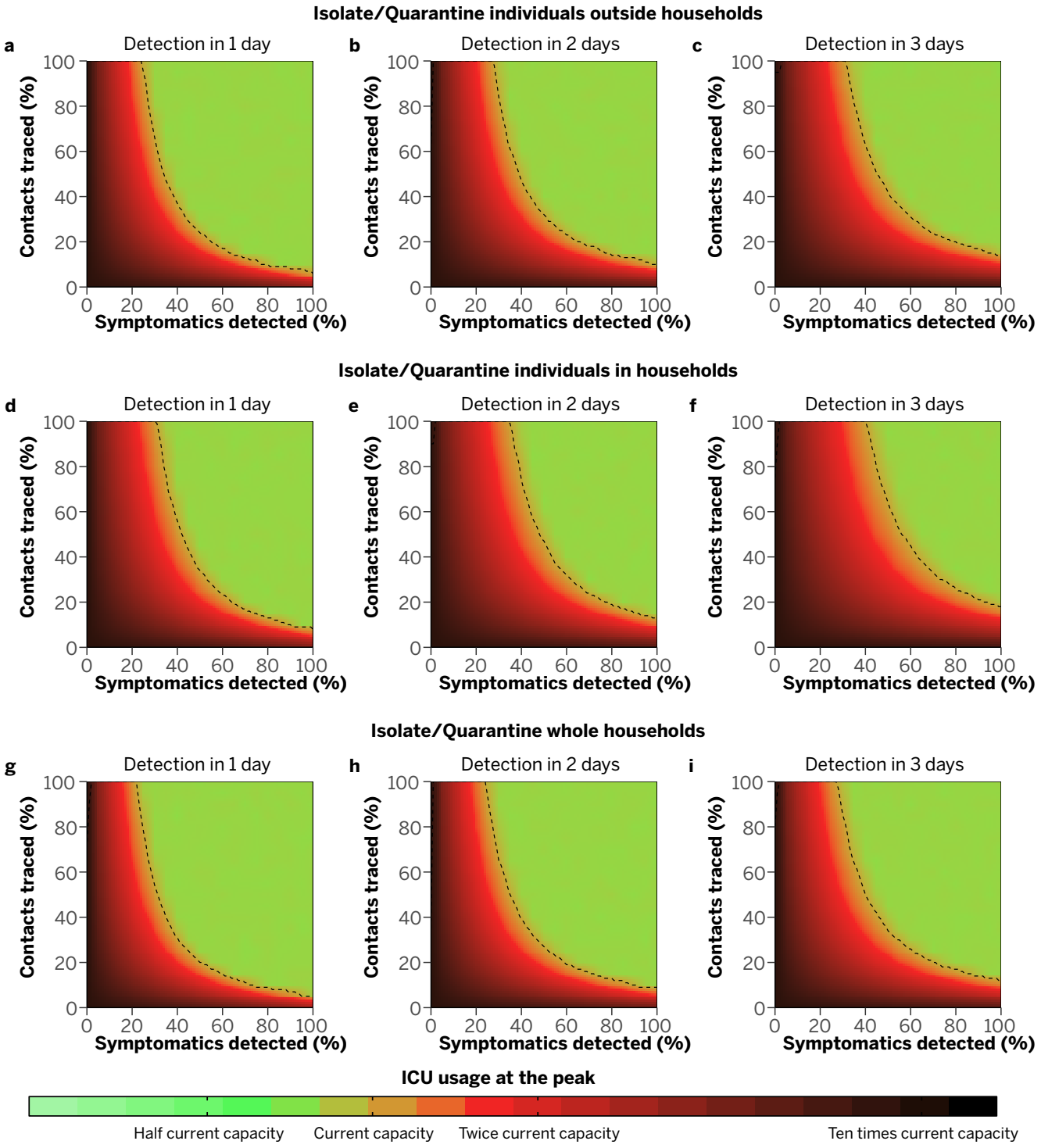

Figure 6: Maximum ICU occupation with different isolation/quarantine strategies. From left to right, each column shows the results when the symptomatic individuals are detected and isolated after 1, 2 or 3 days. From top to bottom, each row shows the results for strategies i), ii) and iii). The dashed line indicates when the ICU occupation excess the current availability.

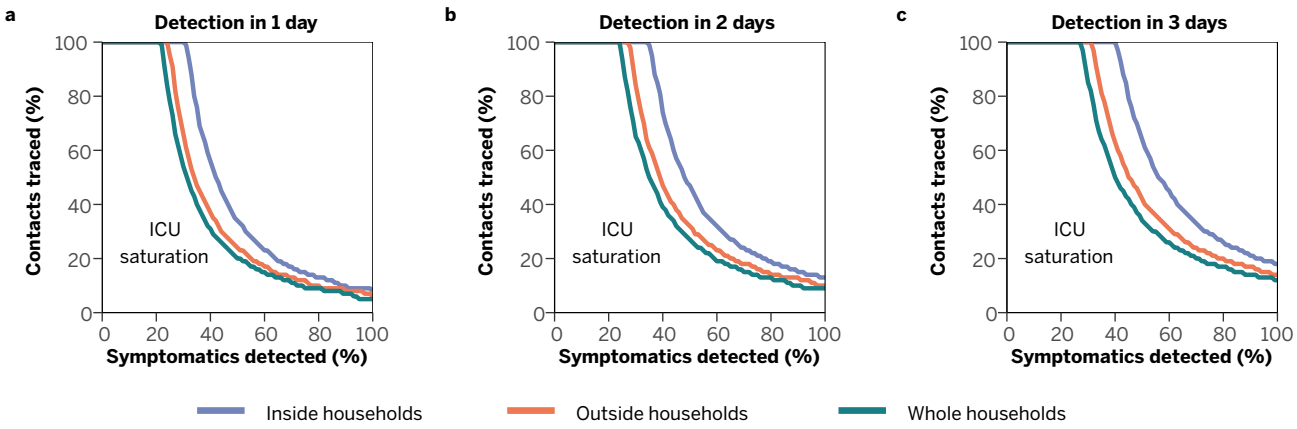

Figure 7: Threshold for ICU saturation - with the current estimated ICU availability - as a function of the fraction of symptomatic individuals detected and the fraction of contacts traced. From left to right, we show the results when the symptomatic individuals are detected and isolated after 1, 2 or 3 days.

