Sociologica. V.14 N.1 (2020)
ISSN 1971-8853

Complex Social Networks are Missing in the Dominant COVID-19 Epidemic Models

Gianluca Manzo*

Submitted: April 18, 2020 – Revised version: May 3, 2020
Accepted: May 4, 2020 – Published: May 20, 2020

Abstract

In the COVID-19 crisis, compartmental models have been largely used to predict the macroscopic dynamics of infections and deaths and to assess different non-pharmaceutical interventions aimed to contain the microscopic dynamics of person-to-person contagions. Evidence shows that the predictions of these models are affected by high levels of uncertainty. However, the link between predictions and interventions is rarely questioned and a critical scrutiny of the dependency of interventions on model assumptions is missing in public debate. In this article, I have examined the building blocks of compartmental epidemic models so influential in the current crisis. A close look suggests that these models can only lead to one type of intervention, i.e., interventions that indifferently concern large subsets of the population or even the overall population. This is because they look at virus diffusion without modelling the topology of social interactions. Therefore, they cannot assess any targeted interventions that could surgically isolate specific individuals and/or cutting particular person-to-person transmission paths. If complex social networks are seriously considered, more sophisticated interventions can be explored that apply to specific categories or sets of individuals with expected collective benefits. In the last section of the article, I sketch a research agenda to promote a new generation of network-driven epidemic models.

Keywords: compartmental models; ego-centered networks; scale-free networks; small-world networks; agent-based computational models.

* Gemass, CNRS, Sorbonne University (France); gianluca.manzo@cnrs.fr; https://orcid.org/0000-0002-4905-2878

1 Introduction[1]

On March 17, 2020, the *New York Times* published an article promising to explain what lied *Behind the virus report that jarred the U.S. and the U.K. to action*; on March 25, 2020, *Science* published a detailed piece on *Mathematics of life and death: How disease models shape national shutdowns and other pandemic policies*; finally, on April 3, 2020, *Nature* followed with a special report titled *The simulations driving the world's response to COVID-19: How epidemiologists rushed to model the coronavirus pandemic*. These articles succeeded in showing, to a large audience, how public authorities' decisions on strategies to contain COVID-19 epidemic has temporally strictly followed the moment when mathematical epidemiologists reported to them on simulations' results of various non-pharmaceutical interventions (given hospitals' care capacity) on COVID-19 spread. These articles, however, are less useful to think critically about the mathematical models they pedagogically describe. Although journalists and scientific writers explicitly stress that the models' numerical predictions may be unreliable, in particular due to the poor quality of data needed for models' calibration, the link between prediction and intervention established by modellers is never explicitly questioned. In particular, there is no critical reflection on the kind of interventions that scientists can imagine *given* their models' substantive assumptions. In addition of being a device to compute, a model is a tool to think. Thus, the way a model is designed is also probably shaping the way we think about the means to intervene in the social world (provided that we consider the model as a legitimate guide for this purpose).

As suggested by Edmonds et al. (2019) in connection with computational models, modellers themselves are often not explicit enough on what purpose their model was designed for. This is problematic, the authors argue, because simplifications that may be acceptable for one purpose, say prediction (i.e., according to the authors, the task of reliable anticipation of numerical unknown quantities), may not be equally inconsequential for another purpose, say explanation (i.e., according to the authors, the task of defining specific chains of events linking initial conditions to a final outcome). This led Edmonds et al. (2019, line 10.2) to argue that using the same model for different purposes has critical implications, and to suggest that it is better to design, perhaps related, but separated models for pursuing different goals. Within the context of the COVID-19 crisis, Squazzoni et al.'s (2020) has recently re-considered Edmonds et al.'s typology of computational models' purposes and discussed the link between prediction and intervention in a crisis context (see, in particular, the appendix, and Table 1). In particular, they suggest that models' predictions may be used to "compare intervention scenarios (including the consequences of doing nothing) (...);" in this sense, prediction "could make a valuable contribution to discussion over interventions." They caution however against several risks of building interventions over predictions like "over-reliance on the model as an 'oracle'," "inappropriate political exposure of developers," "inability of an effectiveness-focused model to forecast policy utility" or still poor "quality of data to calibrate important model parameters."

In this article, I propose to develop further Squazzoni et al.'s remarks (2020) on the relationship between prediction and intervention by focusing on the specific class of mathematical models that are driving policy interventions in the COVID-19 crisis context, i.e., compartmental models (also called SIR models). For reasons that I will explain below, one may accept that these models can be fruitfully employed to predict several aspects of the virus diffusion at the

1. The present text develops a short public-style piece that I published in French in *Le Monde* on April 14, as well as a longer, more academic-oriented essay appeared on April 21 on the French debate website *La Vie des Idées* (https://laviedesidees.fr/Les-reseaux-sociaux-dans-la-lutte-contre-le-Covid-19.html).

macroscopic scale, such as the number of infections, the number of cases requiring critical hospital assistance or the number of deaths. However, I will claim that these models should be regarded as problematic with respect to intervention. In particular, I will argue that their specific content can only lead to one type of intervention, i.e., interventions that indifferently concern large subsets of the population or even the overall population. These models cannot help targeting interventions aimed at surgically isolate specific individuals and actor-to-actor particular diffusion paths. The reason is simple: the variants of SIR models used in the current crisis context address virus diffusion without modelling the topology of social interactions realistically. As we will see, SIR models including explicit representations of network topologies have been advocated as a necessary improvement of classical compartmental models at least since early 2000 (see, for instance, Keeling & Eames, 2005). However, SIR models' variants supporting policy decision-making in the current crisis continue to ignore this stream of the literature; the consequences of not properly considering social networks for intervention are not systematically discussed. My main argument is that this is regrettable, and the debate should be re-oriented toward network-based SIR models as well as toward how appropriate empirical network data to nourish these models can be collected at a very large scale.

To defend my argument, I will first shortly describe the way mathematical models progressively led French authorities to opt for a generalized lockdown. I will then describe the specific models that were so influential on political decisions in France. Then, I will develop the real point of this article: I will explain how social interactions are (mis)represented in these models and the implications of this in terms of intervention design. Finally, I will propose a research agenda for a new generation of epidemic models, which will be ideally better equipped for the next pandemics we will have to face. Reasonably, the fact that I will focus on a specific national case does not limit the scope of my analysis. As shown by *Science*'s and *Nature*'s review articles mentioned above, French experts indeed relied on the same type of models as their counterparts in US and other European countries.

2 Models and COVID-19 French Measures: From March 11 to April 13, 2020

The first three positive cases of COVID-19 in France have been found in January 24, 2020. They were Chinese citizens, all with relevant connections to the Hubei province. During the following weeks, French authorities were convinced that the virus was not present in the French national territory. The two first cases concerning French nationals were probably found on February 12 and 13 in two villages close to the small city of Creil, about 70 kilometres North-East to Paris, where an important French military basis is located. On January 31, 2020, a French military aircraft repatriated 193 French citizens landed from Wuhan. However, the link between this event and the two contaminations in mid-February is still controversial (see *Le Monde*, April 10, pp. 20–21). From February 17 to 21, a public gathering of about two thousand evangelists coming from many locations in the country took place in Mulhouse, about 500 kilometres South-East of Paris, close to German and Swiss borders. The official counting of COVID-19 cases by the French Health Agency started from February 25, when 13 cases were officially recorded. At that time, quarantine for infected persons and partial movement restrictions were recommended only at the local level in the two abovementioned areas where the virus diffusion was reasonably taking place. On February 26, the UEFA Champion League match between Olympic Lyon and Juventus Turin took place as expected in Lyon, about 500 kilometres straight South to Paris.