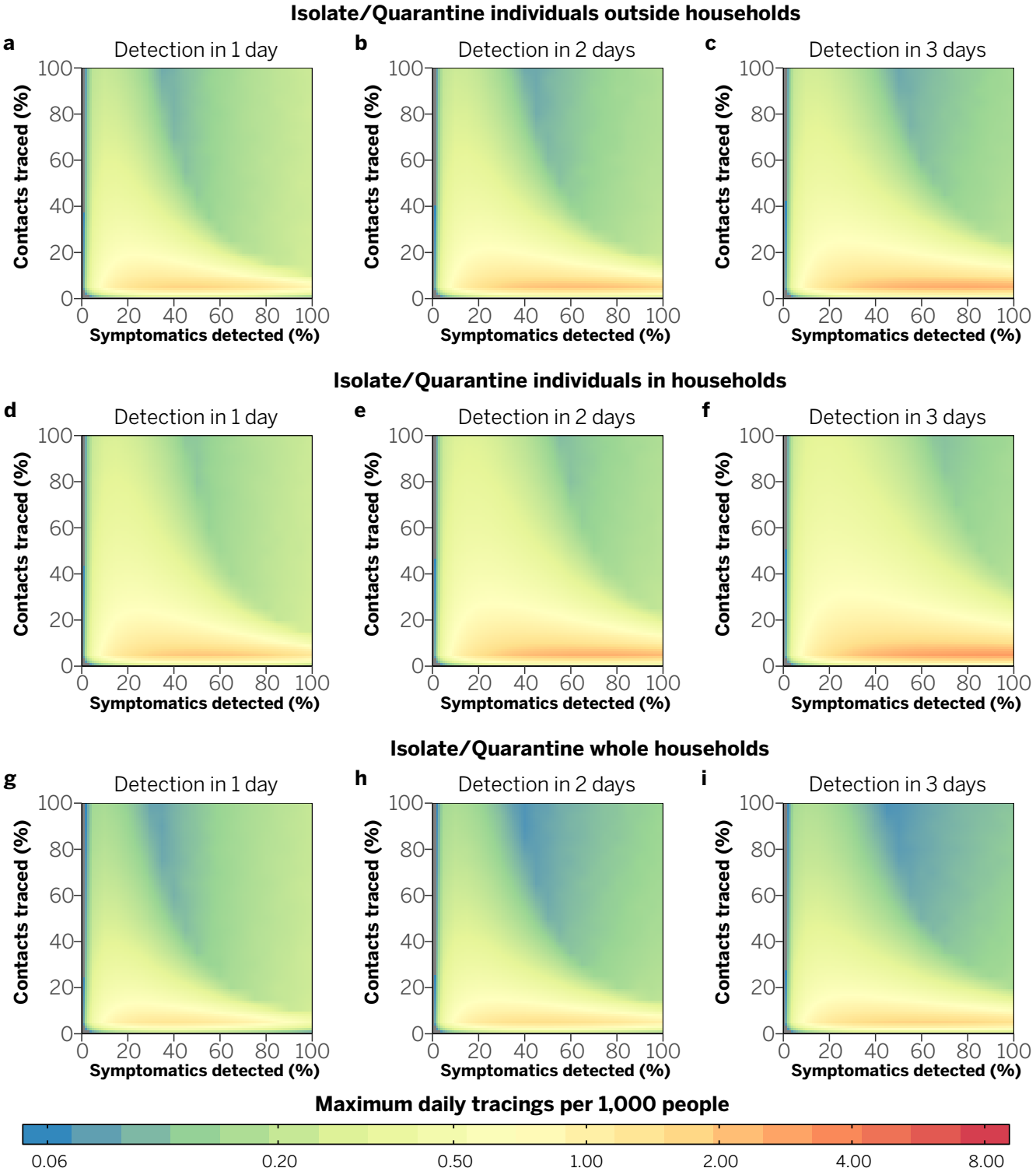

Figure 8: Maximum number of individuals whose contacts are traced in a single day per 1,000 people. From left to right, each column shows the results when the symptomatic individuals are detected and isolated after 1, 2 or 3 days. From top to bottom, each row shows the results for strategies i), ii) and iii).

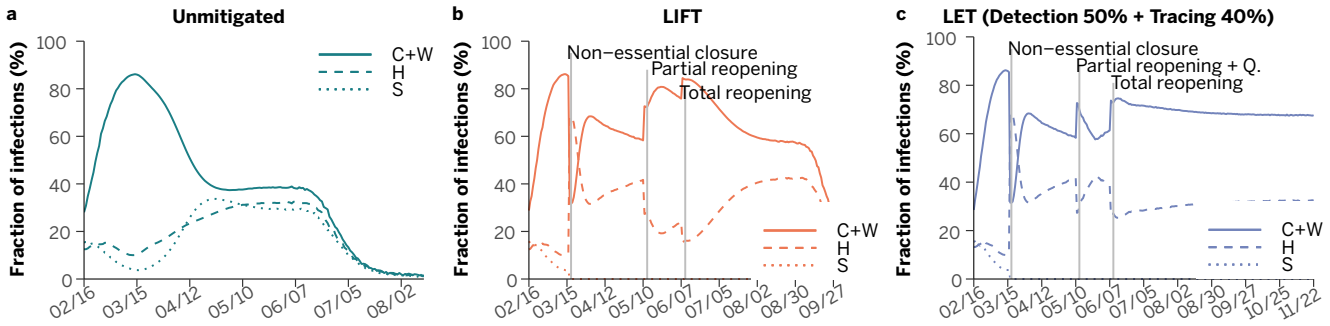

Figure 9: Distribution of infections across layers in the unmitigated, LIFT and LET (with 50% detection of symptomatic individuals and 40% tracing) scenarios. Lines represent the average fraction of infections that take place in each layer out of all the infections that have been produced each day.

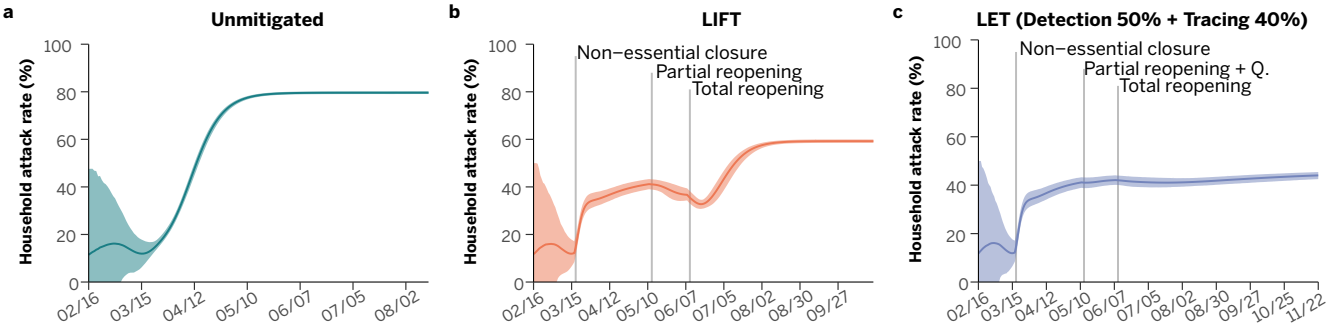

Figure 10: Household attack rate as a function of time in the unmitigated scenario, **a**, LIFT strategy, **b**, and LET strategy with 50% detection of symptomatic individuals and 40% tracing, **c**. Solid lines represent the average value while the shaded region shows the 95% C.I.

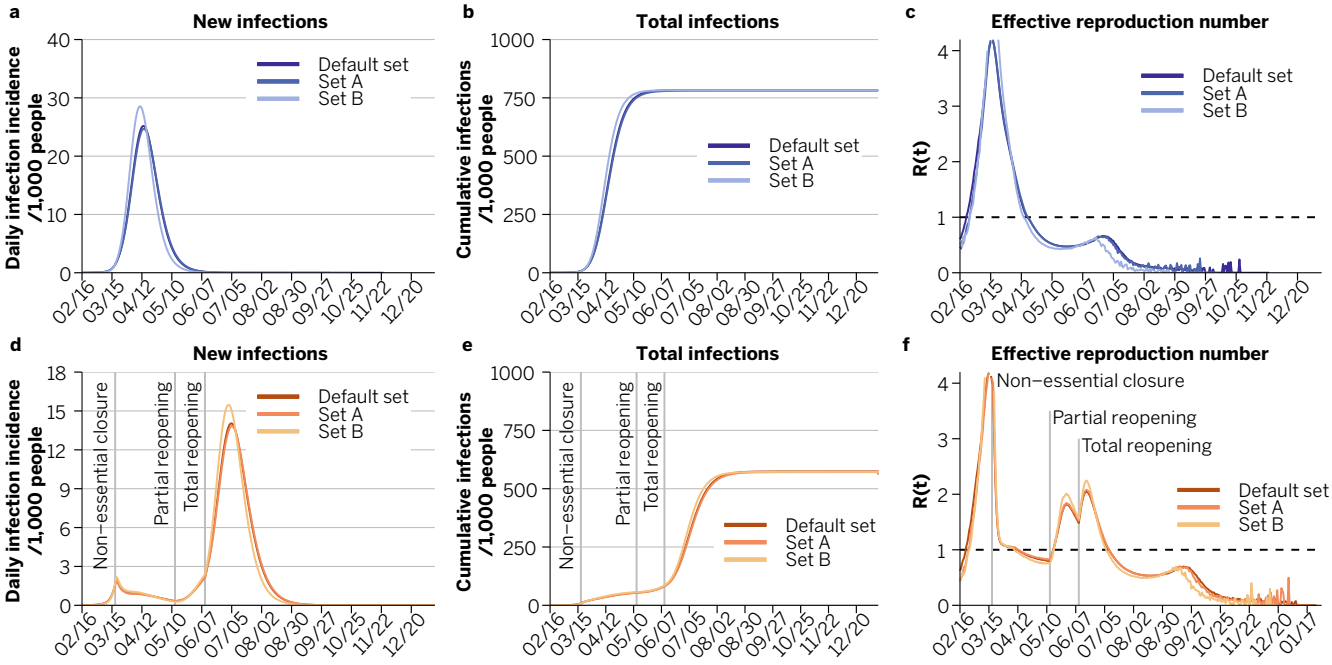

Figure 11: Evolution with larger fraction of asymptomatic individuals (set A) and larger pre-symptomatic transmission (set B) for the unmitigated scenario and LIFT strategy.

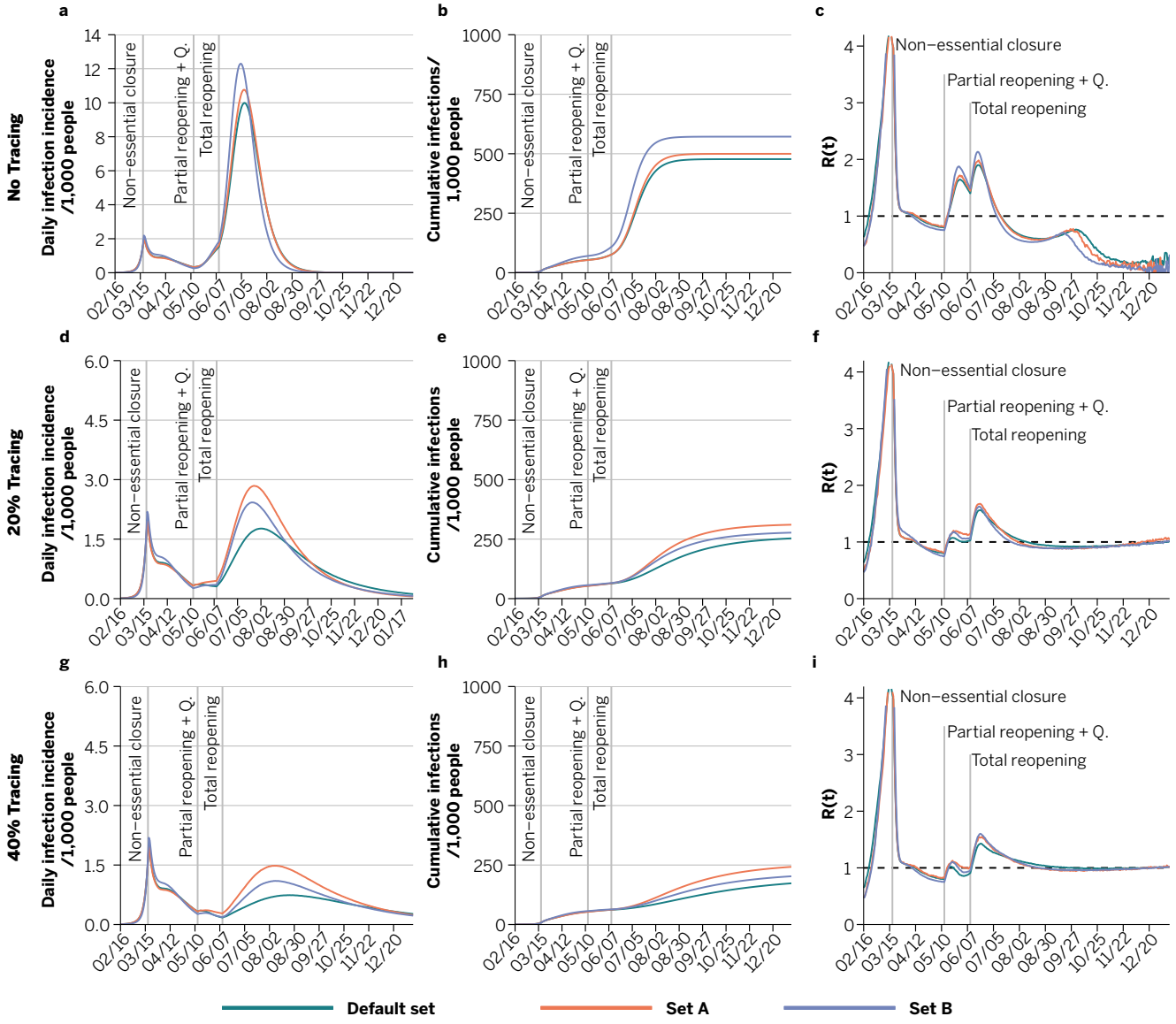

Figure 12: Evolution with larger fraction of asymptomatic individuals (set A) and larger pre-symptomatic transmission (set B) for the LET scenario with 50% of symptomatic detection. The top, mid and bottom rows show the results with 0%, 20% and 40% of tracing, respectively.