On March 11, 2281 confirmed cases of COVID-19 were officially recorded in France. The same day, at the invitation of French President, the Minister of Solidarity and Health formed a scientific committee, bringing together several specialists (i.e., three infectious disease experts, one virologist, one intensive care specialist, one immunologist, two epidemiologists, one anthropologist, one sociologist and one city doctor). According to the formula adopted by the press release (https://solidarites-sante.gouv.fr/IMG/pdf/200311-_conseil_scientifique.pdf), this committee aimed "to enlighten public decisions in the handling of Coronavirus-related health situation." This was a turning point. Following the scientific committee's policy recommendations, a control strategy was initially organized at the national level, including massive and generalized measures restricting individuals' freedom of movement.

On March 12, 2020, the freshly nominated scientific committee released its first judgement ("avis").[2] Among several other elements (following the Italian situation), the judgement explicitly referred to "mathematical models" simulating the evolution of the epidemic with (or without) interventions aiming at drastically reducing social interactions (see p. 2). As we will see later, the committee was especially considering predictions produced by Neil Ferguson's team at the Imperial College London. Among these interventions, the French scientific committee advised "home quarantine for infected persons and all their intra-household contacts," "home confinement of persons above 70 and all health-fragile persons," "reduction of public transportations," "closure" of school system at any level, "homeworking," and the "reduction of all non-essential social and leisure activities" (see p. 5). At 8pm of the same day, in his first Covid-19-related public discourse to the nation, the French President announced the closure of the school system (from Monday, March 16) and firmly invited all French citizens to limit to the minimum any movement outside home.

On March 14, 2020, the scientific committee released a second judgment, explicitly mentioning "known mathematical models" (see p. 1). This judgment recommended to implement more severe "restrictions to the social life," in particular the closure of "all leisure places, including restaurants, pubs, cafés," "shops, except food- and health-related ones," "gym clubs," "swimming pools," "museums," "theatres," "cinemas," and "nightclubs," as well as the suspension of all "church services." At 8pm of the same day, the Prime Minister publicly announced that these measures were effective from midnight.

On March 16, 2020, the scientific committee released a third judgement. It started with the regret that, despite the measures announced during the previous days, "the behaviour of a fraction of French population did not sufficiently change." For this reason, the committee recommended to increase even more the restrictions to individuals' freedom of movement, namely through "the implementation of a strict generalized confinement following the Italian model, at the scale of the country for the entire population, with the possibility to implement even more severe measures within the areas where the virus spread is the strongest" (see point 1, p. 2). Mathematical models again supported the conclusions: "a reduction of at least 60% of social interactions is necessary to contain the epidemic dynamic as shown by some modelling studies" (*ibidem*). At 8 pm of the same day, in his second Covid-19-related public discourse to the nation, the French President announced the generalized home confinements at the national level and the associated systems of sanctions for lack of compliance (from March 18, at noon). These measures were planned to last in two weeks.

On March 23, 2020, the scientific committee released a fourth judgement including two interesting novelties. On the one hand, it was firmly stated that the generalized home confine-

2. All reports can be freely downloaded by date at: https://solidarites-sante.gouv.fr/actualites/presse/dossiers-de-presse/article/covid-19-conseil-scientifique-covid-19.

ment should last for "reasonably six weeks at least;" on the other hand, the committee asked for further restrictions of individuals' freedom of spatial movements. This fourth judgement did not explicitly refer to any mathematical models. It is however important to note that the mathematical models to which the scientific committee was referring in its first judgement (i.e., those studied by Neil Ferguson and his team) were simulated under the hypothesis that the generalized home confinement was in place for "5 months or longer" (see Ferguson et al., 2020, p. 4). Again at 8pm of the same day, the Prime Minister publicly announced more restrictive conditions under which exceptions to leave one's home were tolerated; four days later, on March 27, 2020, he extended the home generalized confinement (under the new more severe conditions) until April 15. In its sixth judgement, released on April 2, the scientific committee, although initiating a reflection on what criteria should be adopted to decide whether a progressive relax of some freedom restrictions was possible, clearly restated that, for the moment, "the priority still is the continuation of a reinforced confinement on the long term" (see p. 7).[3] During his third Covid-19-related public discourse to the nation on April 13, the French President announced that this in fact was the case at least until May 11, 2020!

Thus, given the way political decisions followed the scientific committee's judgements, although the scientific committee explicitly cautioned that mathematical models "should be taken as only one element informing decision-makers" (judgment of March 12, p. 2), the hypothesis that these models worked as a compass for the French authorities is plausible. Pragmatically, in the short run, this made totally sense. Although rare, *post-hoc* assessment of the quality of the macroscopic predictions of this type of models suggests that these predictions are reliable as long as relatively short periods of time, say, one to two weeks, are considered (see, for instance, Camacho et al., 2019). The French scientific committee cannot be blamed for not taking these predictions with prudence. They indeed based their recommendations on "intermediary hypotheses that were sufficient to make visible a highly plausible and strong mismatch between the number of cases needing intensive care assistance and French hospitals' capacities, even when these capacities were reinforced with appropriate reorganization strategies" (judgment of March 12, p. 2). What should be thus critically scrutinized is not the predictive utility of the mathematical models at hand. Under the pressure of the emergency situation, modellers provided policy-makers with rough but precious information on the *macroscopic* epidemic dynamics. In such a crisis context, it seems rational to value more the order of magnitude than the precision of point-estimates.

However, can these models be used to go beyond predictions? More precisely: are they equally useful to provide recipes for intervening on the *microscopic* dynamics of contagions in order to stop it? Presently, this seems to me the crucial question. One should indeed remark that the same type of mathematical models, which guided policy-makers in the fire of the epidemics, now also shape thinking about exit strategies from the generalized lockdown (for France, see Di Domenico et al., 2020; for other countries, see, for instance, Kissler et al., 2020). This is the real issue. Can these models help us to design interventions that are targeted at individuals (or small subsets of them) and person-to-person diffusion paths, which may play a special role in the virus spread? Or, in contrast, do these models increase the probability that scientists and policy-makers can only think in terms of generalized interventions concerning large and undifferentiated groups of people? To answer this question, we need to look into the building blocks of these models that are (partly) determining our everyday life.

3. The scientific committee delivered a fifth judgement on March 30 that I do not consider here because it mainly concerns the specific, especially dramatic, situation of French retirement homes.