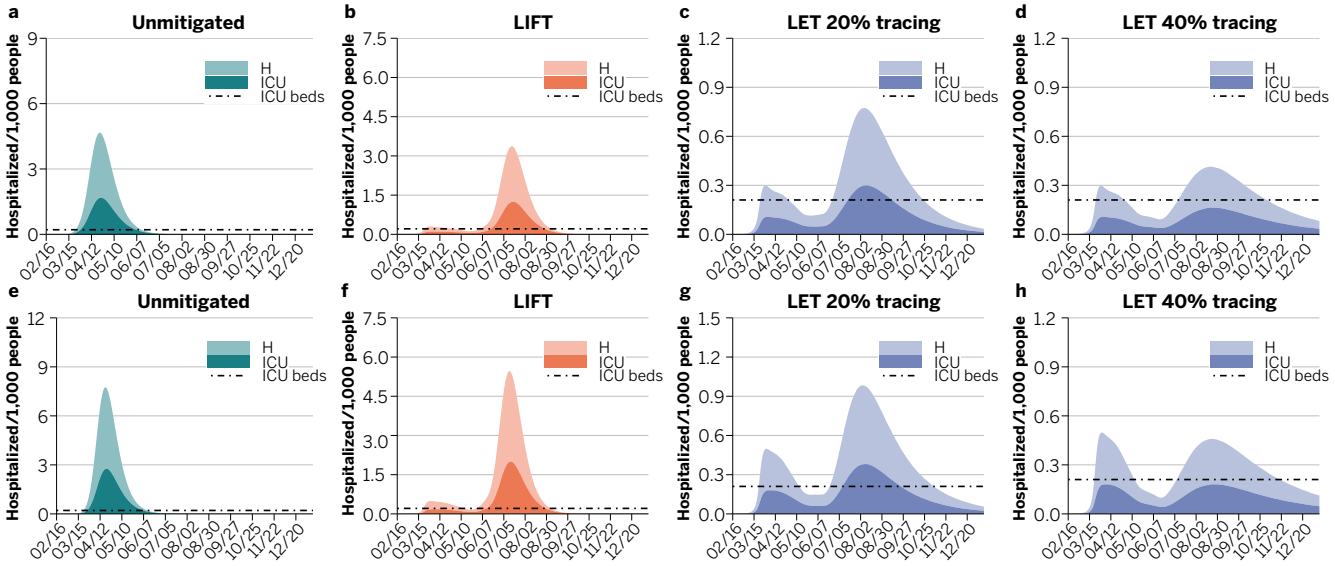

Figure 13: Evolution of the hospitalizations. Top row: set of parameters A (50% of asymptomatic individuals). Bottom row: set of parameters B (fraction of pre-symptomatic transmission $k = 0.50$).

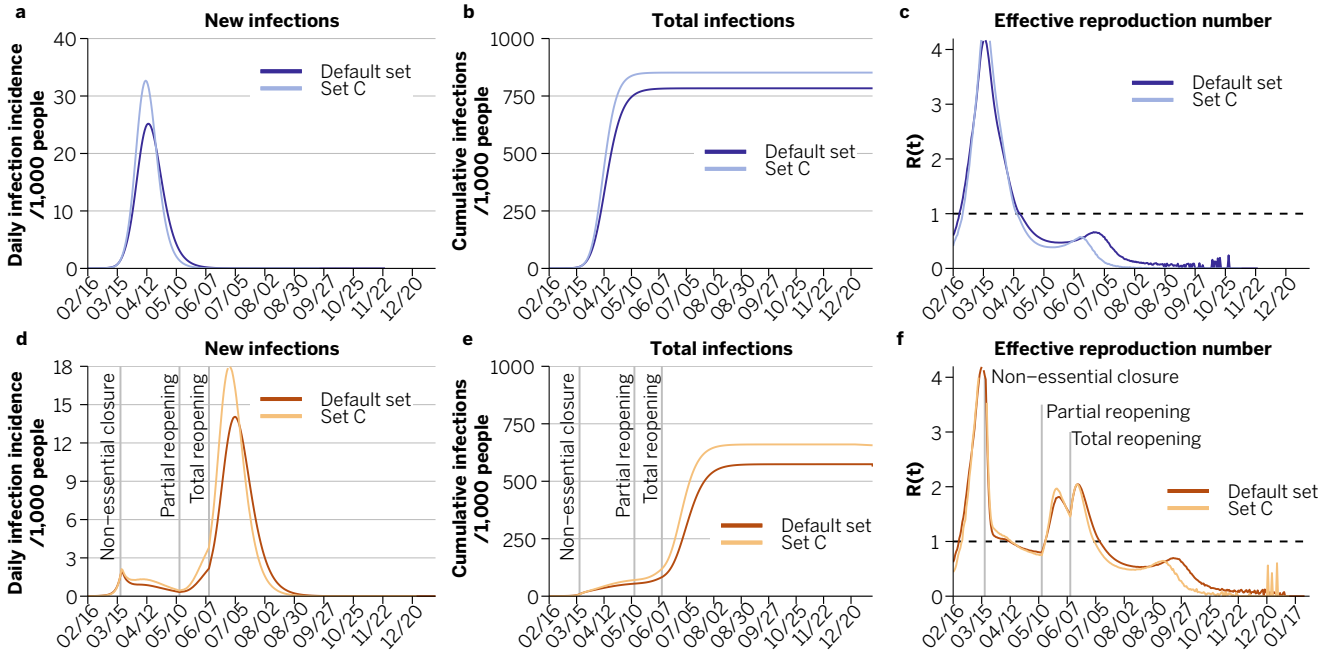

Figure 14: Evolution with larger R_0 (set C) for the unmitigated scenario and LIFT strategy.

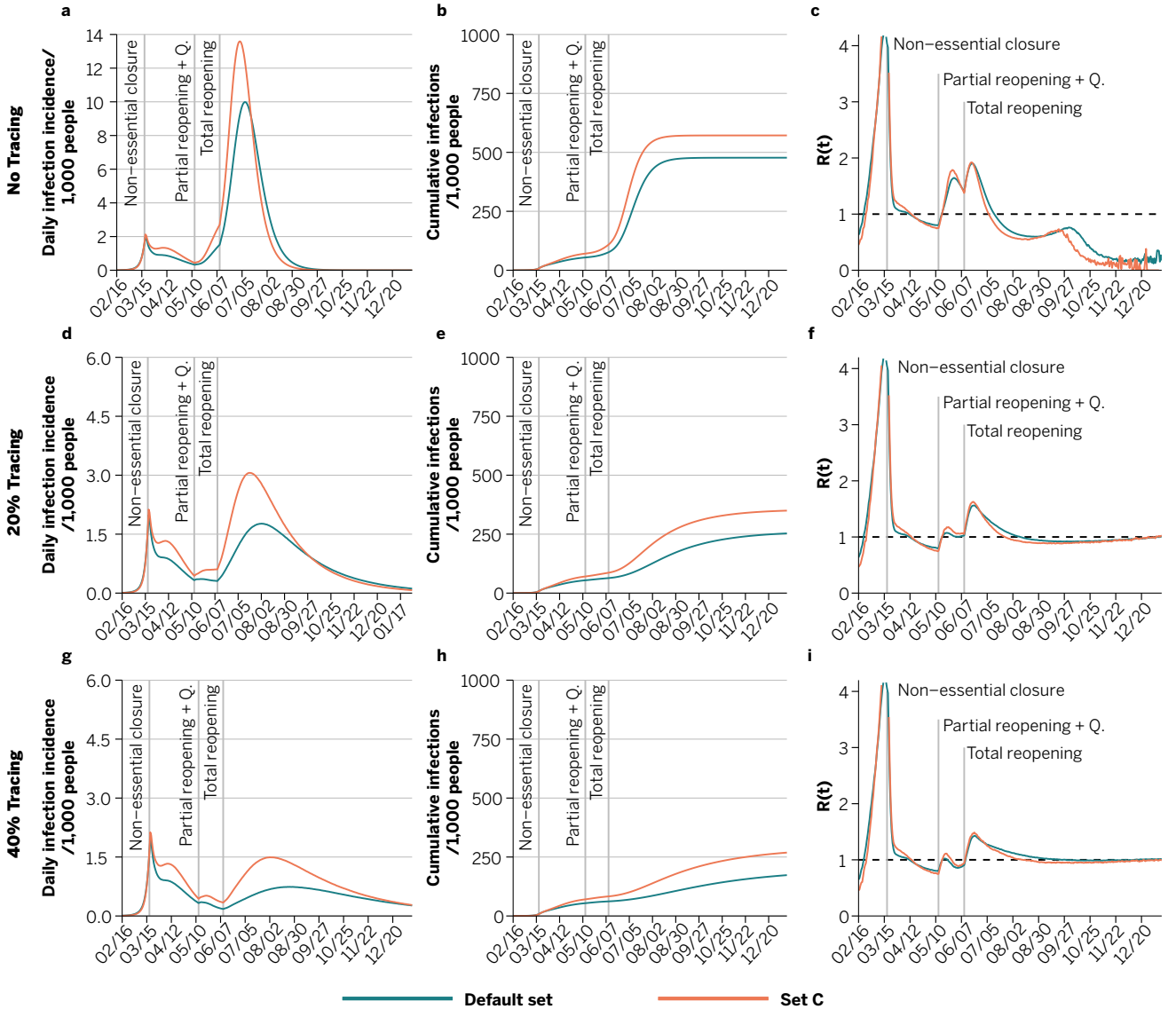

Figure 15: Evolution with larger R_0 (set C) for the LET scenario with 50% of symptomatic detection. The top, mid and bottom rows show the results with 0%, 20% and 40% of tracing, respectively.

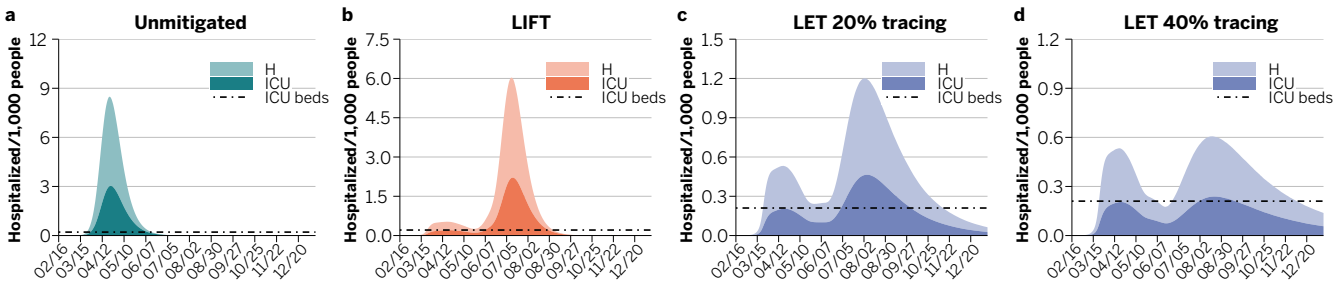

Figure 16: Evolution of the hospitalizations with larger R_0 (set C).

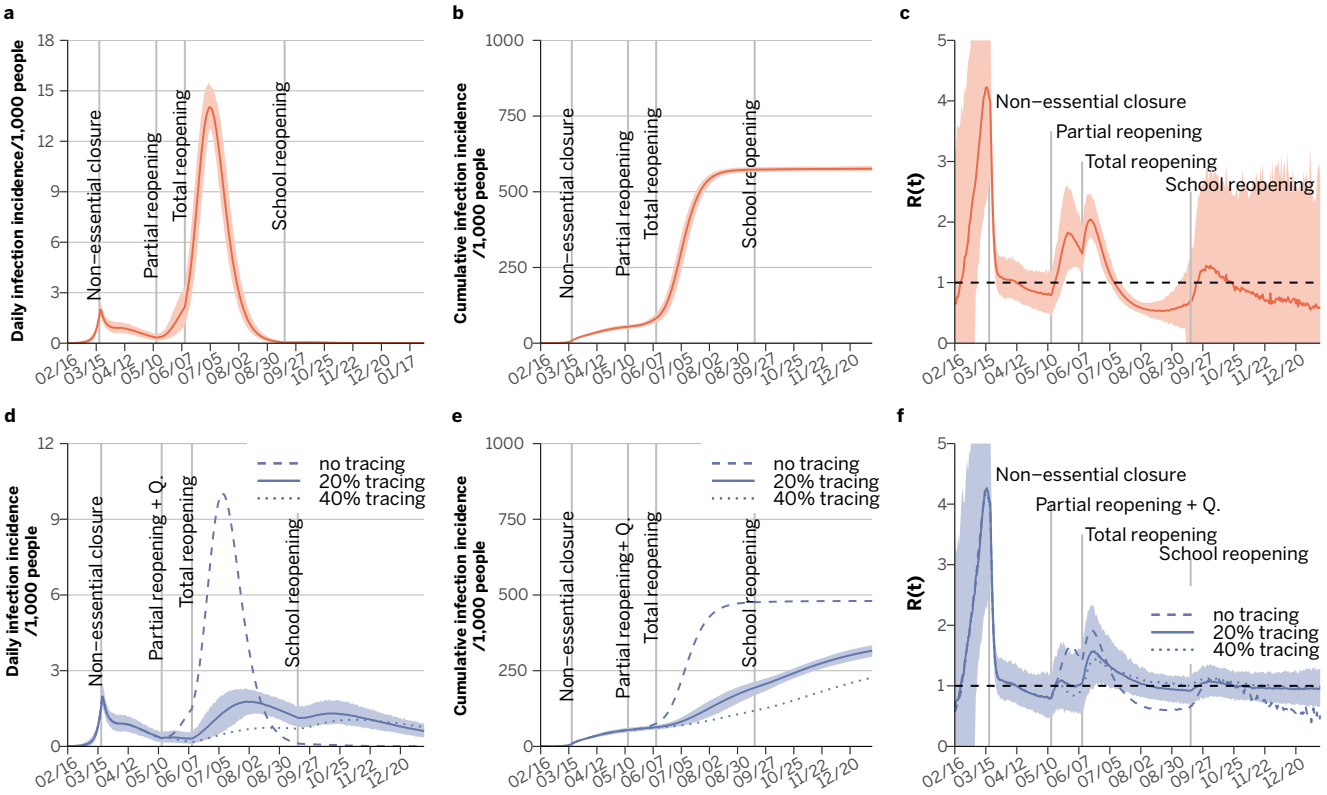

Figure 17: Evolution of the system when 30% of the symptomatic individuals are isolated at their homes and their contacts traced and quarantined in their households, default set of parameters.