3 Ferguson's Microscopic Simulations and SIR Models in Epidemiology

The text of the first judgment — released by the French scientific committee on March 12, 2020 (see, in particular, p. 2) — suggests that French experts illustrated their policy recommendations to French authorities through the simulated scenarios of the number of infections, severe cases, and deaths, without and with more or less drastic non-pharmaceutical interventions, realized by the British mathematical epidemiologist Neil Ferguson, the head of the MRC Centre for Global Infectious Disease Analysis at the Imperial College London. Ferguson and his team started to deliver COVID-19 reports on January 17, 2020: at the moment of writing, fourteen reports have been published. Although the 12-March-judgement of the French scientific committee did not explicitly refer to Ferguson's recent papers (the mention "personal communication" only appeared in the text), it is reasonable to think that the "specific COVID-19 model" to which the committee referred in the text was the model finally published (on March 16) in Ferguson et al. (2020) report n° 9.[4] As shown by the articles in *The New York Times*, *Science* and *Nature* mentioned in the introduction, this report had a worldwide coverage and impact. Thus, to explain why Ferguson's simulations were adopted by French experts, the fact that one member of the French scientific committee was an internationally well-recognized mathematical epidemiologist, who regularly co-authored paper with Neil Ferguson since early 2000, is probably only an additional element.

At this point, a more precise analysis of the content of the model that Ferguson and his team employed to predict the diffusion of COVID-19 in UK and US, is necessary. In Ferguson et al. (2020, see in particular pp. 4–5), the model was presented as an "individual-based simulation model;" it was said that this model relied on a previous model that was designed to study the diffusion of *H5NI* in South Asia and the *influenza* in UK and US (the two articles describing the results of this original model were published in *Nature* on September 8, 2005 and July 27, 2006, respectively). As the authors honestly acknowledged, "the basic structure of the model remains as previously published." Thus, the French scientific committee regarded a model's structure designed at least 15 years ago for a different virus "as still being a reference for pandemic intervention" (see judgment from March 12, 2020, p. 2).

In more detail, Ferguson et al.'s model (2020, pp. 4–5) represents several millions of virtual individuals whose age and family size distributions follow empirically observation in the UK and US. Each virtual individual has a geographical location and is assigned to a school and/or a workplace, whose (spatial) distributions are also empirically calibrated to follow actual distributions. Within this virtual but realistic social environment, each artificial individual can probabilistically be in one of these three different states: "susceptible," "infective," and "recovered" (in this case, supposedly "immunized" at least on the short run). The heart of the model is the transmission mechanism of the virus from an infected virtual agent to a susceptible one. This mechanism combines three elements: 1. the number of days during which the infected individual is really contagious; 2. the probability the infected individual can effectively transmit the virus in an interaction with a susceptible individual (this probability being in turn set differently for infective-asymptomatic and infective-symptomatic individuals); 3. the number of social contacts of the infected individual. The simulation of the epidemic at the macroscopic level can then be performed by setting the value of a number of parameters (in addition to 1–3), among which are the precise timing of infected individuals' infectiousness (which again differs for infective-asymptomatic and infective-symptomatic individuals), the threshold for

4. All reports of Ferguson's team can be freely downloaded at https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/covid-19-reports/.

being considered a severe case requiring hospitalization, the length of hospitalization, and the probability of dying (as a function of the individual's age).

Thus, although Ferguson and his team qualified their model as a "stochastic, spatially structured individual-based simulation," to go back to the label adopted in *Nature* article in 2006, the simulations upon which the impact of non-pharmaceutical interventions on COVID-19 in the UK and US was tested was a clear example of the compartmental models dominating the approach of biologists and epidemiologists working on the diffusion of infectious diseases (for an introduction, see for instance, Keeling & Rohani, 2008). In their generic form, these models essentially study the temporal evolution of the fraction of individuals belonging to a pre-defined number of groups (or types), given a certain probability of state changing per unit time. As their most frequent variants postulate the existence of at least three groups, namely "Susceptible," "Infective" and "Recovered" individuals, these models are also known as SIR models (the label that, for simplicity, I will adopt hereafter).

The particularity of Ferguson et al.'s model (2020) is that they implemented the virus transmission mechanism at the actor-level, thus providing a micro-simulated version of a SIR model. These models are in fact usually specified at the group-scale and only describe the evolution of the size of each group at the aggregate level through various forms of differential equations. This may indeed partly explain the reason of the success of classical SIR models (I will point to a second possible factor related to empirical data in the last section). They are (relatively) simple and can be studied (often analytically) through a powerful and widely theorized mathematical formalism. This a well-known mechanism of formal models' success. Unrealistic models can durably dominate a given research domain because they are supported by a practical mathematical technique: for example, think of macroeconomic general equilibrium models based on unrealistic microscopic assumptions, such as utility maximization and representative agents, but supported by powerful *calculus* techniques (see on this point Kirman, 1992).

Ferguson and his team confirmed themselves the close relationship between their "stochastic, spatially structured individual-based simulation" and meta-population SIR models, when, in the report n° 12, published on March 23, 2020 (see Walker et al., 2020), they considered the aggregate level to extend to 200 other countries their counterfactual analysis of the diffusion of COVID-19, with and without non-pharmaceutical interventions. In this case, as they did not have access to data as rich as those they could exploit for the case of the US and UK, they returned to more classical SIR (and SEIR) models, which confirmed the order of magnitude of the predictions initially generated through the individual-based simulations when applied to the US and the UK. The fact that microscopic implementations of SIR models require richer data is well known. When these data exist, the added value of individual-based SIR models is that they can lead to more precise predictions at smaller spatial scales (like cities or regions) and for smaller subgroups of the population (for this kind of cross-model comparisons, see, for instance, Ajelli et al., 2010).

However, after all, why are SIR models so appealing? The reason seems to be their promise to deliver, and monitor, a fundamental epidemiological quantity, i.e., the basic reproduction rate (noted R_o), which amounts to the average number of individuals that an infected individual can contaminate for as long as s/he is contagious. This is a crucial parameter because, according to standard SIR-model theory, the value of R_o gives us the *critical threshold* of a diffusion process, i.e., the tipping point beyond which the virus spread is supposed to slow down, and, eventually, stop. I will come back to this in the next section. What matters now is that this rate in fact arises from three components that Ferguson's individual-based simulation allows to distinguish at the microscopic level. These are: the period of time during which an

infected individual is contagious, the probability s/he effectively transmits the virus during a given social interaction, and her/his number of social contacts. Obviously, the last factor plays a crucial role within the model's dynamic: the action of the first two factors in fact unfolds again and again at each dyadic interaction. But how do dominant SIR models represent such interactions? This is the crucial point that must be understood to assess what kind of policy interventions these models can in fact deliver to slow down the macroscopic dynamics of contaminations and deaths.

4 Social Interactions in Dominant SIR Models

Within Ferguson et al.'s (2020) individual-based simulations of COVID-19 diffusion in UK and US, each virtual individual has social contacts within the household, at school, in the workplace as well as in the wider community. The latter are assumed to happen "randomly" (*ibidem*, p. 4), even though their probability decreases with the geographical distance between individuals. Moreover, for all virtual individuals assigned to a school, contacts are assumed to be two times more numerous than elsewhere. At the population-level — the authors claimed — such a parameterization leads to a contact matrix where one third of social contacts occur within the household, another third at school and in workplaces, and the remaining third unfolds in the wider community. Within the aggregate SIR models subsequently developed by Ferguson and his team to extend the analysis to 200 new countries, the focus is on one specific feature of social contacts: their age-group structure (see Walker et al., 2020, figure 2d-f, p. 5).

A small detail should be remarked at this point. Although social distancing interventions are usually represented as generically aiming at the reduction of the reproduction rate R_o (i.e. the average number of individuals that an infected individual can contaminate as long as s/he is contagious), these interventions in fact only touch upon one specific component of R_o, i.e., the per-individual number of social contacts. If we carefully read Ferguson team's reports, it is evident that policy scenarios are formulated precisely in terms of the overall proportion of social contacts that this or that intervention is assumed temporally to inactivate (see, for instance, Ferguson et al., 2020, table 2; Walker et al., 2020, pp. 6–7). To some extent, this is surprising. One may in fact imagine also to manipulate another component of R_o, i.e., the probability of virus transmission within a given social interaction. This may be used, for instance, to represent systematic variation in individuals' propensity to respect (worldwide recommended) social-interaction-related basic protective measures (such as sneezing in the elbow or keep appropriate spatial distances). However, this is usually not considered. As it is overtly admitted, adaptive individuals' behaviours are not represented within the model (see, for instance, Ferguson et al., 2020, p. 3, p. 6, p. 15; Walker et al., 2020, p. 3; Flaxman et al., 2020, p. 1). Differential-equation-based SIR models are not very flexible in this respect. I will come back to this point in the last section.

Let us then move on to the crucial question. Which assumptions are we implicitly making when social distancing interventions are conceived as a generic reduction of, say, 75% of all social contacts outside household, which corresponds to the scenario of a population-wide social distancing simulated by Ferguson and his team (see, for instance, Ferguson et al., 2020, table 1)? I will discuss two of these assumptions, which to me seem to shape most profoundly the type of interventions designed by these models.

First, when SIR models, similarly to those built by Ferguson and his team in early March, simulate a temporal reduction of physical social interactions by a given factor, these models imply that all actors are equally important in terms of social connectivity. This ignores one of

the most important empirical findings of research on complex social networks during the last twenty years or so: social networks are highly heterogeneous in terms of degree distributions. This means that while a few actors are able to establish and sustain tens of social interactions, the vast majority of them only have a small amount of social contacts (Newman, 2003, line 3.3). Barabasi (2014, ch. 5), who largely contributed to understand this phenomenon, calls these special actors "connectors" or "hubs." As we have seen, some SIR models take into account possible age-group variations in social contacts: this clearly is an important improvement. However, given the usual size of these subgroups, actor-level heterogeneity in social connectivity is likely to persist. For example, consider sexual contacts. At a given age, and for a given sex, a few individuals continue to have tens of partners (over a period of one year, for instance), whereas the vast majority of individuals only have one, or a few, sexual relationships (see, for instance, Liljeros et al., 2001).

Secondly, in addition to ignoring hubs, interventions that are conceived as generically temporally inactivating a given proportion of physical social contacts, implicitly assume that, given the allowed average number of contacts per person, interactions within a given subgroup and/or social space occur at random. As these interactions are regarded as having the same value, they can be imagined as suddenly disappearing at once. This also ignores a crucial feature of social networks: not all interactions have the same probability of occurring. Contrary to what was initially postulated by random models of network formation, where ties were supposed to form independently with a fixed probability (see, for an overview, Newman, 2003, pp. 96–106), we now know that tie formation in empirical networks follows well-defined path dependencies. Social interactions do not follow random paths: they have a structure and are deeply constrained. Dominant SIR models continue to ignore these regularities both at the local and the global level. Let me briefly discuss these two scales in turn.

Social interactions have a structure at the level of each actor (i.e. ego-centered networks). Let us consider just one example: if two actors have a common contact, then it is more likely that they also establish a contact compared to the probability of this happening had the three ties formed at random. As already suggested by Granovetter's classic paper on weak ties (1973, p. 1362), this pressure toward triadic closure is deeply related to another property of social interactions, i.e., homophily — the tendency of contacts between similar actors to be more frequent than random. And we now know that homophily, and the associated pressure toward triadic closure, can accelerate the creation of contagion clusters (Jackson and Lopez-Pintado 2013). Indeed, if A and B share an infected contact C, they are more likely to contaminate each other too than if they do not have a common infected friend. This property of ego-centered networks may be exploited to identify local paths of contaminations that are more likely than others. Although it would be possible to design SIR-like models incorporating these ego-centered network phenomena (see, Block et al., 2020) and so eventually considering that social interactions are locally structured, this is simply impossible with the type of models that are used in the current crisis context. This limits the type of interventions that can be explored.

However, the structuration of social interactions is not limited to the actor-level. The combination of a myriad of ego-centered networks produces a higher-order structure that also exhibits well-defined properties. In particular, research on "small-world" topologies has shown that small sets of highly connected nodes can be related by long-range ties, increasing the probability that two actors belonging to two distant social circles get in contact, compared to a social world where actors where only locally paired. It was shown that these bridges precisely correspond to the interactions that can rapidly increase disease spread in a population (Watts & Strogatz, 1998). This property could be exploited to identify transmission paths that make

distant contagion clusters communicate: interrupting them probably leads to slow down the diffusion process. Again, although it would be possible to design compartment models incorporating this phenomenon (see, Brethouwer et al., 2020), the variant of SIR model currently dominating the policy decision-making process cannot help reasoning along these lines, because they do not pay attention to the global structure of social interactions.

The important point now is that there is a deep relationship between the high heterogeneity of actor degrees (i.e. their number of links), the high clustering of ego-centered networks and the high reachability of the global network resulting from the combination of such local circles. Indeed, hubs act as the social glue of social networks. Roughly, small hubs create links within and across relatively close local circles; medium hubs create connections between sets of local circles tied together by small hubs; and large hubs (i.e. the very rare nodes with an exceptionally high number of connections) keep the entire set of small and medium communities in communication. This explains why, when considered as a whole, social networks qualitatively resemble a tree-like structure (Barabasi & Bonabeau, 2003), where basic units are more or less internally densely connected communities (Lancichinetti et al., 2010). The typical statistical signature of this hierarchically modular structure is the inverse relationship between the node's degree and its clustering coefficient: the more links a hub has, the less of the hub's contacts are also connected. This suggests that the larger the hub, the larger its capacity to span over otherwise disconnected local circles (see Barabasi, 2014, pp. 232–237). The presence of large, medium and small hubs (i.e., network hierarchy) tying together more or less densely connected local communities (i.e., network modularity) has a crucial consequence: real-world (social) networks are resistant to perturbations concerning lots of low-degree-nodes whereas they are exposed to cascades of negative effects when hubs are attacked (see Barabasi, 2014, ch. 9). Hubs are the "Achilles' heel" of complex social networks.