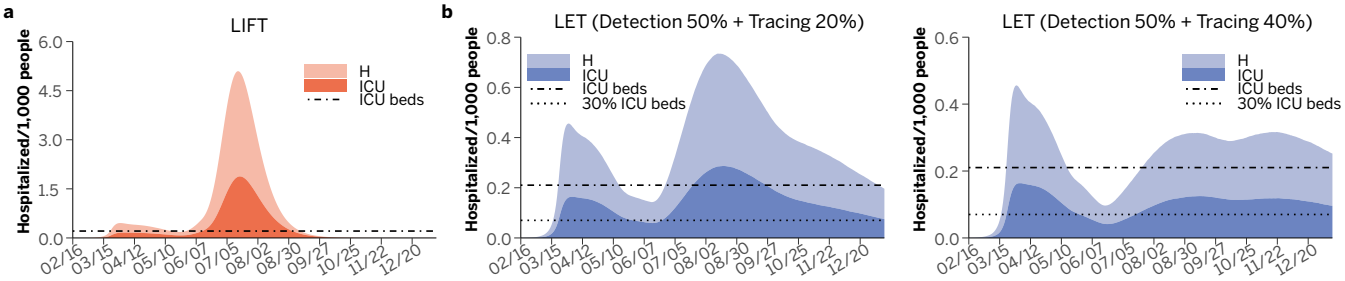

Figure 18: Evolution of the system when 70% of the symptomatic individuals are isolated at their homes and their contacts traced and quarantined in their households, default set of parameters.

Figure 19: Evolution of the system when there is a 1 day delay between the isolation of a symptomatic individuals and the tracing and isolation of its contacts.

Figure 20: Evolution of the system when there is a 2 days delay between the isolation of a symptomatic individuals and the tracing and isolation of its contacts.

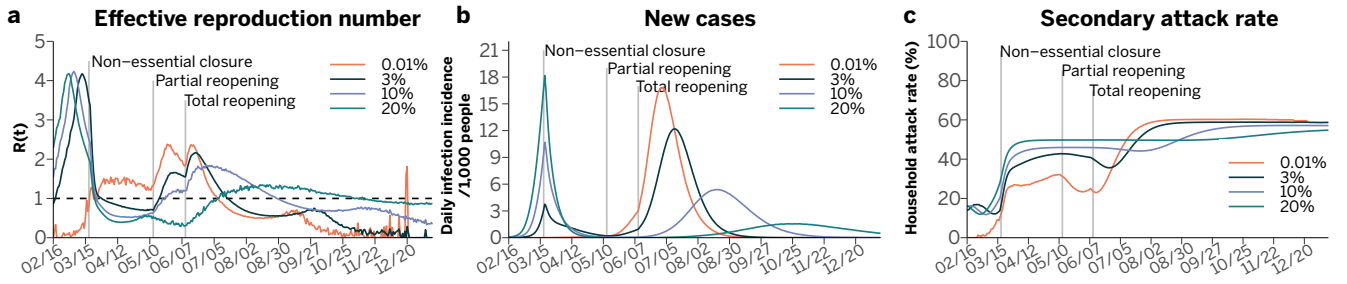

Figure 21: Effective reproduction number, daily incidence and household attack rate for different initial attack rates. The sooner the intervention, the larger the second wave can be.

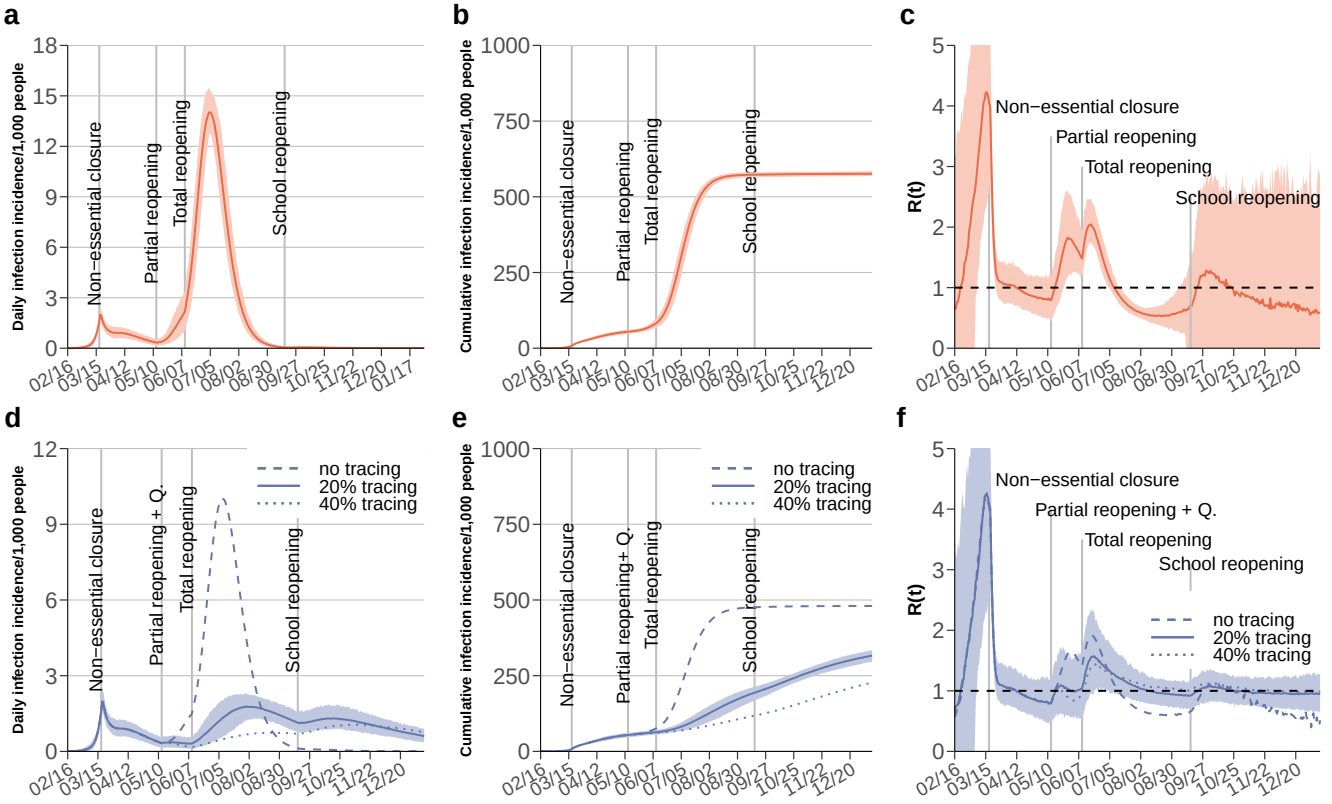

Figure 22: Evolution of the LIFT and LET scenarios when schools are reopened on the 15th of September.