Thus, if one had to draw only one implication from the observation that social interactions are not random but deeply structured, both locally and globally, this would be the following: to slow down a virus spread, we must try to identify and neutralize hubs, i.e., the actors with the highest probability to be involved in the highest number of relevant interactions per unit time. Hubs play a twofold crucial role in the contagion process: first, if they are not infected yet, they have the highest chance to become infected because of their larger social exposure; second, for the same reason, when they are infected, they can in turn quickly contaminate a larger number of not-yet-infected individuals. That is why hubs should be the priority of targeted interventions. To cut their ties is likely to generate larger systemic effects compared to what one may obtain if a large majority of low-degree-actors were isolated (see Barabasi, 2014, ch. 10). If one reaches and neutralizes the hubs, the networks can be indeed fragmented, thus leaving the virus vanish within smaller and smaller local contagious circles; actors with few interactions do not have such a capacity of setting local communities apart. This is a well-known fact for any scholar empirically studying the diffusion of sexually transmitted diseases in small groups: the (relatively) few actors having a large number of sexual contacts (what scholars called "concurrency" as opposed to possibly "serial monogamy") play a crucial role in the diffusion process and must be prioritized in any intervention (see, for instance, Morris & Kretzschmar, 1997; Rocha, Liljeros, & Holme, 2011; Moody & Benton, 2016). It is also true that there are different ways to conceptualize, identify and measure the hubs in network terms. Furthermore, we must recognise that interventions targeting hubs are likely to be more or less effective according to several aspects of the network topology (see, for a systematic analysis of these two points, Montes et al., 2020). However, my point here is that dominant SIR models continue to think in terms of average contacts and random interaction and this makes it impossible in the present

crisis to understand which specific individuals and diffusion paths we need to test, isolate, protect and/or treat (depending on the specific features of the hub and on material and practical changing constraints). Although reasonably constituting a quantitative minority, these cases are probably keys in accelerating the contagion dynamic at the microscopic level.[5]

To push the argument further, one may eventually consider that the formal proof of the crucial importance of taking seriously the structure of social interactions for modelling disease diffusion is due to certain variants of SIR models itself. Although they still seem a minority, SIR models integrating network topologies that have been studied by physicists at least since early 2000 (for a review, see Newman, 2003, pp. 229–233; more recently, see Duan et al., 2015, pp. 809–815). Results are radical. For instance, let us go back to the reproduction R_o, i.e., the average number of individuals that infected individuals can contaminate as long as they are contagious. The principle according to which the value of R_o should go below 1 — the critical threshold often mentioned also in public debate — in order to stop the COVID-19 epidemics, continue to provide the main compass to assess policy interventions. For example, when empirically evaluating how the epidemic reacted in eleven countries to various forms of social distancing, Ferguson's team stated:

> The key aim of these interventions is to reduce the effective reproduction number, Rt, of the infection, a fundamental epidemiological quantity representing the average number of infections, at time t, per infected case over the course of their infection. If Rt is maintained at less than 1, the incidence of new infections decreases, ultimately resulting in control of the epidemic. If Rt is greater than 1, then infections will increase (dependent on how much greater than 1 the reproduction number is) until the epidemic peaks and eventually declines due to acquisition of herd immunity. (Flaxman et al., 2020, p. 3).

But, under which hypotheses, were the mathematical proofs of classical SIR models showing the existence of such critical threshold? Hypotheses are as follows: (a) all infected actors have more or less the same probability to transmit the virus during each interaction; (b) these interactions occur randomly within the social space (a condition usually called "homogenous", "uniform" or "full" random mix). Analytical and/or numerical studies of SIR models relaxing these hypotheses and explicitly including network structures with realistic levels of degree heterogeneity showed that a value of R_o lower than 1 does not guarantee that the epidemic stops. On the contrary: when broad degree distributions are present, every virus can spread irrespectively of the postulated transmission probability (for a literature review on this point, see Barrat, Barthélemy, & Vespignani, 2008, ch. 9). The explanation of this result is again related to the potential action of the hubs. A reproduction rate lower than 1 can well arise from the co-existence of many actors being involved in very few interactions (thus depressing the overall value of R_o) and a few actors with tens of links. According to the specific underlying structure

5. It should be noted that, in addition to intervention, the disproportionate impact of hubs on the spread of a disease may also be crucial to explain deep statistical features of epidemics. Cirillo and Taleb (2020) have recently studied the distribution of deaths from 72 major epidemics of human history and found that the distribution of victims regularly presents extreme fat tails. This led them to claim that "epidemiological models like the SIR differential equations, should never be used for precautionary risk management, which should focus on maxima and tail exposures instead." Although their work is descriptive, they acknowledge that network analysis may help to identify "(…) mechanisms for the spreading of contagion and the existence of super spreaders, a plausible joint cause of fat tails." The analysis above helps to see why the relatively few individuals with a disproportionate high amount of connections may generate cascades of infections whose aggregate consequences can go far beyond those expected within a population of "average" individuals.

of social interactions, it may suffice that one, or a few, of these hubs be infected to have then the virus leaking through the network (see Barabasi, 2014, p. 135). That is why only knowing the distribution of social interactions per individual may not be sufficient to predict the exact unfold of the epidemic: the overall network topology should also be known.[6]

To sum up: the variants of SIR models that still dominate prediction and intervention in the current crisis are based on the (correct) hypothesis that the virus spread is fuelled by social interactions. However, due to the sake of simplicity, modellers do not explicitly represent the structure of these interactions, either at the local scale (i.e. ego-centered networks) or at the global scale (the topology of the resulting network). It is as if we pretended realistically to model car flows at a country level, and potentially associated traffic jams, without also modelling the networks of streets, routes, and freeways. Could this type of models go beyond recommendations advising everyone not to use the car or allowing only specific fractions of the population to take the route at specific times and days? I suspect they could not. One may also anticipate that many drivers would be highly dissatisfied with such generic and undifferentiated instructions. SIR models currently in use put each of us in a similar situation. The lack of route infrastructure within my fictive traffic model corresponds to the absence of the structure of social interactions with dominant SIR models. Similarly to the fictive traffic model that would be incapable to recommend to users specific paths to make traffic more fluid and travel at a higher speed, dominant SIR models do not allow to identify specific actors and diffusion path that should be neutralized, because of their especially high impact on the speed of the virus spread. Consequently, currently dominant SIR models limit us to think in terms of generalized interventions while we are not allowed to move, no matter of our relative impact on contagion dynamics. There is a crucial piece of reality that these models miss to include, and this limits our policy interventions, leaving us with no real choice.

5 An Agenda for a New Generation of Epidemic Models

There are many reasons why integrating the structure of social interactions within models of virus diffusion where there is human-to-human transmission is greatly rewarding. This may help us to think about, design and implement targeted interventions, i.e., interventions that focus on individuals and dyadic transmission paths that are likely to have the largest systemic impact on contagion's speed and penetration rate. This would avoid imposing to large and undifferentiated sectors of the population huge psychological, relational and economic costs. With respect to possible exit strategies from lockdown, attempts at adapting SIR models to the logic of targeted interventions are in place. In particular, research is focusing on the importance

6. It is important to note that, conceptually, we are not obliged to think in terms of social network analysis to pay attention to the disproportionate potential impact of hubs. Epidemiologists admit the possibility of "super spreaders" either as individuals with a special infectivity or as events where one (or a few) individuals can generate a high number of secondary infections in a short time. As shown by control strategies discussed by James, Pitchford and Plank (2007) from this perspective, however, not to consider "superspreading" in combination with social network analysis again leads to interventions that only target social events and/or large and undifferentiated fractions of the population. A similar event-based logic would consist in targeting specific types of places, and among them, specific sub-categories, as a function of how often and/or how intensely a place is publicly attended. This is what is done, for instance, by Benzell, Collis and Nicolaides (2020) for a long list of stores and locations in US by exploiting a massive dataset of geolocations from tens of millions of smartphone devices in. Again, from the network perspective I am endorsing here, this strategy is forced to assume that all individuals are potentially equivalent as a vector of the disease. As we do not know the characteristics of the hubs visiting a given location, their specific network position, their geographical origin when visiting a given place, we can only close entire categories of shops and activities.

of breaking connections that increase communication between distant social circles seems in progress.

For instance, Siegenfeld and Bar-Yam (2020) proposed to think in terms of basic reproduction rate across local communities rather than between individuals. They argued that interventions aiming at the reduction of connections between communities can slow down COVID-19 diffusion. In an actor-level analysis, Brethouwer et al. (2020) developed a SEIR model to explicitly represent a small-world topology; by differentiating actor-to-actor transmission probability through short- and long-range ties, they suggested that interventions targeting the latter type of links could help to slow down the virus spread. These studies are tremendously useful to illustrate the impact that the network topologies can have on the type of alternative policy interventions in epidemic models. Unfortunately, they did not incorporate empirical data on social interactions on a large scale. As a consequence, targeted interventions can only be illustrated in abstract and qualitative terms whereas quantitative precise recommendations cannot be explored. This is the main problem we have to face for the future. In order to consider networks systematically in a new generation of epidemic models, we need much more fine-grained data on social interactions as well as much more massive computational power capacities.

It is probable that these deficiencies could explain why SIR models without realistic networks structures are still so dominant. In addition to a powerful, parsimonious, and well-theorized mathematical formalism (i.e. differential equations), they are less data-demanding. Indeed, many analysts have raised concern on the limitations of modelling heterogeneous social behaviour and accurate structures of social interactions in this type of models at least since the H1N1 pandemic. Epstein (2009), for instance, already formulated a plea for data-driven algorithmic models, in particular agent-based computational models, and defended the technical possibility to have this type of models implemented at a global scale (Parker & Epstein, 2011). However, as forcefully pinpointed by Squazzoni et al. (2020), this type of models was in fact absent from the scientific and public debates on predictions and interventions concerning COVID-19. I suggest that this is primarily due to the absence of fine-gained empirical data on social networks at the country level, as well as to the technical difficulties related to run network-data-driven computational models at a sufficient scale.

This said, despite its dramatic effects on the short run, we should take the current crisis as an opportunity to reconsider the challenge of empirical data and research infrastructures. Scientific and political authorities are now aware that quantitative social sciences urgently need the equivalent of particle accelerators in physics or wheatear forecasting facilities in climate and environment studies. It is natural to consider that in order to advance in the understanding of the molecular structure of a new virus, we need massive rich data and tools with huge computational capabilities. However, it seems unlikely that we are ready to accept the equal importance of providing comparable long-run economic and institutional efforts to adequately model and simulate the diffusion of the same virus through human-to-human interactions, which are (at least) as complex as the infra-virus interactions between molecules and cells. No matter how unrealistic this idea may seem now, the current crisis clearly urges us to understand that without being able to model and simulate heterogeneous human behaviours embedded in complex (and also evolving) social networks at a country level, we simply cannot intervene on contagion dynamics more surgically than what presently suggested by dominant SIR models (even in their individual-based variants).

To advance in such direction, a minimal *agenda* should contain at least four chapters. I am not speaking here only from a specific French situation. Recent calls for a new "outbreak science" in the US (see Rivers et al., 2020), for instance, were based on similar diagnoses on

modelling, data and institutional deficiencies (*ibidem*, pp. 10–15), and led to similar proposals for the future (*ibidem*, pp. 17–20). I believe that the following suggestions reflect a larger problem, not only my main French observation window.

1. We should collect various forms of massive digital and mobility data systematically. These data have the potential to inform network-based model of epidemic spread at a certain level of resolution. For instance, by exploiting the frequency, the length and the reciprocity of phone calls, it is possible to infer the existence of strong and weak ties between actors (see, for instance, Onnela et al., 2007). By exploiting users' geographical location, phone calls help to show that the former type of ties is more likely to connect individuals who are geographically closer (Onnela et al., 2011). However, existing studies comparing network structures inferred from co-presence data and face-to-face interactions involving verbal exchanges in small groups, suggest that spatial co-presence may be a poor indicator of some features of real social networks that are relevant for disease diffusion. For instance, this is the case of actors with a large number of connections, i.e., *hubs* (see Génois & Barrat, 2018). In addition, because of privacy issues under (even more) lively debate in the current crisis context, this type of data is likely to be released and treated at a higher level of aggregation than the individual level. Thus, we may expect that digital data can help to identify the broad structure of long-range ties across more or less geographically closed social circles rather than providing detailed, and directly exploitable, information on the structure of actors' local social networks. This is probably what we should also expect from various forms of mobility data (see, for instance, Aleta et al., 2017).

2. Given massive digital data's limitations, despite their potential, this type of empirical information should not be our only tool to map social interactions. Especially for virus spread, person-to-person interactions is what really matters. Equally, if not more, intense efforts thus should be made in collecting empirical data on off-line social interactions at a country scale. In this respect, more traditional sampling methods should be reinforced, and adapted, to collect data on ego-centered networks. National statistical offices and various public and private poll institutes usually collect a huge amount of information on properties and behaviours of individuals, households and firms. However, at least in France, no comparable effort and public investments have been made on data collection concerning network data. This is a serious problem, in particular now that we start to have statistical methods that could be used to infer global network structures from samples of ego-centered networks (see Smith, 2012). This type of information could then be increased by finding ways to incorporate novel numerical devices, such as wearable wireless sensors to trace real duration and temporal evolution of face-to-face interactions (see Barrat et al., 2014), within more tradition sample design at a country level. We need to design procedures allowing such coupling at a very large scale (on this point, see Stopczynski et al., 2014).

3. In addition to reinforce our data collection procedures, we also need to continue to develop methods to clean and describe complex data structures on social interactions. A major issue here is how to deal with error measurement and omitted observations. This problem equally concerns novel numerical devices for high-granularity face-to-face interactions (Génois et al., 2015) and more traditional sampling approach for ego-centered networks (Smith, Moody, & Morgan, 2017). Considerable progresses have been made in both fields (see, respectively, Sapienza et al., 2018; Heckathorn & Cameron, 2017), but this clearly is a major line of research that should be pursued and reinforced even further.

4. Once properly collected, cleaned and described, fine-grained data on social interactions must then be introduced within methods that allow us to simulate the macroscopic dynamic

of the epidemic when the virus is transmitted through person-to-person interactions among individuals with heterogeneous adaptive behaviours. The judgements delivered by the French COVID-19 scientific committee as well as *scenario* analyses by Ferguson and his team repeatedly remarked that the systemic effect of the proposed interventions crucially depend on the way individuals respond to them. However, no model advising these interventions explicitly include actors' possible behavioural responses. Thus, we face a paradox here: we have our lives determined by interventions that are based on models ignoring *at the same time* the real structure of social interactions and the complexity of human behaviours when uncertainty and risks are both present. The same technique (i.e., agent-based computational models) that can incorporate empirical data on complex network structures can also easily integrate large scale survey and/or experimental data on the heuristics that actors develop under situations of uncertainty, risk or fear (for an interesting example on this actor-related part, see Badham & Gilbert, 2015). However, one of the main challenges is to develop technical infrastructures to make it possible to run simulations combining *simultaneously* real complex networks and actors' behaviours at a large scale. For instance, we can consider the micro-simulation platform described by Grefenstette et al. (2013) and imagine what the inclusion of network topologies and heterogeneous adaptive behaviours would require in computational terms.