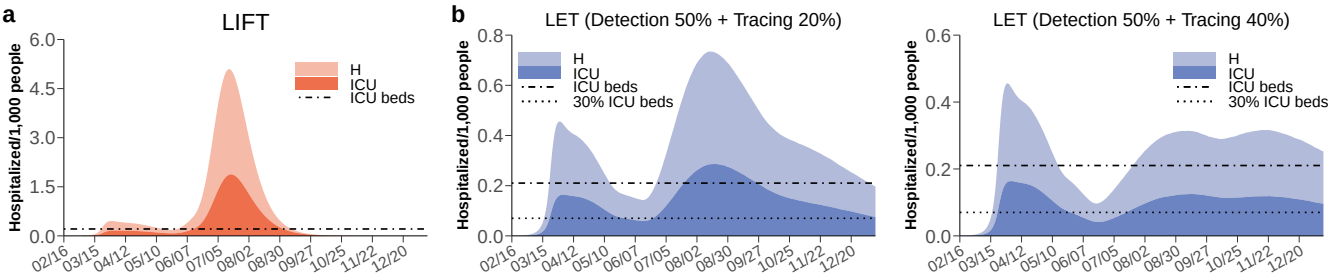

Figure 23: Evolution of hospitalizations in the LIFT and LET scenarios when schools are reopened on the 15th of September.

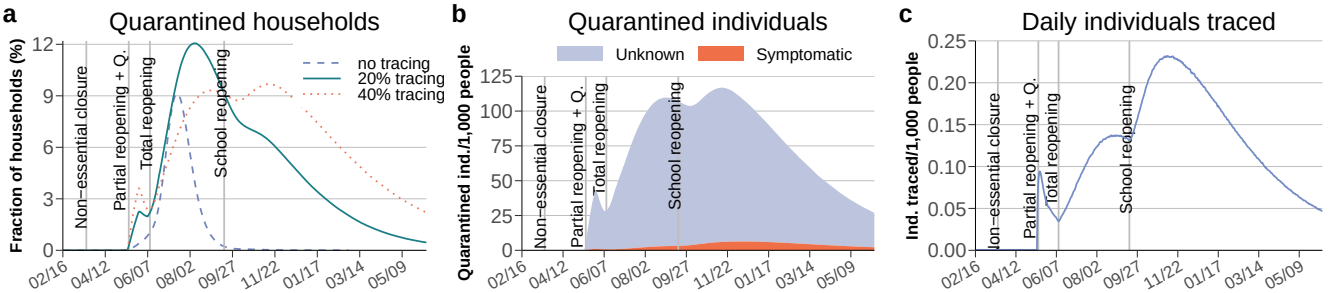

Figure 24: Affordability of the quarantine strategy when schools are reopened on the 15th of September.

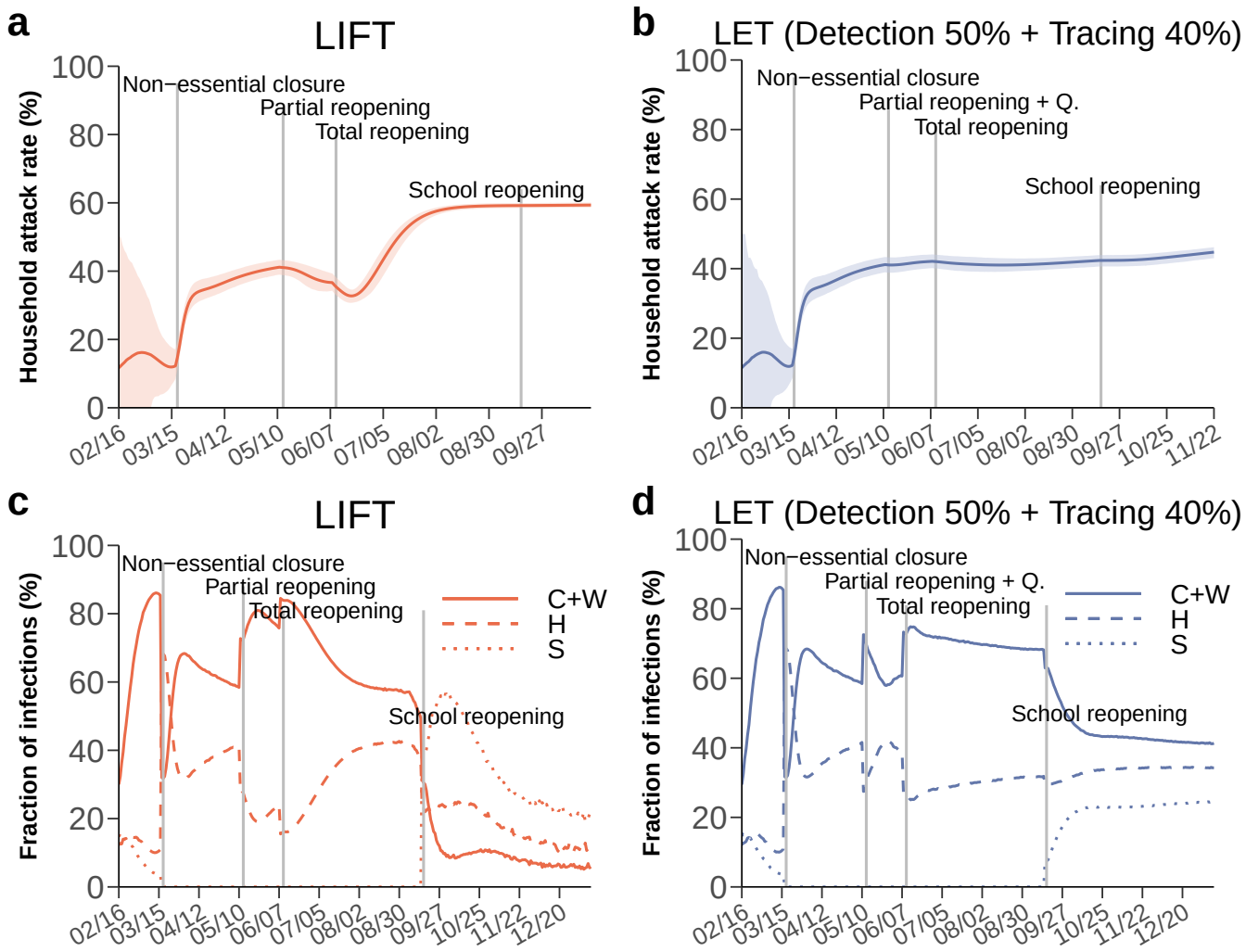

Figure 25: Household attack rate (**a**, **b**) and fraction of infections that take place in each setting (**c**, **d**) when schools are reopened on the 15th of September.