Needless to say, a similar *agenda* requires pro-active scientific policies and extensive public investments. However, the costs of the current crisis will give politicians no excuse for ignoring these challenges and the opinion of experts better prepared to face the next pandemic. It is now clear that the extent to which interventions on a virus spread can be targeted rather than generalized, directly depends on the granularity of data on social interactions and computational power of the required modelling tools. Obviously, making part of our relational life accessible to data and research could be costly. Nevertheless, the choice for the future is now clear. If we do not accept such a relatively small privacy reduction, each of us will be exposed to much larger and painful freedom restrictions.

References

Adam, D. (2020). The Simulations Driving the World's Response to COVID-19. How Epidemiologists Rushed to Model the Coronavirus Pandemic. *Nature*, *580*, 316–318. https://doi.org/10.1038/d41586-020-01003-6

Ajelli, M., Gonçalves, B., Balcan, D., Colizza, V., Hu, H., Ramasco, J.J., Merler, S., & Vespignani, A. (2010). Comparing Large-Scale Computational Approaches to Epidemic Modeling: Agent-Based versus Structured Metapopulation Models. *BMC Infectious Diseases*, *10*(190). http://dx.doi.org/10.1186/1471-2334-10-190

Aleta, A., Hisi, A.N.S., Meloni, S., Poletto, C., Colizza, V., & Moreno, Y. (2017). Human Mobility Networks and Persistence of Rapidly Mutating Pathogens. *Royal Society Open Science*, *4*, 160914. http://dx.doi.org/10.1098/rsos.160914

Badham, J., & Gilbert, N. (2015). *TELL ME Design: Protective Behaviour During an Epidemic*. CRESS Working Paper 2015(2), University of Surrey.

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

Barabási, A-L., & Bonabeau, E. (2003). Scale-Free Networks. *Scientific American*, *288*(5), 60–69. http://dx.doi.org/10.1038/scientificamerican0503-60

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

Barrat, A., Cattuto, C., Tozzi, A.E., Vanhems, P., & Voirin, N. (2014). Measuring Contact Patterns with Wearable Sensors: Methods, Data Characteristics and Applications to Data-Driven Simulations of Infectious Diseases. *Clinical Microbiology and Infection*, *20*(1), 10–16. https://doi.org/10.1111/1469-0691.12472

Benzell, S., Collis, A., & Nicolaides, C. (2020). *Rationing Social Contact During the COVID-19 Pandemic: Transmission Risk and Social Benefits of US Locations*. OSF Preprints, Center for Open Science, April 18. https://doi.org/10.31219/osf.io/d64vm

Block, P., Hoffman, M., Raabe, I.J., Beam Dowd, J., Rahal, C., Kashyap, R., & Mills, M.C. (2020). *Social Network-Based Distancing Strategies to Flatten the Covid 19 Curve in a Post-Lockdown World*. Eprint article, April 15. Available at https://arxiv.org/abs/2004.07052

Brethouwer, J-T., van de Rijt, A., Lindelauf, R., Fokkink, R. (2020). *"Stay Nearby or Get Checked": A Covid-19 Lockdown Exit Strategy*. Eprint article, April 11. Available at https://arxiv.org/abs/2004.06891

Camacho, A., Kucharski, A.J., Lowe, R., Eggo, R.M., & Edmunds, W.J. (2019). Assessing the Performance of Real-Time Epidemic Forecasts: A Case Study of Ebola in the Western Area Region of Sierra Leone, 2014-2015. *PLoS Computational Biology*, *15*(2), e1006785. https://doi.org/10.1371/journal.pcbi.1006785

Cirillo, P., & Taleby, N.N. (2020). Tail Risk of Contagious Diseases. *Nature Physics* (forthcoming). Eprint article, April 18. Available at https://arxiv.org/abs/2004.08658

Di Domenico, L., Pullano, G., Sabbatini, C.E., Boëlle, P-Y., & Colizza, V. (2020). *Expected impact of lockdown in Île-de-France and possible exit strategies*. Report 9, April 12. Available at https://www.epicx-lab.com/covid-19.html

Duan, W., Fan, Z., Zhang, P., Guo, G., & Qiu, X. (2015). Mathematical and Computational Approaches to Epidemic Modeling: A Comprehensive Review. *Frontiers in Computational Science*, *9*(5), 806–826. https://doi.org/10.1007/s11704-014-3369-2

Edmonds, B., Le Page, C., Bithell, M., Chattoe-Brown, E., Grimm, V., Meyer, R., Montañola-Sales, C., Ormerod, P., Root, H., & Squazzoni, F. (2019). Different Modelling Purposes. *Journal of Artificial Societies and Social Simulation*, *22*(3), 1–6. https://doi.org/10.18564/jasss.3993

Enserink, M. & Kupferschmidt, K. (2020). Mathematics of Life and Death: How Disease Models Shape National Shutdowns and Other Pandemic Policies. *Science*, March 25. https://doi.org/10.1126/science.abb8814

Epstein, J.M. (2009). Modelling to Contain Pandemics. *Nature*, *460*(7256): 687. https://doi.org/10.1038/4606687a

Ferguson, N.M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bha- tia, S., Boonyasiri, A., Cucunubá, Z., Cuomo-Dannenburg, G., & Dighe, A. (2020). *Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand*. Report 9. Imperial College COVID-19 Response Team, London, March 16. https://doi.org/10.25561/77482

Flaxman, S., Mishra, S., Gandy, A., et al. (2020). *Estimating the Number of Infections and the Impact of Non-Pharmaceutical Interventions on COVID-19 in 11 European Countries*. Report, Imperial College London. Available at: https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-13-europe-npi-impact/

Génois, M., & Barrat, A. (2018). Can Co-Location Be Used as a Proxy for Face-to-Face Contacts?. *EPJ Data Science*, *7*(11). https://doi.org/10.1140/epjds/s13688-018-0140-1

Génois, M., Vestergaard, C.L., Cattuto, C., & Barrat, A. (2015). Compensating for Population Sampling in Simulations of Epidemic Spread on Temporal Contact Networks. *Nature Communications*, *6*, 8860. https://doi.org/10.1038/ncomms9860

Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology*, *78*(6), 1360–1380. https://doi.org/10.1086/225469

Grefenstette, J.J., Brown, S.T., Rosenfeld, R. et al. (2013). FRED (A Framework for Reconstructing Epidemic Dynamics): An Open-Source Software System for Modeling Infectious Diseases and Control Strategies Using Census-Based Populations. *BMC Public Health*, *13*(940). https://doi.org/10.1186/1471-2458-13-940

Heckathorn, D.D., & Cameron, C.J. (2017). Network Sampling: From Snowball and Multiplicity to Respondent-Driven Sampling. *Annual Review of Sociology*, *43*, 101–119. https://doi.org/10.1146/annurev-soc-060116-053556

Jackson, M.O., & Lopez-Pintado, D. (2013). Diffusion and Contagion in Networks with Heterogeneous Agents and Homophily. *Network Science*, *1*(1), 49–67. https://doi.org/10.1017/nws.2012.7

James, A., Pitchford, J.W., & Plank, M.J. (2007). An Event-Based Model of Superspreading in Epidemics. *Proc. Biol. Sci. Proceedings B*, *274*(1610), 741–747. https://doi.org/10.1098/rspb.2006.0219

Keeling, M.J., & Eames, K.T. (2005). Networks and Epidemic Models. *Journal of the Royal Society Interface*, *2*(4), 295–307. https://doi.org/10.1098/rsif.2005.0051

Keeling, M.J., & Rohani, P. (2008). *Modeling Infectious Diseases in Humans and Animals*. Princeton, NJ: Princeton University Press.