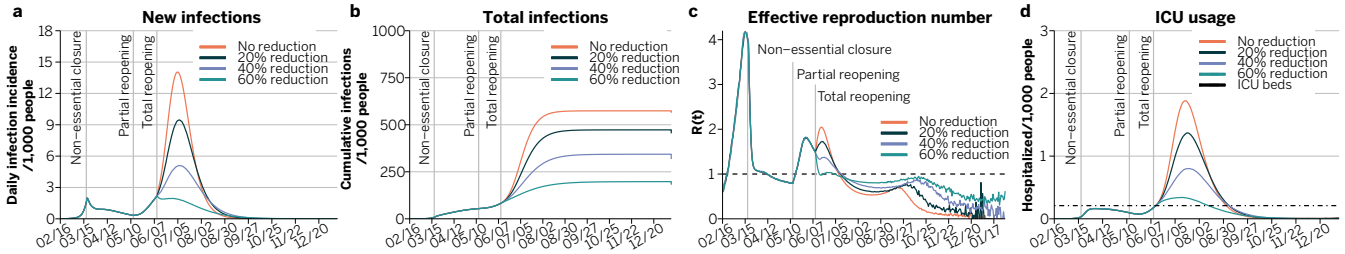

Figure 26: Evolution of the system under the LIFT scenario when the transmissibility is reduced by 20, 40 or 60%.

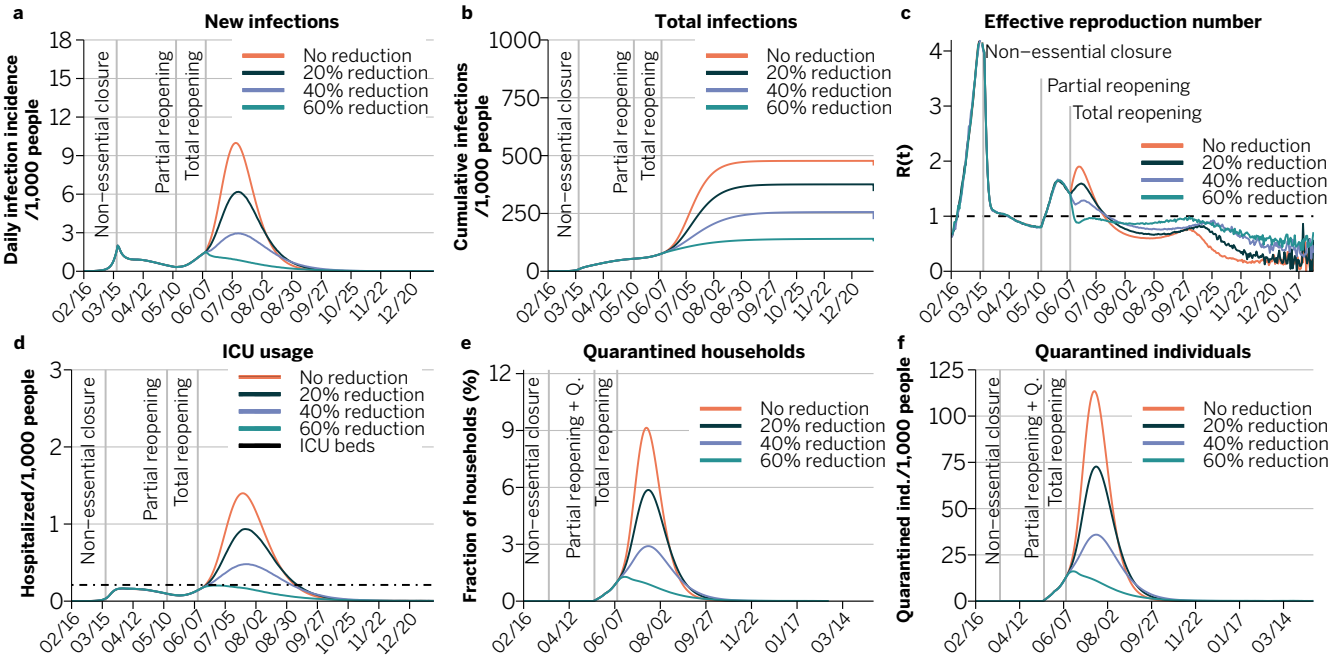

Figure 27: Evolution of the system under the LET scenario when the transmissibility is reduced by 20, 40 or 60%, with 50% symptomatic detection and isolation but without tracing.

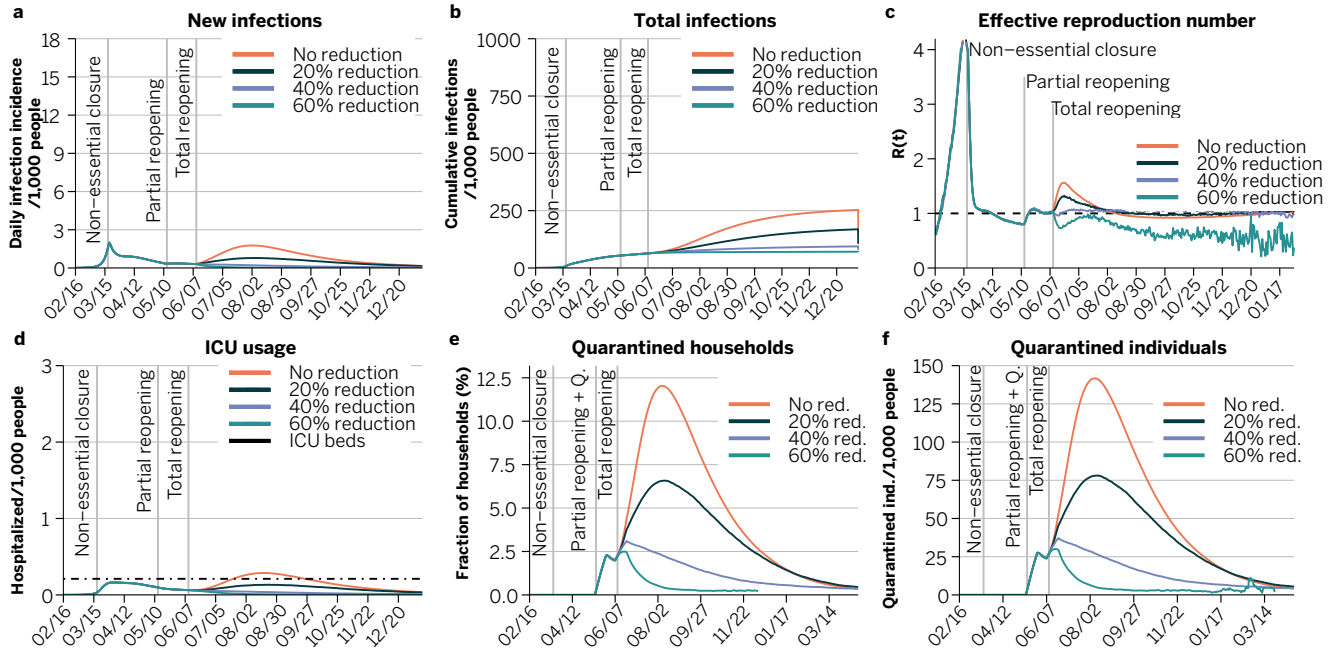

Figure 28: Evolution of the system under the LET scenario when the transmissibility is reduced by 20, 40 or 60%, with 50% symptomatic detection and isolation and tracing of 20% of the contacts.