Kirman, A.P. (1992). Whom or What Does the Representative Individual Represent?. *Journal of Economic Perspectives*, *6*(2), 117–136. https://doi.org/10.1257/jep.6.2.117

Kissler, S., Tedijanto, C., Lipsitch, M., & Grad, Y.H. (2020). Social Distancing Strategies for Curbing the COVID-19 Epidemic 2020. Pre-print article available at: https://www.medrxiv.org/content/10.1101/2020.03.22.20041079v1.full.pdf https://doi.org/10.1101/2020.03.22.20041079

Lancichinetti, A., Kivelä, M., Saramäki, J., & Fortunato, S. (2010). Characterizing the Community Structure of Complex Networks. *PLOS ONE*, *5*(8), e11976. https://doi.org/10.1371/journal.pone.0011976

Landler, M., & Castle, S. (2020). Behind the Virus Report That Jarred the U.S. and the U.K. to Action. *The New York Times*, March 17.

Le Monde (2020). Sur la piste du patient "zéro". April 10, 20–21.

Liljeros, F., Edling, C.R., Nunes Amaral, L.A., Stanley, H.E., & Aberg, Y. (2001). The Web of Human Sexual Contacts. *Nature*, *411*(6840), 907–908. https://doi.org/10.1038/35082140

Manzo, G. (2020). Il faut intégrer la structure des interactions sociales dans les modèles de diffusion de l'épidémie. *Le Monde*, April 14. https://www.lemonde.fr/idees/article/2020/04/14/il-faut-integrer-la-structure-des-interactions-sociales-dans-les-modeles-de-diffusion-de-l-epidemie_6036502_3232.html

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*(3666). https://doi.org/10.1038/s41598-020-60239-4

Moody, J., & Benton, R.A. (2016). Interdependent Effects of Cohesion and Concurrency for Epidemic Potential. *Annals of Epidemiology*, *26*(4), 241-248. https://doi.org/10.1016/j.annepidem.2016.02.011

Morris, M., & Kretzschmar, M. (1997). Concurrent Partnerships and the Spread of HIV. *AIDS*, *11*(5), 641–648. https://doi.org/10.1097/00002030-199705000-00012

Newman, M.E.J. (2003). The Structure and Function of Complex Networks. *SIAM Review*, *45*(2), 167–256. https://doi.org/10.1137/S003614450342480

Onnela, J.-P., Arbesman, S., Gonzalez, M.C., Barabási, A.L. & Christakis, N.A. (2011). Geographic Constraints on Social Network Groups. *PLoS ONE*, *6*(4), e16939. https://doi.org/10.1371/journal.pone.0016939

Onnela J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertesz, J., & Barabási, A.-L. (2007). Structure and Tie Strengths in Mobile Communication Networks. *PNAS*, *104*(18), 7332–7336. https://doi.org/10.1073/pnas.0610245104

Parker, J., & Epstein, J.M. (2011). A Distributed Platform for Global-Scale Agent-Based Models of Disease Transmission. *ACM Transactions on Modeling and Computer Simulation*, *22*(1), 2. https://doi.org/10.1145/2043635.2043637

Rivers, C., Martin, E., Meyer, D., Inglesby, T.V., & Cicero, A.J. (2020). *Modernizing and Expanding Outbreak Science to Support Better Decision Making During Public Health Crises: Lessons for COVID-19 and Beyond*. Report, The Johns Hopkins Center for Health Security. Available at https://www.centerforhealthsecurity.org/our-work/pubs_archive/pubs-pdfs/2020/200324-outbreak-science.pdf

Rocha, L.E.C., Liljeros, F., & Holme, P. (2011). Simulated Epidemics in an Empirical Spatiotemporal Network of 50,185 Sexual Contacts. *PLoS Computational Biology*, *7*(3), e1001109. https://doi.org/10.1371/journal.pcbi.1001109

Sapienza, A., Barrat, A., Cattuto, C., & Gauvin, L. (2018). Estimating the Outcome of Spreading Processes on Networks with Incomplete Information: A Dimensionality Reduction Approach. *Physical Review E*, *98*(1), 012317. https://doi.org/10.1103/PhysRevE.98.012317

Siegenfeld, A.F., & Bar-Yam, Y. (2020). Eliminating COVID-19: A Community-Based Analysis Eprint article, March 23. Available at https://arxiv.org/abs/2003.10086

Smith, J.A. (2012). Macrostructure from Microstructure: Generating Whole Systems from Ego Networks. *Sociological Methodology*, *42*(1), 155–205. https://doi.org/10.1177/0081175012455628

Smith, J.A., Moody, J., & Morgan, J.H. (2017). Network Sampling Coverage II: The Effect of Non-Random Missing Data on Network Measurement. *Social Networks*, *48*, 78–99. https://doi.org/10.1016/j.socnet.2016.04.005

Squazzoni, F., Polhill, J.G., Edmonds, B., Ahrweiler, P., Antosz, P., Scholz, G., Chappin, É., Borit, M., Verhagen, H., Giardini, F., & Gilbert, N. (2020). Computational Models That Matter During a Global Pandemic Outbreak: A Call to Action. *Journal of Artificial Societies and Social Simulation*, *23*(2), 10. https://doi.org/10.18564/jasss.4298

Stopczynski, A., Sekara, V., Sapiezynski, P., Cuttone, A., Madsen, M.M., Larsen, J.E., et al. (2014). Measuring Large-Scale Social Networks with High Resolution. *PLoS ONE*, *9*(4), e95978. https://doi.org/10.1371/journal.pone.0095978

Walker, P.G.T., Whittaker, C., Watson, O. et al. (2020). *The Global Impact of COVID-19 and Strategies for Mitigation and Suppression*. Report, Imperial College London. https://doi.org/10.25561/77735

Watts, D.J., & Strogatz, S.H. (1998). Collective Dynamics of "Small-World" Networks. *Nature*, *393*, 440–442. https://doi.org/10.1038/30918

Gianluca Manzo: Gemass, CNRS, Sorbonne University (France)
https://orcid.org/0000-0002-4905-2878
gianluca.manzo@cnrs.fr; https://www.gemass.org/manzo/
Gianluca Manzo is a research fellow in Sociology at the National Center for Scientific Research (CNRS) in Paris. He applies computational models and social network analysis to the study of social stratification and diffusion dynamics. He edited *Analytical Sociology: Actions and Network* (Wiley, 2014). His most recent (forthcoming) solo-authored book is *Agent-based Models and Causal Inference* (Wiley, 2021).